The Exploration-Exploitation Trade-off in the Newsvendor Problem*

Omar Besbes, Juan M. Chaneton, Ciamac C. Moallemi
Graduate School of Business
Columbia University
Henry R. Kravis Hall
665 West 130th Street
New York, NY 10027
ob2105@columbia.edu, jchaneton15@gsb.columbia.edu, ciamac@gsb.columbia.edu

This version: January 20, 2022

Abstract

When an inventory manager attempts to construct probabilistic models of demand based on past data, demand samples are almost never available: only sales data can be used. This demand censoring introduces an exploration-exploitation trade-off as the ordering decisions impact the information collected. Much of the literature has sought to understand how operational decisions should be modified to incorporate this trade-off. We ask an even more basic question: when does the exploration-exploitation trade-off matter? To what extent should one deviate from a myopic policy that takes the optimal decision for the current period without consideration for future periods? We analyze these questions in the context of a well-studied stationary multi-period newsvendor problem in which the decision-maker starts with a prior on parameters characterizing the demand distribution. We show that, under very general conditions in both perishable and non-perishable settings, the myopic policy will almost surely learn the optimal decision one would have taken with knowledge of the unknown parameters. Furthermore, in the perishable setting, we analyze finite time performance for a broad family of tractable cases. Through a combination of analytical parametric bounds and exhaustive exact analysis, we show that the myopic optimality gap is negligible for many practical instances.

Keywords: demand censoring, inventory management, dynamic learning, newsvendor, myopic policy, exploration-exploitation trade-off, Bayesian analysis.

*The third author was partially supported by NSF Grant CMMI-1235023.

1 Introduction

In many business settings, inventory management can be challenging due to uncertainty about underlying demand. A decision maker must build probabilistic models for future demand which are estimated based on past data. In practical operational situations, however, demand information is rarely accessible. Often, only access to some form of distorted demand is possible, and this distortion is usually impacted by prior operational decisions. An important example is *demand censoring*: in many retail environments, demand is only observable up to the level at which it can be fulfilled. As a result, firms most often only have access to sales data as opposed to the true, underlying demand. In such settings, inventory decisions affect not only immediate profits, but also the amount of information collected for use in future periods.

Such a problem is a special instance of a more general class of dynamic learning problems, in which a decision maker faces the following question: should he optimize his decision using the current information, or should he make a sub-optimal choice with regard to instantaneous rewards that might improve his future knowledge of the system, and lead to better performance in the future? This is known as the *exploration-exploitation* trade-off. The key concept is that decisions not only affect rewards, but also information that is collected to inform future decisions.

Accounting for the informational impact of actions can considerably complicate decision making, and much of the literature addressing the exploration-exploitation trade-off seeks to obtain insights in understanding how this trade-off should impact operational decisions. In the present paper, we focus on a particular setting that has received significant attention in the literature — a multi-period newsvendor problem with demand censoring — and ask an even more basic question: does the exploration-exploitation trade-off matter in the first place, i.e., to what extent does one need to account for future rewards and deviate from myopic decisions? Quite remarkably, we find that myopic decision-making is "sufficient" for all practical purposes in a wide range of settings: there is almost no value of accounting for the impact of ordering decisions on information collection. We also articulate circumstances where this conclusion does not apply, i.e., the conjunction of conditions that are required for the value of accounting for the exploration-exploitation trade-off to be non-trivial.

We consider a multi-period newsvendor problem, where a retailer sells perishable items[1] and has to decide how many units to order for each time period. Unmet demand is censored. We assume that the retailer does not have complete knowledge of a vector of parameters characterizing the underlying distribution, but rather learns about those as sales realizations are observed. We consider a Bayesian formulation where one may summarize the state of the system through the

[1] We anchor our analysis around this case but also extend some results to the non-perishable case.

current belief over the parameters characterizing the demand distribution.

As we discuss in the literature review below, it is clear from existing studies that ignoring the exploration-exploitation trade-off is suboptimal and that one should order more than the myopic solution under fairly general conditions. This feature of optimal policies is known as "information stalking" (Lariviere and Porteus, 1999). However, although the literature finds that the *policy* of information stalking is optimal in general, little focus has been been given to quantitatively analyzing the *value* of information stalking. This is despite the fact that numerical experiments in earlier studies suggest that it may not be significant.[2]

Motivated by this, in the present paper, we analyze a fundamental question that arises in the presence of demand censoring. What is the value of optimally accounting for the exploration-exploitation trade-off?

Main contributions. The main contribution of the present paper is to articulate a case for the viability and appeal of the myopic policy in Bayesian inventory problems. To do so, we proceed in two main steps.

1. In a first step, we analyze the problem under very general conditions with regard to the unknown parameters, the prior, and the demand distribution. In particular, the prior-demand pair need not admit any conjugate structure. In this general setting, we establish that the myopic policy learns almost surely in the long run the optimal decision one would have taken with full knowledge of the unknown parameters. In particular, this result shows that the myopic policy is always a viable alternative if the time horizon is large. The approach taken to prove the result is novel and different than typical approaches in parametric learning problems where a basic building block is to establish that one can learn the true underlying parameters. Our approach does not rely on learning the true parameters directly but on the structure of the myopic policy in the specific context of the newsvendor problem, and the analysis of martingales in Banach spaces. We further establish that the result continues to hold for the case of non-perishable inventory.

2. In a second step, we focus on the finite performance of the myopic policy. To do so, we narrow down the scope to one of the few known tractable versions of the censored newsvendor problem, Weibull demand with one unknown parameter, and a prior with a gamma distribution. We evaluate the relative performance of the myopic policy compared to the optimal policy, a

[2]See, for example, the experiments of Ding et al. (2002) and Chen (2010); the latter considers a case with non-perishable inventory and the performance of the myopic policy is very rarely below 99% of the optimal. The numerical results in these papers show that the myopic policy can, in the specific instances considered, be near optimal. However, the generality of this fact is not established exhaustively — it is not the focus of those papers. The broad near optimality of myopic policies appears to be fundamentally unacknowledged in the literature.

quantity we refer to as the *myopic optimality gap*.

We develop a parametric upper bound on the *myopic optimality gap* through a recursive argument. The upper bound captures the dependence on the problem parameters such as the time horizon, the service rate, the uncertainty in the prior as well as the shape of the demand distribution. We show that the upper bound becomes negligible for any extreme cases of the input parameters, and characterize the exact rate of convergence to zero. In particular, we study the bound as the time horizon increases to infinity, as the service level increases to one or decreases to zero, and as the parameter uncertainty decreases. Quite notably, even when the inventory manager selects a low service level and demand is very often censored, the myopic optimality gap still vanishes to zero.

We then evaluate the myopic optimality gap in an *exact* manner for a grid of **all** input parameters: time horizon, parameter uncertainty, service level, Weibull parameter. We show that it is negligible for most parameter combinations, often on the order of a fraction of a percentage point. The gap may only be significant when the problem primitives satisfy at the same time three very specific conditions: (i) high demand distribution uncertainty, (ii) low noise in demand realizations and (iii) high myopic censoring probability. If any of these conditions is violated, *there is almost no value in accounting for the exploration-exploitation trade-off*.

In addition to the above, we note that implementing myopic policies is always relatively straightforward while the computation of optimal policies may be intractable. In other words, the findings of this paper make a case that the myopic policy should not only be a computationally attractive but may also be a viable candidate heuristic in cases where the optimal policy is intractable.

By analyzing the evolution of ordering decisions and costs on simulated trajectories when the optimal policy is tractable, we develop intuition for the near optimality of the myopic policy. We find that it does not arise from "flatness" in the cost-to-go function of the problem. Indeed, we observe single-period costs vary substantially for both the optimal and myopic policies as they learn the demand. Instead, the similarity in performance stems from the fact that in most cases, both policies learn at approximately the same rate. The exception to this arises in the fragile confluence of conditions described above, where it is both possible to quickly learn demand and necessary (from a cost perspective) to do so; in these cases the optimality gap is non-trivial.

The framework we use can also be leveraged to analyze other questions. For example, we show how one may leverage the bounds we develop to bound the cost of censoring, the difference when operating in the presence of demand censoring versus when operating while observing all demand samples. Furthermore, in an exact analysis of the cost of censoring akin to the one conducted for the myopic optimality gap, we find that it is always significantly higher than the myopic optimality gap. This suggests that effort to improve information collection about lost sales may be more valuable than analytic or computational effort to pin down the optimal dynamic policy in the presence of censoring.

Related literature. A Bayesian approach to inventory management with observable demand was first proposed by Scarf (1959), leading to a formulation as a Markov Decision Process (MDP), with the state given by the current knowledge about the demand distribution. In such a case, decisions do not affect collection of information, and there is no exploration-exploitation trade-off; the main issue is to deal with the dimensionality of the state space. In its most general formulation, the problem has an infinite dimensional state; it can, however, be written as a finite dimensional MDP if the demand distribution is chosen from a parametric conjugate family. See also the work of Azoury (1985) for another early reference on the topic. Lovejoy (1990), in the absence of demand censoring, but when inventory is non-perishable argued for the near-optimality of a myopic policy. In that case, the policy is myopic with regard to the anticipation of stock and information evolution in future periods. In particular, there is no notion of exploration-exploitation trade-off in the class of problems considered, as decisions do not affect the information that is collected. In our work, the main focus is on demand censoring and the exploration-exploitation trade-off it induces. In particular, a myopic policy in our setting is one that does not anticipate the implications of demand censoring for information collection and its impact on future performance.

When demand observations are censored, decisions and information collection are coupled, and the problem becomes more challenging. Most of the literature on inventory control with Bayesian learning has focused on understanding the censored case. The main difficulty is that Bayesian updates lose their conjugate property in the presence of censoring for most distribution families, and hence the problem becomes intractable. Many studies have therefore focused on showing structural properties of the optimal policies. Ding et al. (2002)[3] study the censored newsvendor problem for general distributions and show that, under certain conditions, the optimal order quantity is greater than the quantity one would order to minimize current single period cost, also referred to as the myopic quantity. This result was then generalized by Bensoussan, Çakanyildirim and Sethi (2009) for arbitrary parametric distributions. The main point is that the exploration-exploitation trade-off leads to collecting more information for future periods. This is, broadly speaking, as far as the general case has been characterized.

There exists, however one well known family of distributions that preserves the conjugate property even under censoring. First introduced by Braden and Freimer (1991), the *newsvendor* dis-

[3] See also the complementary note of Lu et al. (2005*b*).

tributions have been frequently used as a benchmark to study the problem in a more tractable fashion. Lariviere and Porteus (1999) study the non-perishable inventory problem with newsvendor demands, and they find that, if demands follow a *Weibull* distribution — a particular case of a newsvendor distribution — a state space reduction technique can be applied and the problem can be solved by backward induction. Bisi et al. (2011) analyze the same case for perishable and non perishable inventory and they develop a specific recursive formula for the optimal solution in the perishable case. They also show that the Weibull is the only case in which such space reduction can be applied.

The effects of demand censoring have been studied in a wide range of environments, focusing on different aspects of the problem. In particular, many extensions of the base newsvendor case have been analyzed recently. Dai and Jerath (2013) study the impact of inventory restrictions and demand censoring in sales force compensation contracts, and Heese and Swaminathan (2010) analyze a similar problem when the firm has complete control over the sales effort. Chen and Han (2013) analyze the effects of learning and censoring on the supply side. Jain et al. (2013) study the newsvendor problem with censored demands when sales transaction timing is also available; see also Caro and Gallien (2010) for the use of such timing data. Mersereau (2013) focuses on the conjunction of demand censoring with inventory record inaccuracies. Most of these papers rely on specific model assumptions in order to circumvent the complexity of the analysis. From the computational point of view, heuristics have been developed in the literature to overcome the intractability of the problem for general prior/demand combinations. Chen (2010) and Lu et al. (2005*a*) propose heuristics and provide bounds on the optimal quantity.

All of the above papers focus on modeling uncertainty in a Bayesian framework. There is also a stream of literature that analyzes the design of policies under different informational assumptions, when there is no parameterization of the demand distribution. In such cases, performance is measured through regret, e.g., the gap between the performance of a policy and that of an oracle with knowledge of the demand distribution. Kunnumkal and Topaloglu (2008) and Huh and Rusmevichientong (2009) are example of such approaches. In that line of work, Besbes and Muharremoglu (2013) is related in spirit to the present study as it measures the impact of demand censoring on performance. However, it does so under different informational assumptions (non-parametric) and through the lens of minimax regret, and only measures performance of policies through the growth rate of regret as opposed to finite time analysis. Under such assumptions, there is no universal notion of an exploitation-only policy (as such a policy would need to be defined with respect to an estimation method) and hence no exact metric to quantify universally and precisely the exploration-exploitation trade-off. In the present Bayesian setting, the decision-maker has a prior on the demand distribution, and the exploitation-only policy is uniquely defined, allowing one to exactly quantify the exploitation-exploration trade-off. In turn, we are able to quantify the myopic optimality gap for arbitrary time horizons through fine parametric bounds that account for the time horizon but also for the specific informational structure operating regime (across different probabilities of being censored).

At a high level, the paper is also related to Gaur et al. (2015) who show that stockouts have a limited impact on the inference process of a choice model in textbook retailing. Finally, we refer to the recent review of Chen and Mersereau (2013) for a broader overview of the literature on demand censoring.

Our paper relates to studies that have highlighted classes of learning problems where myopic policies can perform well. In the context of dynamic pricing with unknown demand parameters, Besbes and Zeevi (2011) and Broder and Rusmevichientong (2012) highlight the "well-separated" case in which a myopic policy is near-optimal as no incomplete learning may occur. Indeed, at any price the possible demand curves do not intersect and as a result all prices are informative regarding the underlying true parameters. In related work, den Boer and Zwart (2015) show that even when the demand curves are not well-separated, then if there are capacity constraints, then variation in prices will be induced naturally through those, and a myopic policy will be near-optimal. The three studies above operate in a frequentist framework and analyze the regret of a myopic policy (compared to an optimal policy with knowledge of the true demand curve) through its growth rate with the time horizon. All studies rely on showing that the parameter estimates are suitably close to the true parameters, and in turn the regret of a myopic policy is small. In the present paper, we operate in a Bayesian framework, so the decision-maker has more information at hand. Furthermore, our general result regarding the almost sure convergence of myopic decisions does not rely on a "well-separated" assumption as our approach is different. We actually do not even track the parameter estimates and rely on a martingale argument to establish that convergence to the correct decisions needs to take place. In particular, it could be that the demand distributions are not well-separated but in a newsvendor problem, this is not critical as one, roughly speaking, only needs to learn a particular quantile. Finally, given a tractable family of problems (Weibull-Gamma), we are also able to evaluate numerically the worst case myopic optimality gap for the full parameter space (which is not possible in the frequentist studies above as there is no handle on the optimal policy). This highlights the quality of the myopic heuristic not only for long time horizons but for any time horizon.

The rest of this paper is organized as follows. In Section 2 we define the model and problem primitives. In Section 3, we conduct an asymptotic analysis of the decisions induced by the myopic policy. In Section 4, we turn to a finite time analysis, restricting attention to a tractable family of prior-demand distributions. In Section 5 we develop parametric upper bounds on the myopic optimality gap and Section 6 is then devoted to an exact analysis. Section 7 presents an analysis of the cost of censoring. Section 8 concludes. All proofs are collected in the Online Supplement to this paper.

2 Problem Formulation

We consider a multi-period newsvendor problem defined over a probability space $(\Omega, \mathcal{F}, P)$, consisting of a discrete-time sequence of ordering periods indexed by $t = 1, 2, \ldots$. We assume zero lead-time. At the beginning of each period, the retailer may order units, demand is then realized and is fulfilled to the extent possible. If the decision maker runs out of stock, demand beyond the initial inventory level is lost. On the other hand, we assume that inventory is perishable, i.e., if items remain, these are discarded.[4]

The probability measure P is defined as follows. We assume that the demands are independent and identically distributed, belonging to a family of distributions parameterized by a vector of unknown parameters θ, which takes values in the parameter set Θ. Given vector θ, the demand distribution has a probability density function denoted by $f(\cdot|\theta)$ and a cumulative distribution function denoted by $F(\cdot|\theta)$. Following a Bayesian approach, we assume that the decision maker has a prior distribution over θ.

The retailer incurs a cost $h > 0$ for each unit of unsold product and a penalty cost $p > 0$ for each unit of unmet demand. In other words, the single period cost given stocking decision $y \geq 0$ and realized demand $D \geq 0$ is

$$L(y, D) := h(y - D)^+ + p(D - y)^+.$$

Let $r := p/(h+p)$ denote the *critical ratio*, also sometimes referred to as the *service level* . Without loss of generality, we assume that $h = 1$ and parameterize the cost structure through $r \in (0, 1)$, i.e.,

$$L(y, D) := (y - D)^+ + \frac{r}{1 - r}(D - y)^+.$$

The retailer can only observe sales, rather than full demand realizations. In period t, the observed sales are given by $D_t \wedge y_t := \min\{D_t, y_t\}$, where D_t is the demand realized and y_t the ordered quantity. We denote by $\mathcal{F}_t$ the filtration generated by the censored demand process, that is, for $t \geq 1$,

$$\mathcal{F}_t := \sigma\big(D_1 \wedge y_1, y_1, \ldots, D_t \wedge y_t, y_t\big).$$

[4]In Section 3.3, we explore the non-perishable case.

Note that the information collected about the demand up to period t is impacted by the past stocking decisions $y_1, y_2, \ldots, y_{t-1}$.

We assume that $\mathbb{E}[D_1] < \infty$, which implies that $\mathbb{E}[L(y, D_t)] < \infty$, for all $t = 1, \ldots$ and any $y \geq 0$. We denote by $\mathcal{P}^c$ the set of non-anticipatory policies with respect to the censored information system, that is, policies for which the decision y_t prescribed in period t is $\mathcal{F}_{t-1}$-measurable for all $t \geq 1$.

The objective is to minimize the cumulative expected cost over a finite time horizon of length T, and the optimal value is given by

$$V_T^* = \inf_{\mu \in \mathcal{P}^c} \sum_{t=1}^{T} \mathbb{E}^{\mu}\left[L(y_t, D_t)\right], \tag{1}$$

where the expectation is taken assuming that decisions are made according to the policy μ.

Since the items are perishable, decisions across periods are only coupled through the information collected about the unknown parameter θ of the demand distribution. Indeed, given knowledge of θ, the optimal decision is simply to minimize the single-period cost, i.e.,

$$y^i(\theta) := \underset{y \geq 0}{\operatorname{argmin}}\ \mathbb{E}\left[L(y, D_t) | \theta\right]. \tag{2}$$

We refer will to y^i as the *informed order quantity*. On the other hand, in the absence of full knowledge of the demand parameter θ, When deciding on an order quantity in period t, the decision maker has to balance the impact of this decision on the current single-period cost with the impact on future costs that stems from the information collection process. This leads to the *exploration-exploitation* trade-off.

Computing an optimal policy that solves the optimization problem (1) is intractable in general, leading to the need for heuristics. One of the simplest possible such heuristics that is also computationally appealing is the *myopic* policy that seeks to minimize the single period cost at each time period, given the past information at hand. In other words, the myopic policy focuses exclusively on exploitation, ignoring any exploration considerations. Formally, for each time period $t \geq 0$, the myopic policy prescribes the order size

$$y_t^m := \underset{y \geq 0}{\operatorname{argmin}}\ \mathbb{E}\left[L(y, D_t) | \mathcal{F}_{t-1}\right]. \tag{3}$$

As we see, both the informed and myopic policies admit similar structure, minimizing single period costs, however the latter does so only with past information observations rather than full knowledge of the demand distribution. In order to guarantee the existence and uniqueness of the informed and myopic policies, we will make the following technical assumption:

Assumption 1. *For P-almost every $\theta \in \Theta$, the cumulative demand distribution $F(\cdot|\theta)$ satisfies the following conditions:*

1. *$F(\cdot|\theta)$ is continuous over $\mathbb{R}_+$.*
2. *$F(\cdot|\theta)$ is strictly increasing over $\mathbb{R}_+$. That is, for all $0 \le x < y$, $F(x|\theta) < F(y|\theta)$.*

We will assume throughout that Assumption 1 holds, and we formalize the characterization and uniqueness of the informed and myopic decisions below.

Proposition 1. *For P-almost every θ, the informed order quantity $y^i(\theta)$ satisfying (2) exists and is unique, and further this order quantity uniquely satisfies*

$$P\left(D_t \le y^i(\theta)\middle|\theta\right) = r. \tag{4}$$

Further, for each $t \ge 1$, P-almost surely, the myopic order quantity y_t^m satisfying (3) exists and is unique, and further this order quantity uniquely satisfies

$$P\left(D_t \le y_t^m|\mathcal{F}_{t-1}\right) = r. \tag{5}$$

We note that the requirements of Assumption 1 are not necessary, however, and can be weakened. For example, Assumption 1 requires that the cumulative demand distribution be strictly increasing over all of $\mathbb{R}_+$. If instead the demand has support over a fixed subset of $\mathbb{R}_+$ for all θ and is strictly increasing on the subset, this is sufficient.

3 Asymptotic Analysis

Our goal is to establish the asymptotic, long run optimality (as $T \to \infty$) of the myopic policy. We will do so by comparison to the informed policy, which would be optimal given full knowledge of the demand distribution. Indeed, our ultimate goal will be to establish that, under general conditions, P-almost surely, $y_t^m \to y^i(\theta)$ as $t \to \infty$. In other words, we will establish that asymptotically the myopic policy learns and converges to the optimal order quantity given knowledge of the demand distribution.[5] Note that this does not necessarily imply that the myopic policy asymptotically learns the true underlying demand distribution (i.e., learns θ), but rather that it learns a statistic of the underlying demand distribution (the r-fractile) that is sufficient for optimal decision making.

[5] Related is the work of Bensoussan, Royal and Sethi (2009), who show the same convergence in the case of exponentially distributed demands with a Gamma prior. This special case is subsumed by our result.

We proceed in two steps. First, we will establish that in the absence of any additional assumptions, the myopic policy orders in a way that it is censored asymptotically at the same rate as the informed policy. Then, under mild regularity assumptions on the demand distribution, we will show that this implies that, asymptotically, the myopic order quantity converges to the informed order quantity. In this way, the myopic policy is asymptotically optimal in a very general setting.

3.1 Asymptotic Rate of Censoring

To start, observe that (4) implies that the informed policy is asymptotically censored on a fraction $1-r$ of the periods. Formally,

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}_{\{D_t \le y^i(\theta)\}} = r, \tag{6}$$

P-almost surely. This is an immediate consequence of the strong law of large numbers. The following result guarantees that this also holds for the myopic policy:

Theorem 1. *Under the myopic policy,*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}_{\{D_t \le y^m_t\}} = r, \tag{7}$$

P-almost surely.

Theorem 1 is a consequence of the *martingale* strong law of large numbers. In particular, because of (5), the process

$$W_t := \sum_{\tau=1}^{t} \left(\mathbb{I}_{\{D_\tau \le y^m_\tau\}} - r \right), \quad \forall\, t \ge 0, \tag{8}$$

is an $\mathcal{F}_t$-adapted zero mean martingale. This, in combination with the martingale strong law of large numbers, implies that, P-almost surely, $W_T/T \to 0$ as $T \to \infty$. This in turn yields (7).

3.2 Asymptotic Order Quantity

Theorem 1 establishes that, in the absence of any additional assumptions, the myopic policy is censored at the same rate as the optimal informed policy over the long run. While this is encouraging, it does not quite imply that the ordered quantity converges. In order to make further progress, we will need an additional technical assumption on the demand distribution.

We begin with some definitions. Define Σ to be the set of Lebesgue measurable subsets of the positive interval $\mathbb{R}_+$, and define $\mathcal{B}$ to be the Banach space of signed finite measures on $(\mathbb{R}_+, \Sigma)$, equipped with the total variation metric

$$\|\mu\|_{\mathsf{TV}} := \sup_{A\in\Sigma}\ |\mu(A)| + |\mu(A^c)|,$$

for $\mu \in \mathcal{B}$. Note that if $\mu, \nu \in \mathcal{B}$ are also probability measures, then

$$\|\mu\|_{\mathsf{TV}} = 1, \quad \|\mu - \nu\|_{\mathsf{TV}} = 2\sup_{A\in\Sigma}\ |\mu(A) - \nu(A)|.$$

Now, we can define the $\mathcal{B}$-valued random variable $\mu_\theta \colon \Omega \to \mathcal{B}$ corresponding to the (unknown) demand distribution by

$$\mu_{\theta(\omega)}(A) := \int_A dF\big(x|\theta(\omega)\big) = P\left(D_t \in A|\theta(\omega)\right),$$

for each sample path $\omega \in \Omega$ and set $A \in \Sigma$ (we will sometimes suppress the explicit dependence on ω). In other words, μ_θ is the random measure corresponding to the distribution of demand given θ.

We will make the following assumption for the rest of this section:

Assumption 2 (Strong P-Measurability)**.** *Assume that the random measure $\mu_\theta \colon \Omega \to \mathcal{B}$ is strongly P-measurable. That is, there exists a sequence of functions (f_n), for $n \geq 1$, with each $f_n \colon \Omega \to \mathcal{B}$ taking the form*

$$f_n(\omega) = \sum_{i=1}^{N_n} \mathbb{I}_{A_i^n}(\omega) b_i^n,$$

where each $b_i^n \in \mathcal{B}$ is a measure, and each $A_i^n \subset \Omega$ is a subset of the probability space, for each $1 \leq i \leq N_n$, and for almost every $\omega \in \Omega$,

$$\lim_{n\to\infty} \|f_n(\omega) - \mu_{\theta(\omega)}\|_{\mathsf{TV}} = 0.$$

In other words, μ_θ is the pointwise limit (in $\mathcal{B}$) of a set of measure-valued functions each taking finitely many values.

Strong P-measurability (see, e.g., Hytönen et al., 2016) is a technical condition which guarantees a mild degree of regularity of the demand distribution across realizations of θ. This assumption is very mild, for example it is satisfied for classical newsvendor distributions that will be defined in Section 4.1. We will defer further discussion of this assumption to Section B.2 of the Online Supplement.

We have the following result:

Theorem 2. *Under Assumption 2, P-almost surely,*

$$\lim_{t \to \infty} y_t^m = y^i(\theta).$$

Theorem 2 establishes that the myopic policy is asymptotically optimal in the sense that the myopic order quantities converge to that of the informed policy, i.e., the optimal order quantity with full knowledge of the demand distribution. This holds in a remarkably general setting. For example, under the present hypotheses, it is not clear that θ and thus the entire underlying demand distribution can be determined asymptotically. However, under the myopic policy, the r-fractile of this distribution can be determined, and this is a sufficient statistic for optimal decision making.

While the proof of Theorem 2 is provided in Section B.3 of the Online Supplement, we outline here the main steps and key ideas of this proof.

1. Consider the random measure μ_t corresponding to the posterior distribution of demand after observing sales in period t, i.e.,

$$\mu_t(A) := P\left(D_t \in A | \mathcal{F}_t\right),$$

for all $A \in \Sigma$. Under Assumption 2, this is a ($\mathcal{B}$-valued) bounded martingale, and a Banach space version of the martingale convergence theorem establishes that μ_t converges P-almost surely to a limiting distribution μ_∞ (which may be different than the true, underlying distribution μ_θ). This pointwise convergence occurs under the total variation metric for $\mathcal{B}$.

2. Since $\|\mu_t - \mu_\infty\|_{\mathsf{TV}} \to 0$ P-almost surely, it must be the case that the r-fractile of each μ_t converges to that of μ_∞. Therefore, under the characterization of the myopic order quantity from Proposition 1, it must be the case that y_t^m converges also.

3. Theorem 1 demonstrates that the myopic policy is censored a fraction $1 - r$ of the time. However, an application of the Glivenko-Cantelli theorem implies that the only fixed order quantity with this property is $y^i(\theta)$. Since y_t^m converges to a fixed order quantity, it must be then that $y_t^m \to y^i(\theta)$.

3.3 Non-Perishable Case

The asymptotic optimality result developed above also holds in the non-perishable case, where excess inventory carries over from one period to the next.

Here, we denote by $x_t \geq 0$ the inventory at the beginning of period $t \geq 1$. We denote by $y_t \geq 0$

the *order-up to*[6] quantity for period t, so that $z_t := x_t \vee y_t := \max(x_t, y_t)$ is the post-replenishment inventory. Given demand realization D_t, sales of $D_t \wedge z_t$ are observed, and the cost $L(z_t, D_t)$ is incurred. The initial inventory for the next period is given by

$$x_{t+1} = z_t - (D_t \wedge z_t) = (z_t - D_t)^+ := \max(z_t - D_t, 0).$$

We denote by

$$\mathcal{F}_t := \sigma\big(x_1, D_1 \wedge z_1, y_1, \ldots, D_t \wedge z_t, y_t\big)$$

the information filtration generated by the censored demand process. The objective we consider is to minimize the long-run average expected cost

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[L(z_t, D_t)\right].$$

Note that relative to the perishable problem formulation of Section 2, we have abused notation to re-define y_t as an order-up to quantity (rather than an order quantity) and $\mathcal{F}_t$ as a filtration in the presence of carry-over inventory.

In this setting, we define a myopic decision-maker to be unconcerned with the impact of decisions on future information or future inventory, and as a result it is clear that a myopic order-up to quantity can be defined according to (3), exactly as before. It is less obvious but also well-known that, when the demand parameter θ is known, ordering up to the quantity $y^i(\theta)$ given by (2) is optimal, just as in the perishable case. This is because, when the demand distribution is known, the demand realizations are i.i.d. and the constraint that $z_t \geq x_{t-1}$ is not binding.[7] In other words, the multi-period perishable newsvendor under known demand is completely equivalent to the non-perishable case (see, for example, the discussion in Huh and Rusmevichientong, 2009). Moreover, under Assumption 1, Proposition 1 continues to hold.

In order to analyze the myopic policy in the non-perishable setting, first observe that the following analog of Theorem 1 continues to hold in the non-perishable case:

Theorem 1NP. *In the non-perishable case, under the myopic policy, P-almost surely,*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}_{\{D_t \leq y_t^m\}} = r.$$

[6]Note that describing the inventory decision by an order quantity or an order-up to quantity are equivalent, in this section we will take the latter point of view.

[7]In the initial inventory is high, this constraint may bind. However, once $x_t \leq y^i(\theta)$ for some t, this will continue for all future periods, so in any event $y^i(\theta)$ is the optimal decision in the long run.

Theorem 1NP holds because the process W_t defined in (8) remains an $\mathcal{F}_t$-adapted martingale in the non-perishable case — the key here is that the information filtration is strictly larger than for the perishable case since $z_t \geq y_t^m$ under the myopic policy. Hence, the event that $D_t \leq y_t^m$ is $\mathcal{F}_t$-measurable. Given this observation, Theorem 1NP follows from the same arguments as in the proof of Theorem 1 and is omitted.

Note that Theorem 1NP guarantees that even in the non-perishable case, the myopic policy will be censored asymptotically exactly on a fraction $1 - r$ of the periods. Then, the analysis of Section 3.2 can be repeated, to obtain the following analog of Theorem 2 in the non-perishable case:

Theorem 2NP. *In the non-perishable case, under Assumption 2, P-almost surely,*

$$\lim_{t \to \infty} y_t^m = y^i(\theta).$$

In order words, in the non-perishable case also, the myopic order quantity asymptotically converges to the optimal informed order quantity. The proof of Theorem 2NP is omitted since it is exactly the same as that of Theorem 2.

4 Dynamic Programming Analysis

Our next goal is to analyze more finely the finite time performance of the myopic policy. As mentioned in the introduction, it has been shown under general conditions that it is optimal to "order more". That is, in a multi-period problem with censored observations, a decision maker will want to stock at a level higher than would be optimal to minimize the current single-period cost. Intuitively, it is desirable to explore by ordering more so as to reduce the likelihood of censoring and collect more information about the demand distribution for use in future periods. However, precisely pinning down the optimal order quantity is, in general, difficult. As a result, the exact form of what type of exploration ought to be conducted, and the value associated with it, are not known. To quantify the exploration-exploitation trade-off, we will isolate the value of exploring (i.e., ordering more) by comparing the performance of an optimal policy to that of the myopic policy which sequentially minimizes current single-period expected costs, fully ignoring the impact of decisions on information collection.

The cumulative cost of the myopic policy, formally defined in (3), is given by

$$V_T^m = \sum_{t=1}^{T} \mathbb{E}\left[L(y_t^m, D_t)\right].$$

By definition, $V_T^* \leq V_T^m$ and we are interested in qualifying the relative gap between V_T^m and V_T^*, that is, the relative additional cost incurred by ignoring the exploration-exploitation trade-off and applying a full exploitation policy. We refer to this as the *myopic optimality gap* (MOG), formally defined as

$$\mathsf{MOG} := \frac{V_T^m - V_T^*}{V_T^*}.$$

An upper bound on the MOG. The direct analysis of the MOG appears intractable in general. In the present section, we introduce an upper bound on the MOG that is more tractable. If demand were observable, the collection of information would be unaffected by the policy used by the decision maker. We denote by

$$\mathcal{F}_t^o := \sigma(D_1, y_1, \dots, D_t, y_t)$$

the filtration associated with the full observation process. Let $\mathcal{P}^o$ be the set of admissible policies for the full observation problem, that is, policies for which prescribed decisions at period t are $\mathcal{F}_{t-1}^o$-measurable. The optimal cumulative cost for the uncensored problem is given by

$$V_T^o = \inf_{\mu \in \mathcal{P}^o} \sum_{t=1}^{T} \mathbb{E}^\mu \left[L(y_t, D_t) \right].$$

In the uncensored case, the information collected about the demand distribution is independent of past decisions and there is no exploration-exploitation trade-off. Hence, in this case, the problem is easy to solve: the optimal decision at time t is to minimize the current single-period cost. Moreover, clearly $\mathcal{P}^c \subseteq \mathcal{P}^o$, and it follows that V_T^o is a lower bound on V_T^*, that is,

$$V_T^o \leq V_T^*. \tag{9}$$

We define the *myopic cost of censoring* (MCC) as

$$\mathsf{MCC} := \frac{V_T^m - V_T^o}{V_T^o}.$$

It can be interpreted as the relative increase in cost stemming from censoring, when a myopic policy is applied and we have

$$\mathsf{MOG} = \frac{V_T^m - V_T^*}{V_T^*} \leq \frac{V_T^m - V_T^o}{V_T^o} = \mathsf{MCC},$$

That is, the MCC provides an upper bound on the MOG.

Next, we restrict our analysis to a subset of prior-demand distributions pairs that are conjugate to evaluate the MOG and MCC. In Section 5, we focus on characterizing the upper bound, MCC, through upper and lower bounds. The upper bounds apply directly to the MOG. From an analytical perspective, this approach circumvents the necessity to determine the optimal solution to the dynamic program associated with the censored demand problem. Instead, our analysis in this section relies on the performance of myopic policies (in the censored and uncensored cases), which are more amenable to analysis. In Section 6, we conduct exact analysis on the MOG.

4.1 Gamma-Weibull Demand and Dynamic Programming Formulation

Newsvendor distributions. We now assume that a scalar parameter $\theta > 0$ is unknown. Before ordering in period t, the knowledge about θ can be summarized by the current prior distribution density[8] $\pi_{t-1}(\theta) := \pi(\theta|\mathcal{F}_{t-1})$. The current prior is updated at every period, following Bayes' rule. One general class of distributions that preserves the conjugate property even under censoring are the so-called *newsvendor distributions.* A random variable is said to belong to the newsvendor family if its cumulative distribution function is given by

$$F(z|\theta) := 1 - e^{-\theta d(z)}, \quad \text{for all } z > 0, \tag{10}$$

where $d\colon \mathbb{R}_{++} \to \mathbb{R}_{++}$ is a differentiable, nondecreasing, and unbounded function with $d(0) = 0$. The parameter space is defined to be $\Theta := \mathbb{R}_{++}$ and the prior distribution of θ is assumed to be a Gamma distribution with shape parameter $S \in \mathbb{R}_{++}$ and rate parameter $a \in \mathbb{R}_{++}$. Formally, the prior distribution on θ has a density given by

$$\pi(\theta|a, S) := \frac{S^a \theta^{a-1} e^{-S\theta}}{\Gamma(a)}, \quad \text{for all } \theta > 0.$$

It has been shown (Braden and Freimer, 1991) that this family of distributions preserves its structure even under censored observations, that is, if $\pi_t(\cdot)$ is a Gamma distribution, then so is the posterior $\pi_{t+1}(\cdot|y, \pi_t)$ with or without censoring. In particular, given a demand level z and order quantity y, the update rules in the censored case for Gamma distribution hyperparameters (a, S) are given by

$$a' = a + \mathbb{I}_{\{z<y\}}, \qquad S' = S + d(z \wedge y), \tag{11}$$

where $\mathbb{I}_{\{\cdot\}}$ denotes the indicator function and $\cdot \wedge \cdot$ denotes the minimum. Equation (11) offers a natural interpretation of the hyperparameters (a, S): the scale hyperparameter S accounts for the

[8]In order to keep our notation intuitive, we index prior beliefs using a forward numbering scheme (i.e., we index by the current time), while we index cost-to-go functions using a backward numbering scheme (i.e., we index by the remaining time horizon). The choice of indexing should be obvious by the context.

sum of (a function of) the *sales* observed in the system, while the shape hyperparameter a counts the number of fully observed demand realizations. The shape hyperparameter plays a central role in our analysis as it measures the level of information about the unknown demand parameter θ: the coefficient of variation of the demand parameter θ given hyperparameters (a, S) is

$$\mathrm{CV}(\theta|a, S) = 1/\sqrt{a}. \tag{12}$$

Hence, higher values of a indicate less uncertainty regarding the value of θ.

Assumptions. For tractability, our analysis will focus on *Weibull distributions*, i.e., newsvendor distributions with

$$d(z) := z^{\ell}, \quad \text{for } \ell > 0.$$

Note here that the parameter ℓ is assumed to be known to the decision-maker and the only uncertainty pertains to the parameter θ. For the Weibull family, the Bayesian update (11) is written as $a_{t+1} = a_t + \mathbb{I}_{\{z<y\}}$ and $S_{t+1} = S_t + (z \wedge y)^{\ell}$ and the *predictive distribution* (the distribution of demand conditional on the current belief (a, S) for the parameter θ) has density

$$m(z|a, S) = \frac{aS^a \ell z^{\ell-1}}{(S + z^{\ell})^{a+1}}, \tag{13}$$

and cumulative distribution

$$M(z|a, S) = 1 - \frac{S^a}{(S + z^{\ell})^a}, \tag{14}$$

for $z > 0$. The predictive distribution is integrable (i.e., the demand has finite expectation) if and only if $a\ell > 1$, and therefore we will assume that this condition is satisfied throughout the rest of the paper. Note that it suffices to verify that the initial prior distribution of θ satisfies $a\ell > 1$, since the update rule (11) then guarantees that this will continue to hold at all future times.

Finite dimensional dynamic programs. Given current hyperparameters (a, S) we denote by $V_T^o(a, S)$, $V_T^m(a, S)$, and $V_T^*(a, S)$ the optimal cost function under uncensored demand, the myopic cost function under censored demand, and the optimal cost function under censored demand, respectively. Similarly, we define $C(a, S)$ to be the optimal single-period cost given current hyperparameters (a, S), that is,

$$C(a, S) := \min_{y \geq 0} \; \mathbb{E}\big[L(y, D)\big]. \tag{15}$$

In equation (15), the expectation is taken with respect to the pair (D, θ), that follows a newsvendor distribution with parameters (a, S). Unless otherwise stated, this will be the convention for expectation terms in the remainder of the paper.

With this notation in mind, we can decompose the cost functions of interest using standard dynamic programming backward induction.[9] To begin, the optimal cost function under censored demand must satisfy the Bellman equation

$$V_T^*(a,S) = \min_{y\geq 0} \left\{\mathbb{E}\left[L(y,D)\right] + \mathbb{E}\left[\mathbb{I}_{\{D<y\}} V_{T-1}^*\big(a+1, S+D^\ell\big)\right] + P\left\{D\geq y\right\} V_{T-1}^*\big(a, S+y^\ell\big)\right\},$$

for all (a,S) and all $T\geq 1$, with the terminal condition $V_0^*(a,S)=0$, for all (a,S). Similarly, the myopic cost function under censored demand must satisfy

$$V_T^m(a,S) = C(a,S) + \mathbb{E}\left[\mathbb{I}_{\{D<y^m\}} V_{T-1}^m\big(a+1, S+D^\ell\big)\right] + P\left\{D\geq y^m\right\} V_{T-1}^m\big(a, S+(y^m)^\ell\big), \quad (16)$$

for all (a,S) and all $T\geq 1$, with the terminal condition $V_0^m(a,S)=0$, for all (a,S). In (16), the myopic decision y^m is defined to be a solution to (15), i.e., $y^m \in \operatorname*{argmin}_{y\geq 0}\ \mathbb{E}\big[L(y,D)\big]$. Note that y^m depends on (a,S), although we suppress this dependence in our notation. Finally, the optimal cost function when demand samples are observable must satisfy

$$V_T^o(a,S) = C(a,S) + \mathbb{E}\left[V_{T-1}^o\big(a+1, S+D^\ell\big)\right], \quad (17)$$

for all (a,S) and all $T\geq 1$, with the terminal condition $V_0^o(a,S)=0$, for all (a,S). Here, we have used the fact that the optimal policy in this case is myopic.

5 Parametric Upper Bounds and Structural Analysis of the MCC

We aim to develop an upper bound on $V_T^m(a,S) - V_T^o(a,S)$. Rewriting the recursion for the myopic policy in the censored case (16) in a way that parallels the recursion for the observable demand case (17), one obtains, for all (a,S) and $T\geq 1$,

$$V_T^m(a,S) = C(a,S) + \mathbb{E}\left[V_{T-1}^m\big(a+1, S+D^\ell\big)\right] + \Gamma_{T-1}(a,S), \quad (18)$$

where

$$\Gamma_{T-1}(a,S) := (1-r)V_{T-1}^m\big(a, S+(y^m)^\ell\big) - \mathbb{E}\left[V_{T-1}^m\big(a+1, S+D^\ell\big)\mathbb{I}_{\{D\geq y^m\}}\right].$$

Combining (17) and (18), one has that

$$V_T^m(a,S) - V_T^o(a,S) = \mathbb{E}_D\left[V_{T-1}^m(a+1, S+D^\ell) - V_{T-1}^o(a+1, S+D^\ell)\right] + \Gamma_{T-1}(a,S). \quad (19)$$

[9] The existence of an optimal policy in this setting can be shown by applying Proposition 3.4 of Bertsekas and Shreve (1978).

The correction term $\Gamma_{T-1}(a,S)$ may be interpreted as the additional cost incurred over the next period due to the presence of censoring.

As a first step towards bounding the performance difference $V_T^m(a,S) - V_T^o(a,S)$, we bound $\Gamma_{T-1}(a,S)$ as follows.

Lemma 1. *Suppose that the demand distribution is a newsvendor distribution. For all (a,S), and all $T \geq 1$,*

$$\Gamma_T(a,S) \leq \sum_{k=1}^{T} (1-r)^k \left\{ C\big(a, S(1-r)^{-k/a}\big) - C^o_{T-k+1}\big(a, S(1-r)^{-k/a}\big) \right\},$$

where $C^o_t(a,S)$ is the future expected single-period cost, t periods from now, when demands are uncensored, i.e.,

$$C^o_t(a,S) := \mathbb{E}\left[C\big(a+t, S + D_1^\ell + \ldots + D_t^\ell\big)\right]. \tag{20}$$

Notably, Lemma 1 offers a bound on $\Gamma_T(a,S)$ in which all terms depend only on future costs in the *uncensored* setting. We next use Lemma 1 recursively to provide an explicit bound on $V_T^m(a,S) - V_T^o(a,S)$ by exploiting the Weibull structure[10].

Theorem 3. *Suppose that the demand distribution is Weibull with parameter ℓ. Then, for any $\ell > 0$,$T \geq 1$,$r \in (0,1)$, $a > 0$, $S > 0$, with $a\ell > 1$,*

$$V_T^m(a,S) - V_T^o(a,S) \;\; \leq \;\; S^{1/\ell}\left[Q(a,r,l)\right]^{1/\ell} \, \frac{\lambda - \lambda^{T+1}}{1-\lambda} \left[\log\left(1 + \frac{T}{a - 1/\ell}\right) + \frac{1}{a - 1/\ell}\right], \tag{21}$$

where $\lambda := (1-r)^{1-\frac{1}{a\ell}}$ and $Q(a,r,\ell)$ is a function depending only on a, r and ℓ, such that $Q(a,r,l) = O(1/a)$ as $a \to \infty$, when r and ℓ are fixed.

The theorem provides a bound on the (absolute) myopic cost of censoring as a function of the problem parameters: the time horizon T, the "uncertainty" parameter a and the service level r. We first provide a high level overview of the proof of the result and then analyze in detail the parametric dependence.

The proof of the result is based on two steps. We first develop an intermediate bound that connects the (absolute) myopic cost of censoring to the difference between the observable demand cost and the expected total cost when the demand parameter θ is known: $V^o_{T-1}(a,S) - (T-1)C^o_\infty(a,S)$, where

$$C^o_\infty(a,S) := \mathbb{E}\left[\min_{y \geq 0} \; \mathbb{E}\left[L(y,D)\big|\theta\right]\right]$$

[10]In what follows, given functions $f(\cdot)$ and $g(\cdot) > 0$ we say that $f(x) = O\big(g(x)\big)$ as $x \to a$ if $\limsup_{x \to a} |f(x)|/g(x) < \infty$.

is the expected optimal single-period cost assuming θ is known.[11] The latter difference can be interpreted as the increase in cost stemming from the fact that θ is unknown, in an uncensored environment. In a second step, we bound the latter difference based on the problem parameters by analyzing information accumulation and its implications on performance in the observable demand case. A key feature of the bound is that it captures the dependence on the problem parameters: the time horizon T, the economic trade-offs captured by r, the prior information captured by a and the demand distribution shape, captured by ℓ. We analyze some of these dependencies in detail next.

Parametric dependence. We first start with the time horizon dependence. Note that if $T = 1$, information collection does not play any role, the myopic policy is optimal, and we have

$$\frac{V_T^m(a,S) - V_T^o(a,S)}{V_T^o(a,S)} = 0 \qquad \text{when} \qquad T = 1.$$

We now analyze the performance for general time horizons.

Proposition 2 (Time horizon dependence)**.** *Suppose that the demand distribution is Weibull with parameter ℓ. Then, for any $\ell > 0$, $r \in (0,1)$, $a > 0$, $S > 0$, with $a\ell > 1$,*

$$\frac{V_T^m(a,S) - V_T^o(a,S)}{V_T^o(a,S)} = O(T^{-1}\log T) \qquad \textit{as} \qquad T \to \infty.$$

The result captures the dependence of the MCC on the time horizon T. The fact that the MCC is equal to zero when $T = 1$ simply stems from the fact that, in a one period problem, the censored and uncensored problems are identical, and the optimal decision is myopic. For large time horizons, the MCC vanishes at rate $O(T^{-1}\log T)$. This reflects that: (i) Over time, in the censored environment, the optimal order will become closer to the myopic order (cf. Theorem 2), as in the uncensored environment. (ii) The cumulative effects of deviations from a myopic solution in the censored case can only affect performance minimally, not by more than $O(T^{-1}\log T)$ in relative terms. In particular, the MOG cannot be large in both the extreme cases when $T = 1$ and $T \to \infty$ and converges to zero rapidly.

Proposition 3 (Prior information dependence)**.** *Suppose that the demand distribution is Weibull with parameter ℓ. Then for any $\ell > 0$, $r \in (0,1)$, $S > 0$, $T \geq 1$,*

$$\frac{V_T^m(a,S) - V_T^o(a,S)}{V_T^o(a,S)} = O(1/a) \qquad \textit{as} \qquad a \to \infty.$$

[11] Comparing with the definition of $C_T^o(a,S)$ in (20), note that $C_\infty^o(a,S) = \lim_{T\to\infty} C_T^o(a,S)$.

In other words, the MCC converges to zero at a fast rate as $a \to \infty$. This captures the behavior of the MCC in a regime in which there is very little prior uncertainty about the demand distribution parameter θ. In this case, the presence of censoring does not affect performance significantly, since there is little additional information to be captured.

Proposition 4 (Service level dependence)**.** *Suppose that the demand distribution is Weibull. Then for any $\ell > 0, a > 0, S > 0, T \geq 1$ with $a\ell > 1$,*

$$\frac{V_T^m(a,S) - V_T^o(a,S)}{V_T^o(a,S)} = O\left((1-r)^{1-1/a\ell}\right) \qquad as \qquad r \to 1^-,$$

$$\frac{V_T^m(a,S) - V_T^o(a,S)}{V_T^o(a,S)} = O\left(r^{1/\ell}\right) \qquad as \qquad r \to 0^+,$$

The result gives a characterization of the asymptotic properties of the MCC as the service level r becomes close to 0 or 1, that is, as the holding cost becomes arbitrarily large with respect to the penalty cost and vice versa.

If the penalty cost is large (i.e., r is close to 1), the myopic order quantity will be high (corresponding the r-percentile of the predictive distribution) as the decision-maker attempts to mitigate the large penalties associated with stockouts. This implies that the myopic policy will often observe full demand realizations, and one expects that the presence of censoring should not impact performance significantly.

Note here that when r is close to zero, a newsvendor with knowledge of θ would incur optimal expected cost close to zero. Without the knowledge of θ, the myopic newsvendor will be very often censored and it is not possible *a priori* to characterize the magnitude of the MCC as this quantifies relative performance. Remarkably, however, when r is close to zero, and hence, when the myopic quantity leads the decision maker to be censored very often, the performance implications of being myopic are still very limited; the MCC shrinks to zero at a rate $O(r^{1/\ell})$. This stems from the fact that the holding cost is very high and every unit of unconsumed inventory becomes very expensive, making exploration (or "ordering more" than myopic) very expensive.

As a corollary of the results above, both the myopic optimality gap and the cost of censoring become negligible at the boundaries of the input parameters of the problem, namely when $T = 1$, $T \to \infty$, $a \to \infty$, $r \to 0$, and $r \to 1$. Intuitively, as $r \to 0$, and $r \to 1$, the newsvendor trade-off becomes weaker and weaker and the one period optimal ordering quantity converges to zero and infinity respectively. As $T \to \infty$ or $a \to \infty$, there is less and less uncertainty about the unknown parameter. In Section 6, we quantify more finely the MOG by using an exact analysis to compute it for various ranges of the parameters of the problem.

Remark. In Section D of the Online Supplement, we provide a lower bound on MCC that highlights that the upper bound presented in Theorem 3 above captures the appropriate parametric dependence.

6 Exact Analysis of MOG

6.1 Scalability

A notable feature of the problem when demand has a Weibull distribution is that the single period optimal cost possesses the so-called scalability property, namely that

$$C(a, S) = S^{1/\ell} C(a, 1), \quad \text{for all } a > 1/\ell,\ S > 0.$$

As observed by Azoury (1985) and Lariviere and Porteus (1999) this property can be extended to the optimal and full observation cost functions: for any $a > 1/\ell$, $S > 0$, $T \geq 1$

$$V_T^*(a, S) = S^{1/\ell} V_T^*(a, 1), \quad V_T^o(a, S) = S^{1/\ell} V_T^o(a, 1).$$

This separability allows the exact cost functions to be determined in the optimal and full observation cases given only knowledge of $V_T^*(a, 1)$ and $V_T^o(a, 1)$, respectively, for all values of $a > 1/\ell$. Further, exact recursions have been developed for these two quantities in the existing literature (Azoury, 1985; Lariviere and Porteus, 1999; Bisi et al., 2011), and we will use those to evaluate these costs.

A similar reasoning yields that the myopic cost function $V_T^m(a, S)$ also possesses the scalability property, and an exact recursion can be used to compute it, as summarized in the next result.

Proposition 5. *Suppose that demand distribution is Weibull. Then, for all $a > 1/\ell$, $S > 0$, $T \geq 1$,*

$$V_T^m(a, S) = S^{1/\ell} V_T^m(a, 1).$$

In addition, for all $a > 1/\ell$,

$$V_T^m(a, 1) = C(a, 1) + \frac{a\ell}{a\ell - 1} \left(1 - (1 - r)^{1 - \frac{1}{a\ell}}\right) V_{T-1}^m(a + 1, 1) + (1 - r)^{1 - \frac{1}{a\ell}} V_{T-1}^m(a, 1).$$

In the recursive equation provided by Proposition 5, $V_T^m(a, 1)$ can be computed exactly given $V_{T-1}^m(a+1, 1)$ and $V_{T-1}^m(a, 1)$. This implies that $V_T^m(a, 1)$ can be evaluated using backwards induction, starting with the boundary condition $V_1^m(a, 1) = C(a, 1)$. The value of $C(a, 1)$ itself has no closed form expression, but can be approximated to arbitrary accuracy by numerical integration for any a. This is analogous to the situation for the value function of an optimal policy in the censored and full observation cases.

Hence, using the scalability property, the cost functions $V_T^m(a,S)$, $V_T^*(a,S)$, and $V_T^o(a,S)$ can be computed numerically using simplified dynamic programming recursions that are exact up to errors from numerical integration. In particular, no discretization of the state space is necessary.

6.2 Parametric Setup

We are interested in analyzing the behavior of the MOG across a broad range of input parameters. In the case of Weibull demand, the input parameters are given by:

- the Weibull parameter ℓ;
- the shape parameter a of the Gamma prior distribution;
- the scale parameter S of the Gamma prior distribution;
- the time horizon T; and
- the service level r.

Given the scalability property discussed in Section 6.1, the value of the MOG does not depend on the scale parameter S, since it is a cost difference normalized relative to a baseline cost. We are therefore left with four parameters that summarize the three relevant dimensions of the problem: time (T), cost structure / service level (r), and the demand characteristics (a, l). We consider ranges for time and service level given by

$$1 \le T \le 100, \qquad r \in \{0.1, \dots, 0.9, 0.99\}.$$

The demand parameters (a, ℓ) are more difficult to directly interpret. In order to clarify their role, it is convenient to consider two measures of demand uncertainty.

- Conditional on the prior distribution hyperparameters (a, S), we measure the aggregate uncertainty in the next period demand (with predictive distribution (13)–(14)) through the coefficient of variation of demand, i.e., $\mathrm{CV}(D|a,S)$. Note that $\mathrm{CV}(D|a,S)$ is a function of (a, ℓ), but does not depend on the shape parameter S[12]. In particular, one may establish that

$$\mathrm{CV}(D|a,S) = \sqrt{\frac{\mathbb{E}_{a,S}[D^2]}{(\mathbb{E}_{a,S}[D])^2} - 1} = \sqrt{\frac{(2/\ell)\mathrm{Beta}(2/\ell, a-2/\ell)}{((1/\ell)\mathrm{Beta}(1/\ell, a-1/\ell))^2} - 1},$$

  where $\mathrm{Beta}(\cdot,\cdot)$ refers to the Beta function. In the exponential case ($\ell = 1$), for example, one has that $\mathrm{CV}(D|a,S) = \sqrt{a/(a-2)}$ for $a > 2$.

[12]Indeed, if the random variable D follows the predictive distribution (14) with parameters (a, S), then $S^{1/\ell}D$ follows (14) with parameters $(a, 1)$. Thus, changing the parameter S corresponds to rescaling the random variable, leaving the coefficient of variation unchanged.

- The aggregate uncertainty is in part driven by the fact that, even absent uncertainty regarding θ, the demand realizations themselves are random. This can be quantified by the coefficient of variation of demand given the parameter θ, i.e., $\mathrm{CV}(D|\theta)$. Here, demand is assumed to be distributed according to the distribution (10). Note that $CV(D|\theta)$ depends on ℓ but does not vary with θ. Indeed, if the random variable D follows the demand distribution (10) with parameter θ, then $\theta^{-1/\ell}D$ follows (10) with parameter $\theta = 1$. Thus, changing the parameter θ corresponds to a rescaling, and leaves the coefficient of variation unchanged. One may establish that

$$\mathrm{CV}(D|\theta) = \sqrt{\frac{\mathbb{E}[D^2|\theta]}{(\mathbb{E}[D|\theta])^2} - 1} = \sqrt{\frac{\Gamma(1+2/\ell)}{(\Gamma(1+1/\ell))^2} - 1},$$

  where $\Gamma(\cdot)$ is the Gamma function. In particular, in the exponential case ($\ell = 1$) one has that $\mathrm{CV}(D|\theta) = 1$ for all $\theta > 0$.

We define the *uncertainty ratio* (UR) as

$$\mathrm{UR}(a,\ell) := \frac{\mathrm{CV}(D|a,S)}{\mathrm{CV}(D|\theta)}.$$

$\mathrm{UR}(a,\ell)$ is always greater or equal than 1 and is the ratio of the overall, aggregate uncertainty in the next period demand, to the uncertainty that would remain if θ was perfectly known. From the above discussion, UR is a function of only (a,ℓ). For example, in the exponential case ($\ell = 1$), we have that $\mathrm{UR} = \sqrt{a/(a-2)}$ for $a > 2$.

We will parameterize demand uncertainty through input parameters (UR,ℓ). Note that we use UR rather than a because it is directly interpretable as the relative value of uncertainty arising from the fact that θ is unknown. For example, if $\mathrm{UR} = 3$, the demand uncertainty would be reduced by a factor of three if θ were known.

On the other hand, we interpret the Weibull parameter ℓ as measuring the uncertainty of demand realizations given knowledge of θ (Note that the parameter ℓ also affects the shape of the distribution). From (10), it is clear that larger values of ℓ lead to faster decaying tails for the demand distribution. Moreover, from the above discussion, $\mathrm{CV}(D|\theta)$ depends only on ℓ. If, for example, ℓ increases, the coefficient of variation of the Weibull distribution (independent of θ) decreases, and hence there is less variation in demand realizations. The uncertainty about θ, however, remains constant, cf. (12). Because θ can be learned via demand observations, the potential value of exploration can be significantly affected by changing ℓ.

6.3 Analysis of the Myopic Optimality Gap

First, we consider the behavior of the MOG, i.e., the relative sub-optimality of the myopic policy as compared to an optimal policy.

6.3.1 General Weibull case

We are interested in analyzing the MOG. To obtain a broader understanding of the behavior and magnitude of the MOG, we next evaluate it for different values of the uncertainty ratio UR and service level r. In Figure 1, for varying choices of (r, UR) and varying values of the shape parameter ℓ, we depict the worst-case myopic optimality gap $\mathsf{MOG}_{\mathrm{wc}}$ over time horizons $1 \leq T \leq 100$, i.e.,

$$\mathsf{MOG}_{\mathrm{wc}} := \max_{1 \leq T \leq 100} \frac{V_T^m(a,1) - V_T^*(a,1)}{V_T^*(a,1)}. \tag{22}$$

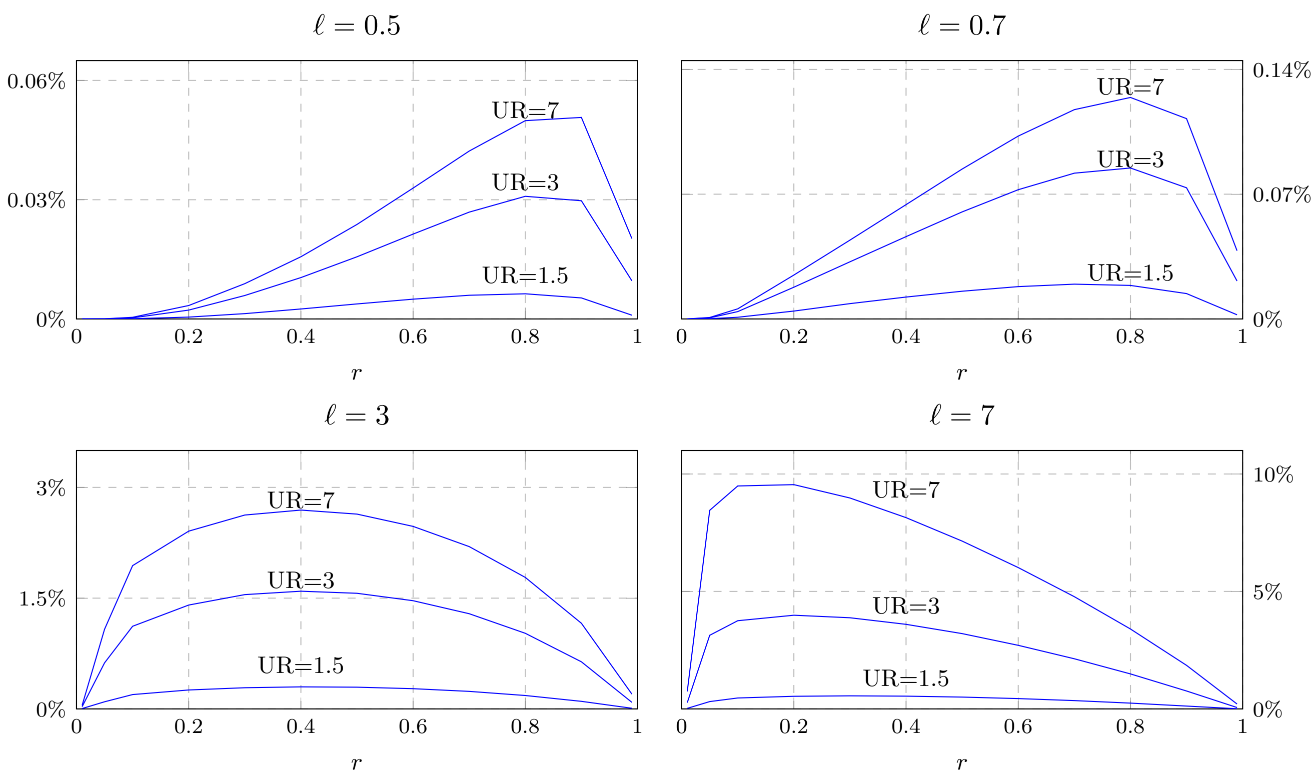


Figure 1: **The worst-case myopic optimality gap $\mathsf{MOG}_{\mathrm{wc}}$ as a function of ℓ, r, and UR.**

Observe that the $\mathsf{MOG}_{\mathrm{wc}}$ tends to zero for values of the service level r close to 0 or 1, consistent with Proposition 4. It is also clear that the $\mathsf{MOG}_{\mathrm{wc}}$ decreases as the uncertainty ratio UR decreases, consistent[13] with Proposition 3. However, the most remarkable fact is the magnitude of the $\mathsf{MOG}_{\mathrm{wc}}$:

[13]Note that, for the exponential case, $\mathrm{UR} = \sqrt{a/(a-2)}$, and hence $a \to \infty$ as $\mathrm{UR} \to 1$. A similar conclusion holds

the maximum value of the MOG over all parameters tested is below 0.6% for $\ell = 0.5$, 0.14% for $\ell = 0.7$ and 3% for $\ell = 3$. It is notable that for these cases, this holds for *any* service level r. In particular, even when r is low and censoring occurs often, there is almost no value of deviating from the myopic policy. *In summary, when ℓ is not too high, the value of exploration beyond the myopic ordering quantity (or of "ordering more") is negligible independent of the problem parameters.* The MOG increases with the value of ℓ, with the worst cases in the instances tested given by $\ell = 7$ and UR $= 7$. Except for these latter cases, the MOG is minimal. We explain the intuition behind these results in the next section.

6.3.2 Analysis of MOG results

We now explore in detail the intuition behind the results presented in the last sections. We start by analyzing, on a sample path basis, the evolution over time of the myopic and optimal order quantities. We measure the level of learning achieved by either policy as the distance between the prescribed order quantity and the optimal order when θ is known. Specifically, if we let $\{y_t^m(\mathbf{D}_1^{t-1})\}_{t=1,\dots,T}$ denote the sequence of orders for a particular sample path $\mathbf{D}_1^T := (D_1, \dots, D_T)$, and $y(\theta) := F^{-1}(r|\theta)$ the optimal order when θ is known. We quantify errors in ordering quantities according to

$$\mathsf{MyopicError}_t := \mathbb{E}\left[\left|\frac{y_t^m(\mathbf{D}_1^{t-1}) - y(\theta)}{y(\theta)}\right|\right], \quad \mathsf{OptError}_t := \mathbb{E}\left[\left|\frac{y_t^*(\mathbf{D}_1^{t-1}) - y(\theta)}{y(\theta)}\right|\right].$$

To assess the downstream impact that these errors have on costs, we define an alternative error measure. Let $C(y|\theta) := \mathbb{E}[L(y, D)|\theta]$ be the expected cost of ordering y, when the demand distribution parameter θ is known. We then define

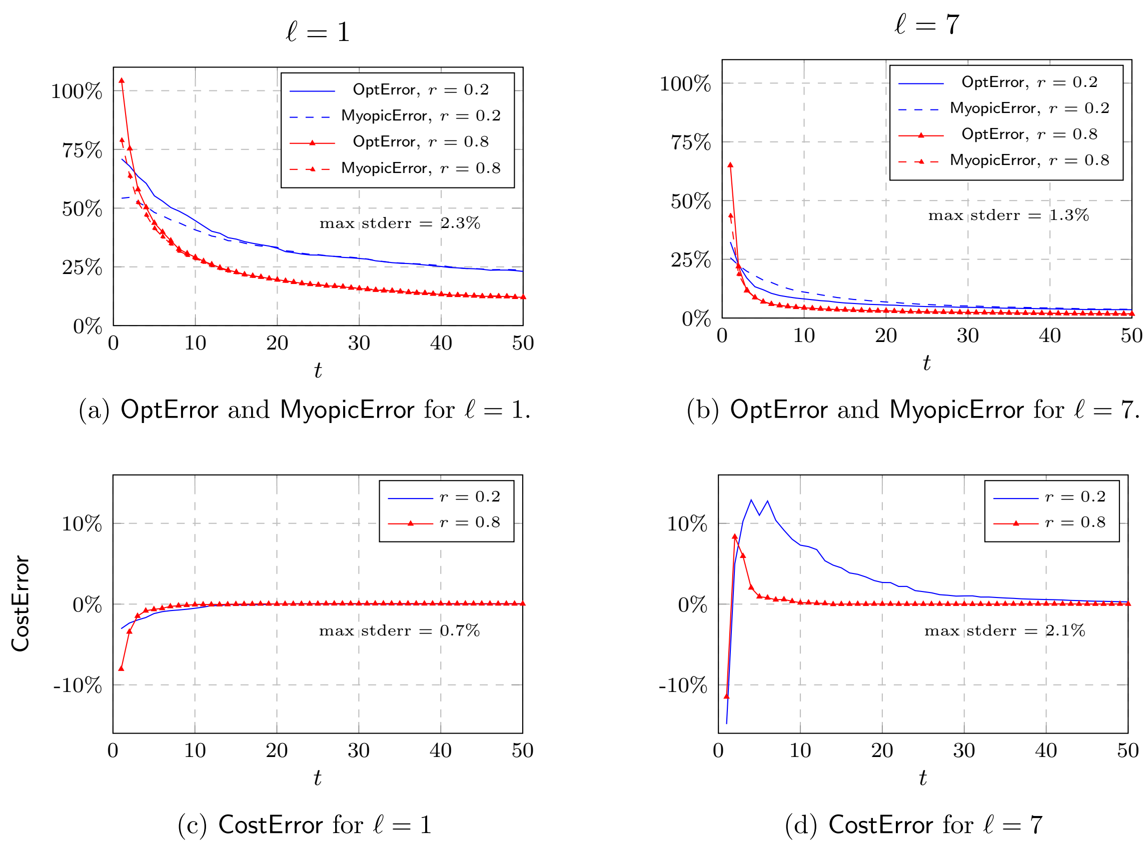
$$\mathsf{CostError}_t := \mathbb{E}\left[\frac{C\big(y_t^m(\mathbf{D}_1^{t-1})|\theta\big) - C\big(y_t^*(\mathbf{D}_1^{t-1})|\theta\big)}{C\big(y_t^*(\mathbf{D}_1^{t-1})|\theta\big)}\right].$$

In words, $\mathsf{CostError}_t$ represents the relative distance in single-period cost between myopic and the optimal policies, in terms of the true expected cost given θ.

For brevity, we fix $T = 50$, UR $= 7$, and consider the $\ell = 1$ (exponential) and $\ell = 7$ cases for two values of r, $r = 0.2$ and $r = 0.8$. For each case we estimate the values of $\mathsf{MyopicError}_t$, $\mathsf{OptError}_t$ and $\mathsf{CostError}_t$ through Monte Carlo simulation. The results are depicted in Figure 2.

Consider first the $\ell = 1$ case. From Figure 2(a) two main conclusions can be drawn: first, there is non-trivial learning taking place, as there is at least a two-fold reduction in $\mathsf{MyopicError}_t$ and $\mathsf{OptError}_t$ for all cases; since both policies prescribe higher orders for $r = 0.8$ the learning

for the general Weibull case.



(a) OptError and MyopicError for $\ell = 1$.

(b) OptError and MyopicError for $\ell = 7$.

(c) CostError for $\ell = 1$

(d) CostError for $\ell = 7$

Figure 2: **OptError_t, MyopicError_t and CostError_t as a function of t, for** $T = 50$**,** UR $= 7$ $\ell \in \{1, 7\}$**, and** $r \in \{0.1, 0.8\}$**.**

rate is faster in that case. Second, the curves of the optimal and myopic policies quickly become indistinguishable. In other words, both policies learn at very similar rates. This fact is reflected in the cost picture, Figure 2(c). The cost difference actually favors the myopic policy in the first periods (i.e., $\text{CostError}_t < 0$) and then the optimal policy slowly takes over (using better information gleaned from the initial higher orders) through the rest of the time horizon[14].

Consider now the case of $\ell = 7$. Recall that, in these examples, UR $= 7$ and hence this case corresponds to the the highest curve in the bottom-right plots of Figure 1. These are the only cases where the MOG becomes significant. In these cases the problem primitives satisfy two conditions: (i) there is little noise associated with demand (i.e., ℓ is large) and (ii) there is high uncertainty about the unknown demand parameter θ (i.e., UR is large). Suppose that we keep UR fixed, and hence the ratio between the two sources of uncertainty in the system is constant. When ℓ is large, the coefficient of variation of the demand distribution is small and hence most of the uncertainty in the system comes from the prior distribution, that is, from the fact that θ is unknown. We are then

[14]While it is not apparent from Figure 2(c) given the scale, CostError_t is positive in both cases for $t \geq 20$.

in a situation where, if θ were known, demand itself would be highly predictable (roughly speaking, almost deterministic) and the subsequent demand-supply mismatch costs would be very low. This is evidenced in the steepness of the learning curves in Figure 2(b): particularly in the $r = 0.8$ case, most of the uncertainty around the true optimal order $y(\theta)$ vanishes after a few periods. Note that, in contrast to the $\ell = 1$ case, for $r = 0.2$ the learning rate of the myopic policy is markedly slower then that of the optimal quantity. This stems from the fact that, even though very few demand observations suffice to nearly learn θ, the myopic censoring probability is high due to the low r and hence the optimal policy can make a significant difference by ordering more at the beginning. This effect is clearly observed in the evolution of $\mathsf{CostError}_t$ in Figure 2(d): the optimal policy incurs a substantially higher cost in the first periods ($\mathsf{CostError}_t < -10\%$) in order to quickly learn θ and then this produces a substantial advantage in the following periods, where $\mathsf{CostError}_t > 0$. Overall the total cost difference, for the $r = 0.2$ case, is no longer negligible, close to 10%, as shown in the bottom-right picture of Figure 1.

To summarize, the prior discussion shows that the MOG can indeed be significant, but only if the problem satisfies simultaneously very specific conditions: (i) there is little noise associated with demand, (ii) there is high uncertainty about the unknown demand parameter θ and (iii) the holding cost h is high relative to the penalty cost p. In this situation, there is a high potential gain from exploring, and hence one may not ignore the exploration-exploitation trade-off. From a practical point of view, however, when the three conditions above are satisfied, one faces a problem of a qualitatively different nature than the typical newsvendor problem we started with: in these cases, demand is, roughly speaking, deterministic and can be essentially learned exactly with very few uncensored observations. Outside this specific case, the results in this section show that the MOG is almost uniformly negligible.

6.3.3 The Myopic Policy Versus Other Heuristics

The objective of the paper is mainly to establish that, the myopic policy, while possibly the simplest heuristic, is a viable alternative that performs well over all time horizons and for a very broad range of parameters. However, one may ask if alternative heuristics could be used. Some heuristics have been proposed in the literature. For example, in the non-parametric setting, Huh and Rusmevichientong (2009) propose a stochastic-gradient based heuristic that one could also apply in our present setting. However, a comparison of performance to the myopic policy would be highly unfair as the myopic policy can make use of much more information through the knowledge of the demand distribution and the prior. A more fair comparison is to heuristics proposed in the Bayesian setting. Chen (2010) proposes heuristics for the non-perishable inventory case, one of which (proposed in §5.1) applies to the perishable case. In particular, it suggests to solve an equation based on the observable case to obtain a prescription for the censored case. The ordering quantity prescribed should solve

$$\mathbb{E}_{a,S}\big[L(y,D)\big] + \mathbb{E}\left[V^o_{T-1}\big(a+1, S+D^\ell\big)\right] = (1+\rho)V^o_T(a,S), \qquad \text{s.t. } y \geq y^m_t(a,S), \tag{23}$$

where $\rho \geq 0$ is a tuning parameter. We will call it the ρ-inflation heuristic. By definition, the prescribed ordering quantity is greater than the myopic ordering quantity.

In Figure 3, the uncertainty ratio UR is set to 3 and for different values of the Weibull parameter ℓ and varying choices of the service level r, we depict the worst case optimality gap for the myopic policy and the ρ-inflation heuristic for two representative values of the tuning parameter ρ.

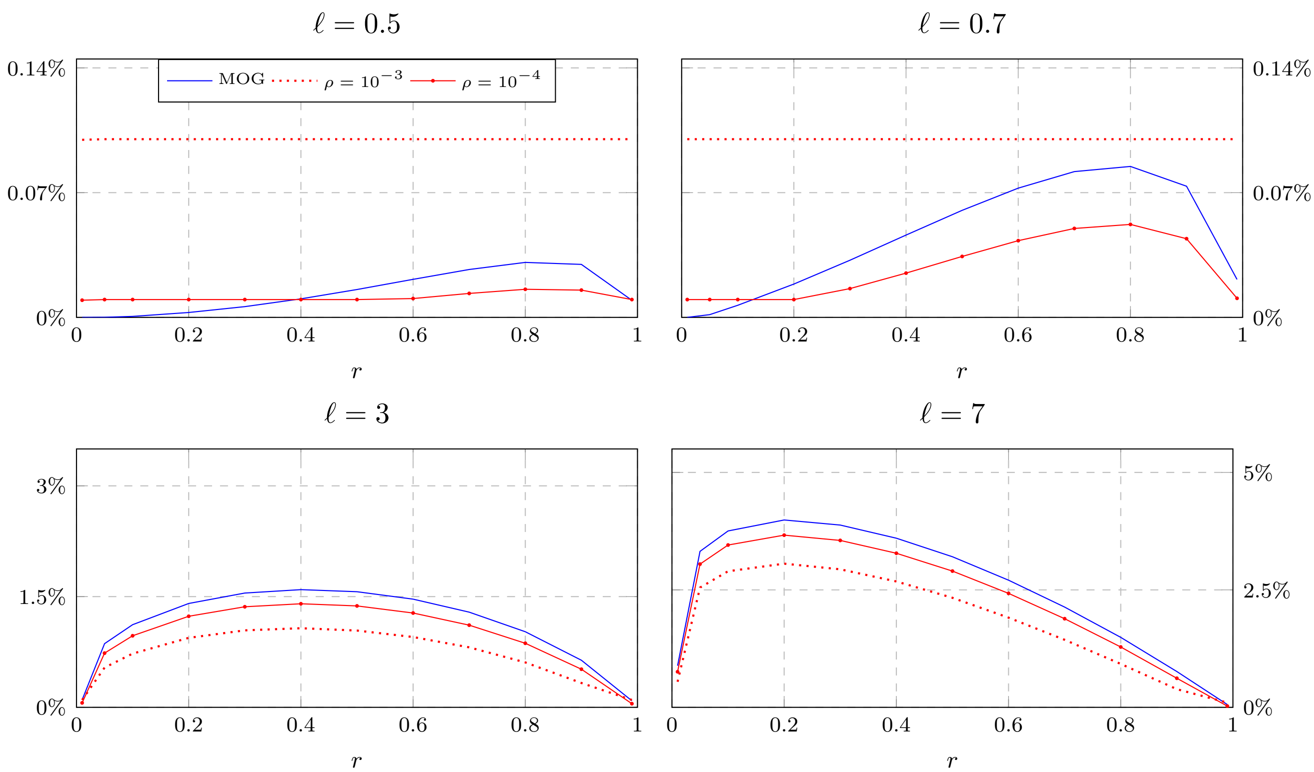


Figure 3: **The worst-case optimality gap for the myopic policy and the ρ-inflation heuristic as a function of ℓ, r for UR $= 3$.**

We observe that the $\rho-$inflation heuristic appears to perform better than the myopic policy for high values of ℓ ($\ell = 3$ and $\ell = 7$) but that for smaller values of ℓ, there is no uniform dominance. In particular, for $\rho = 0.001$, the myopic policy performs uniformly better and for $\rho = 0.001$, the myopic policy performs better for low values of r and worse for high values of r. More important than the relative performance is the fact that both policies achieve very small optimality gaps uniformly over the parameter space. This further reinforces the viability and appeal of the myopic policy given the fact that it does not require the computation of any value function to implement it, and that it possesses strong convergence properties as established in the earlier sections.

7 Analysis of the Cost of Censoring

The current framework and paper analysis can also be used to quantify the impact of censoring on the performance compared to what would be possible if demand samples were fully observable. This would be of interest if, e.g., a firm is considering investments in technology to track lost sales.

In order measure the impact of censoring on performance, we introduce the *Cost of Censoring* (COC), defined as the relative difference between the optimal costs in the censored and uncensored systems, that is,

$$\mathsf{COC} := \frac{V_T^* - V_T^o}{V_T^o}.$$

One may note that

$$\mathsf{COC} = \frac{V_T^* - V_T^o}{V_T^o} \leq \frac{V_T^m - V_T^o}{V_T^o},$$

and hence the MCC bounds the COC and all the upper bounds developed on the MCC in Section 5 apply to the COC.

We next assess the cost of censoring, that is, the relative cost of going from an uncensored to a censored environment. As in Section 6.3, we will consider the worst-case cost of censoring $\mathsf{COC}_{\mathrm{wc}}$ over values of the time horizon $1 \leq T \leq 100$, i.e.,

$$\mathsf{COC}_{\mathrm{wc}} := \max_{1 \leq T \leq 100} \frac{V_T^*(a,1) - V_T^o(a,1)}{V_T^o(a,1)}. \tag{24}$$

In Figure 4, we plot $\mathsf{COC}_{\mathrm{wc}}$ as a function of the uncertainty ratio UR and the service level r, for different values of the Weibull parameter ℓ.

Figure 4 shows that the overall shape and the behavior in extreme cases of the cost of censoring is similar to that of the myopic optimality gap . The main difference between the results in Figure 4 and those shown in Figure 1 is given by the magnitude of the gaps. The $\mathsf{COC}_{\mathrm{wc}}$ appears to be much larger than the $\mathsf{MOG}_{\mathrm{wc}}$.[15] To make a clearer comparison, Figure 5 combines the results of Figures 1 and 4 in single plot: the red lines represent the worst case cost of censoring $\mathsf{COC}_{\mathrm{wc}}$ and the increment of the blue areas represents the worst case myopic optimality gap $\mathsf{MOG}_{\mathrm{wc}}$, starting from $\mathsf{COC}_{\mathrm{wc}}$ (note that the sum of the two has no physical meaning and the plot is intended only

[15] As in Section 6.3, one can observe that the most extreme cases are given in the lower right plot of Figure 4, particularly with higher values of UR. One can apply the same reasoning as before to explain the high gap values in this context: if there is high θ uncertainty and very low demand noise, an uncensored demand observation essentially reveals future demand values, leading to a problem of a different nature.

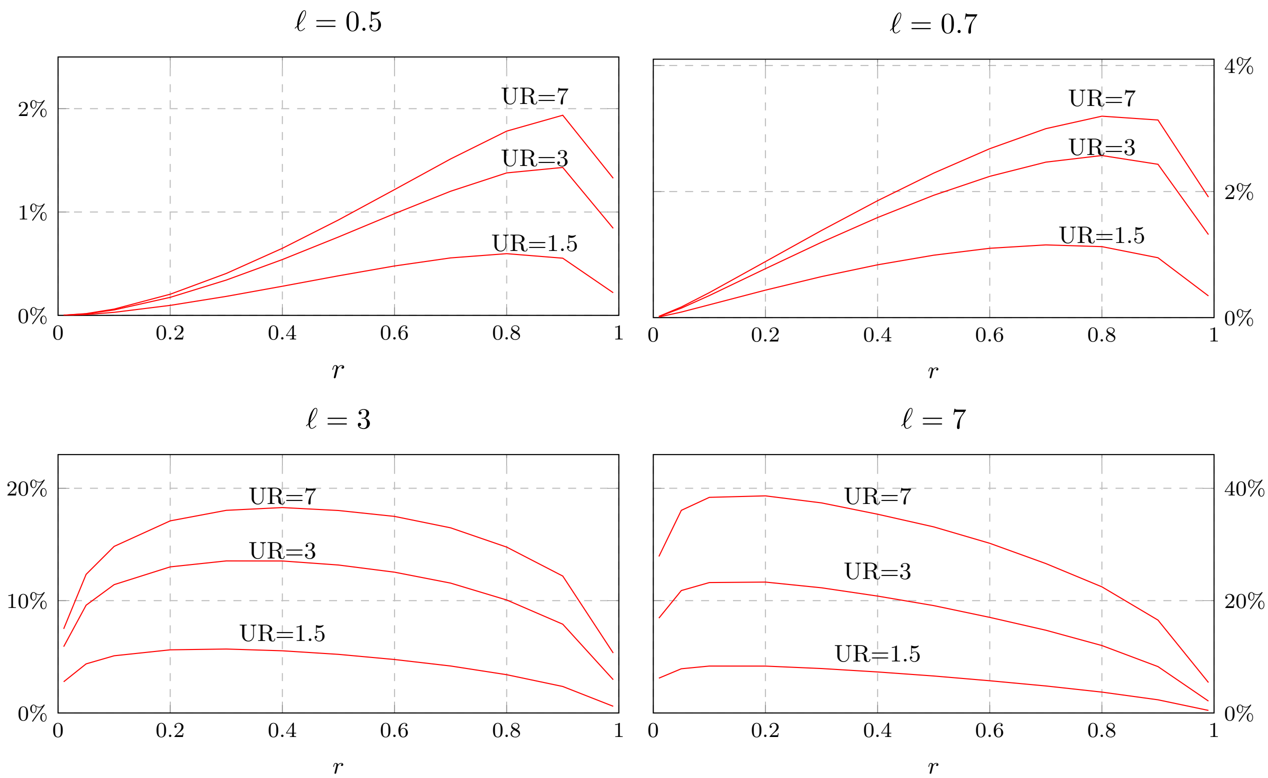


Figure 4: **The worst-case cost of censoring $\mathsf{COC}_{\mathrm{wc}}$ as a function of ℓ, r and UR.**

for visual purposes to be able to compare the relative order of magnitude of COC and MOG). As we can see from the results, the MOG is not only low in absolute terms, but also relative to the COC.

Implications. At a high level, most practitioners are well aware of censoring but rarely fully recognize the exploration-exploitation trade-off, focusing more on attempting to record lost sales. The comparison above is informative in the following sense. It shows that the exploration-exploitation trade-off and the need for forward looking policies introduced by demand censoring (with the computational complexity that might be associated with it) is, for all practical purposes, a second order problem compared to the value that might be generated by investing in processes and technology to uncensor (even partially) demand. An interesting future research direction would be to quantify analytically the differences between the MOG and the COC.

8 Concluding Remarks

In the present paper, we study the implications of demand censoring in inventory problems on optimal or near-optimal decision-making, focusing on the perishable, or newsvendor, case. In par-

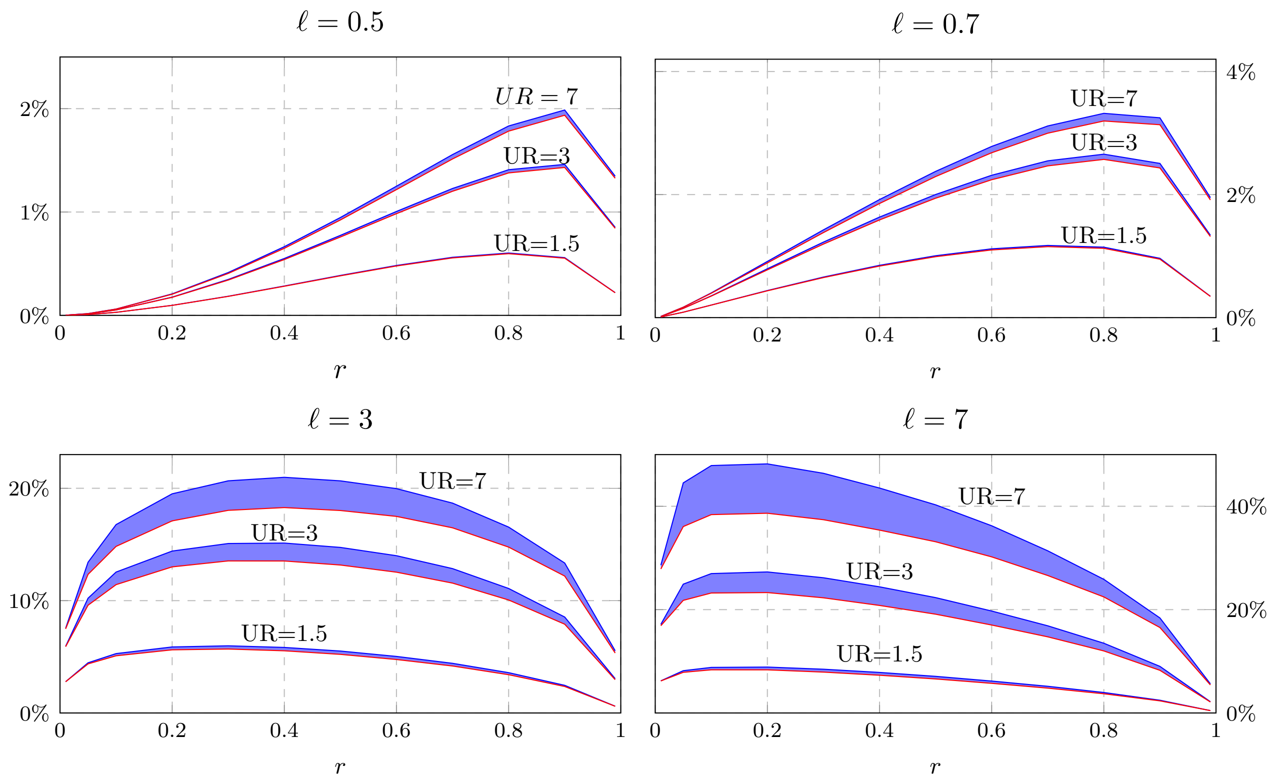


Figure 5: **Comparison of the worst-case cost of censoring $\mathsf{COC}_{\mathrm{wc}}$ and the worst case myopic optimality gap $\mathsf{MOG}_{\mathrm{wc}}$ as a function of ℓ, r and UR.** The red lines represent the $\mathsf{COC}_{\mathrm{wc}}$ and the blue areas represent the $\mathsf{MOG}_{\mathrm{wc}}$, displayed with $\mathsf{COC}_{\mathrm{wc}}$ as a baseline.

ticular, we study how censoring affects decisions, and in particular how the exploration-exploitation trade-off introduced by censoring affects an (otherwise optimal) myopic policy. Through a combination of long-run asymptotic analysis of decisions, and finite time analysis of performance in a more restricted family, we find that, for practical purposes, there is virtually no trade-off between instantaneous performance and information collection in this case: being myopic is "essentially" as good as optimal. Operationally, this surprising fact implies that there is no need to develop and apply optimal policies. The myopic policy is one of the easiest policies to apply in practice, and the present results suggest that it is a viable heuristic to apply in general, yielding near-optimal performance. Furthermore, it is worth noting that in general cases with non-conjugate families of distributions, the dynamic optimization problem becomes infinite dimensional and obtaining an optimal policy is highly intractable, so it is not clear what other alternative one could use. An interesting avenue of future research would be to characterize finite time performance of the myopic policy for such general cases.

References

Abramowitz, M. et al. (1964), *Handbook of Mathematical Functions*, Vol. 55, Courier Dover Publications.

Azoury, K. S. (1985), 'Bayes solution to dynamic inventory models under unknown demand distribution', *Management Science* **31**(9).

Bensoussan, A., Çakanyildirim, M. and Sethi, S. P. (2009), 'Censored newsvendor model revisited with unnormalized probabilities', *Journal Of Industrial And Management Optimization* **5**(2).

Bensoussan, A., Royal, A. and Sethi, S. P. (2009), 'Bayesian and adaptive controls for a newsvendor facing exponential demand', *Risk Decision Anal. 1* **37**(11).

Bertsekas, D. P. and Shreve, S. E. (1978), *Stochastic Optimal Control: The Discrete Time Case*, Vol. 139, Academic Press New York.

Besbes, O. and Muharremoglu, A. (2013), 'On implications of demand censoring in the newsvendor problem', *Management Science* **59**(6), 1407–1424.

Besbes, O. and Zeevi, A. (2011), 'On the minimax complexity of pricing in a changing environment', *Operations Research* **59**(1), 66–79.

Bisi, A., Dada, M. and Tokdar, S. (2011), 'A censored-data multiperiod inventory problem with newsvendor demand distributions', *Manufacturing Service Oper. Management* **13**(4).

Braden, D. J. and Freimer, M. (1991), 'Informational dynamics of censored observations', *Management Science* **37**(11).

Broder, J. and Rusmevichientong, P. (2012), 'Dynamic pricing under a general parametric choice model', *Operations Research* **60**(4), 965–980.

Caro, F. and Gallien, J. (2010), 'Inventory management of a fast-fashion retail network', *Operations Research* **58**, 257–273.

Chen, F. and Han, Y. (2013), Rethinking the value of demand learning in a supply chain with censored demand. Presentation, M&SOM Conference.

Chen, L. (2010), 'Bounds and heuristics for optimal bayesian inventory control with unobserved lost sales', *Operations Research* **58**(2), 396–413.

Chen, L. and Mersereau, A. (2013), Analytics for operational visibility in the retail store: The cases of censored demand and inventory record inaccuracy. Working paper.

Dai, T. and Jerath, K. (2013), Salesforce compensation with inventory considerations. Forthcoming in *Management Science*.

den Boer, A. V. and Zwart, B. (2015), 'Dynamic pricing and learning with finite inventories', *Operations Research* **63**(4), 965–978.
URL: *http://dx.doi.org/10.1287/opre.2015.1397*

Ding, X., Puterman, M. L. and Bisi, A. (2002), 'The censored newsvendor and the optimal acquisition of information', *Operations Research* **50**(3).

Gaur, V., Muthulingam, S. and Lee, J. (2015), Stockout-based substitution and inventory planning in textbook retailing. Working paper.

Heese, S. and Swaminathan, J. M. (2010), 'Inventory and sales effort management with unobservable lost sales', *European Journal of Operational Research* **207**(3), 1263–1268.

Huh, W. T. and Rusmevichientong, P. (2009), 'A non-parametric asymptotic analysis of inventory planning with censored demand', *Mathematics of Operations Research* **34**, 103–123.

Hytönen, T., van Neerven, J., Veraar, M. and Weis, L. (2016), *Analysis in Banach Spaces*, Vol. 1, Springer.

Jain, A., Rudi, N. and Wang, T. (2013), Demand estimation and ordering under censoring: Stock-out timing is (almost) all you need. Working paper.

Kunnumkal, S. and Topaloglu, H. (2008), 'Using stochastic approximation methods to compute optimal base-stock levels in inventory inventory control problems', *Operations Research* **56**, 646–664.

Lariviere, M. and Porteus, E. L. (1999), 'Stalking information: Bayesian inventory management with unobserved lost sales', *Management Science* **45**(3), 346–363.

Lovejoy, W. S. (1990), 'Myopic policies for some inventory models with uncertain demand distributions', *Management Science* **36**(6), 724–738.

Lu, X., Song, J.-S. and Zhu, K. (2005*a*), Inventory control with unobservable lost sales and bayesian updates. Working paper.

Lu, X., Song, J.-S. and Zhu, K. (2005*b*), 'On "the censored newsvendor and the optimal acquisition of information"', *Operations Research* **53**(6), 1024–1026.

Mersereau, A. (2013), Demand estimation from censored observations with inventory record inaccuracy. Working paper.

Qi, F. and Losonczi, L. (2010), 'Bounds for the ratio of two gamma functions', *Journal of Inequalities and Applications* **2010**, 204.

Scarf, H. (1959), 'Bayes solutions of the statistical inventory problem', *The Annals of Mathematical Statistics* **30**(2), 490–508.

Wendel, J. (1948), 'Note on the gamma function', *The American Mathematical Monthly* **55**(9), 563–564.

Online Supplement to "The Exploration-Exploitation Trade-off in the Newsvendor Problem"

Omar Besbes, Juan M. Chaneton, Ciamac C. Moallemi
Graduate School of Business
Columbia University
Henry R. Kravis Hall
665 West 130th Street
New York, NY 10027
ob2105@columbia.edu, jchaneton15@gsb.columbia.edu, ciamac@gsb.columbia.edu

This version: January 20, 2022

A Proofs for Section 2

Lemma A1. *Suppose demand $D \in \mathbb{R}_+$ is a random variable with cumulative distribution function $F(\cdot)$, where $F(\cdot)$ satisfies*

1. *$F(\cdot)$ is continuous over $\mathbb{R}_+$.*
2. *$F(\cdot)$ is strictly increasing over $\mathbb{R}_+$. That is, for all $0 \leq x < y$, $F(x|\theta) < F(y|\theta)$.*

Then, there exists an order quantity $y^* \in \mathbb{R}_+$ that uniquely satisfies

$$F(y^*) = r, \tag{A.1}$$

$$y^* := \underset{y \geq 0}{\operatorname{argmin}} \ \mathbb{E}\left[L(y, D)\right]. \tag{A.2}$$

Proof. Clearly $F(0) = 0$ and $\lim_{x \to \infty} F(x) = 1$. Since $F(\cdot)$ is continuous, applying the intermediate value theorem, there exists exists $y^* \in \mathbb{R}_+$ with $F(y^*) = r$. Since $F(\cdot)$ is strictly increasing, y^* is unique.

Now, since $F(\cdot)$ is strictly increasing, the distribution of D has no atoms, and hence $L(y, D)$ is differentiable in y for almost every D. Further, for such D,

$$\left| \frac{\partial}{\partial y} L(y, D) \right| \leq 1 + \frac{r}{1 - r}.$$

By the dominated convergence theorem,

$$\frac{\partial}{\partial y}\mathbb{E}\left[L(y,D)\right] = F(y) - \frac{r}{1-r}\left(1-F(y)\right) = \frac{F(y)-r}{1-r}.$$

Therefore, since y^* uniquely solves (A.1), y^* is the unique value of y satisfying the first order conditions for the optimization problem in (A.2). Since the objective is convex, y^* must also be the unique global minimizer. □

Proof of Proposition 1. First, consider the informed policy. Under Assumption 1, the cumulative demand distribution $F(\cdot|\theta)$ satisfies the hypothesis of Lemma A1 for almost every θ. The result immediately follows.

For the myopic policy, define the conditional cumulative demand distribution

$$\begin{aligned} F_{t-1}(x) &:= P\left(D_t \le x | \mathcal{F}_{t-1}\right) = \mathbb{E}\left[\mathbb{I}_{\{D_t\le x\}}\middle|\mathcal{F}_{t-1}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{I}_{\{D_t\le x\}}\middle|\mathcal{F}_{t-1},\theta\right]\middle|\mathcal{F}_{t-1}\right] = \mathbb{E}\left[F(x|\theta)|\mathcal{F}_{t-1}\right], \end{aligned}$$

for $x \in \mathbb{R}_+$. Since, under Assumption 1, $F(\cdot|\theta)$ is almost surely continuous and strictly increasing, the dominated convergence theorem implies that $F_{t-1}(\cdot)$ must also be continuous. Now, consider $0 \le x < y$. From Assumption 1, $F(x|\theta) < F(y|\theta)$ almost surely. Taking expectations over θ conditional on $\mathcal{F}_{t-1}$, $F_{t-1}(\cdot)$ must be strictly increasing. Therefore, we can apply Lemma A1 and the result follows. □

■ □ ■

B Proofs for Section 3

B.1 Asymptotic Rate of Censoring

Proof of Theorem 1. Define the process

$$W_t := \sum_{\tau=1}^{t}\left(\mathbb{I}_{\{D_\tau \le y_\tau^m\}} - r\right), \quad \forall\, t \ge 0.$$

Observe that W_t is $\mathcal{F}_t$-measurable, and that by Proposition 1, W_t is a martingale.

Moreover, since $|W_t| \le t$, we have that $\mathbb{E}[W_t^2] \le t^2$. Then, W_t is a square integrable martingale, with quadratic variation process

$$\langle W\rangle_t := \sum_{\tau=1}^{t}\mathbb{E}\left[(W_\tau - W_{\tau-1})^2\middle|\mathcal{F}_{\tau-1}\right] = t\cdot r(1-r).$$

By the martingale strong law of large numbers, as $T \to \infty$,

$$\frac{W_T}{\langle W \rangle_T} \to 0 \text{ almost surely on the event that } \langle W \rangle_T \to \infty.$$

This implies that $W_T/T \to 0$ almost surely, and the result follows. $\square$

■ □ ■

B.2 Strong P-Measurability

Assumption 2 requires strong P-measurability of the random demand measure μ_θ. In other words, μ_θ is the pointwise limit (in $\mathcal{B}$, i.e., under the total variation metric) of a set of simple measure-valued functions, each taking finitely many values. If the parameter space Θ is discrete, this is obviously trivially true. More generally, if Θ is not discrete, this requirement forces a mild degree of regularity between the demand distributions for various values of θ, and mainly serves to rule out pathological cases. For example, the following theorem illustrates a simple sufficient condition to ensure strong P-measurability in the case where the parameter space Θ is finite dimensional.

Theorem B1. *Suppose that $\Theta \subset \mathbb{R}^K$, and that for almost every $\theta_1, \theta_2 \in \Theta$,*

$$\|\mu_{\theta_1} - \mu_{\theta_2}\|_{\mathsf{TV}} \leq L\left(\|\theta_1 - \theta_2\|\right), \tag{B.1}$$

where $L\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous, nondecreasing function with $L(0) = 0$. Then, the random measure $\mu_\theta\colon \Omega \to \mathcal{B}$ is strongly P-measurable.

Proof. As required, we will construct a sequence of random measures (f_n) for $n \geq 1$ with each $f_n\colon \Omega \to \mathcal{B}$ taking finitely many values. Given $n \geq 1$, define the set

$$\Theta_n := \{\theta \in \Theta \ : \ \|\theta\| \leq n\}.$$

Since Θ_n is a bounded set in $\mathbb{R}^K$, it is contained in a compact set. Then, by the Heine-Borel property we can construction a collection of disjoint sets $A_1^n, \ldots, A_{N_n}^n \subset \mathbb{R}^K$ that covers Θ_n, such that each set A_i^n has the property that the distance between any two points is at most 2^{-n}. Select b_i^n to be any point inside $A_i^n \cap \Theta_n$, and define

$$f_n(\omega) = \sum_{i=1}^{N_n} \mathbb{I}_{A_i^n}(\omega) b_i^n.$$

Then, for almost every $\omega \in \Omega$ with $\theta(\omega) \in \Theta_n$, by (B.1),

$$\|f_n(\omega) - \mu_{\theta(\omega)}\|_{\mathsf{TV}} \leq L\left(2^{-n}\right).$$

Since every $\theta \in \Theta$ will be contained in Θ_n for all except finitely many n, clearly

$$\lim_{n\to\infty} \|f_n(\omega) - \mu_{\theta(\omega)}\|_{\mathsf{TV}} = 0,$$

for almost every $\omega \in \Omega$. □

■ □ ■

We illustrate Theorem B1 by establishing strong P-measurability in the setting of classical newsvendor distributions defined in Section 4.1:

Corollary B1. *Suppose the cumulative demand distribution takes the form*

$$F(z|\theta) := 1 - e^{-\theta d(z)}, \quad \textit{for all } z > 0,$$

where $d\colon \mathbb{R}_{++} \to \mathbb{R}_{++}$ *is a differentiable, nondecreasing, and unbounded function with* $d(0) = 0$*, and* θ *takes values in the parameter space* $\Theta := \mathbb{R}_{++}$ *(under an arbitrary prior distribution). Then, the random demand measure* μ_θ *is strongly* P*-measurable.*

Proof. First, consider $0 < \theta_1 < \theta_2 < \infty$. Define

$$x^* := \inf\left\{x > 0 \;:\; d(x) \geq \frac{\log(\theta_2/\theta_1)}{\theta_2 - \theta_1}\right\}.$$

It is easy to verify that the demand density $f(z|\theta) := \theta d'(z)e^{\theta d(z)}$ satisfies

$$f(x|\theta_1) \leq f(x|\theta_2), \quad \forall\, x \leq x^*; \qquad f(x|\theta_1) \geq f(x|\theta_2), \quad \forall\, x \geq x^*.$$

Applying this, we have that

$$\begin{aligned}
\|\mu_{\theta_1} - \mu_{\theta_2}\|_{\mathsf{TV}} &= 2\sup_{A\in\Sigma} |\mu_{\theta_1}(A) - \mu_{\theta_2}(A)| \\
&= 2\sup_{A\in\Sigma} \left|\int_A \big[f(x|\theta_1) - f(x|\theta_2)\big]\,dx\right| \\
&= 2\big[F(x^*|\theta_2) - F(x^*|\theta_1)\big] \\
&= 2\left[\exp\left(\frac{-\log(\theta_2/\theta_1)}{\theta_2/\theta_1 - 1}\right) - \exp\left(\frac{-\log(\theta_2/\theta_1)}{1 - \theta_1/\theta_2}\right)\right].
\end{aligned}$$

Define the change of coordinates $\tilde{\theta}_i := \log \theta_i$, for $i = 1, 2$. Then,

$$\|\mu_{\tilde{\theta}_1} - \mu_{\tilde{\theta}_2}\|_{\mathsf{TV}} = L\left(|\tilde{\theta}_1 - \tilde{\theta}_2|\right),$$

for all $\tilde{\theta}_1, \tilde{\theta}_2 \in \tilde{\Theta} := \mathbb{R}$, where

$$L(\delta) := 2\left[\exp\left(\frac{-\delta}{e^{\delta} - 1}\right) - \exp\left(\frac{-\delta}{1 - e^{-\delta}}\right)\right], \quad \forall\, \delta \geq 0.$$

We can now apply Theorem B1. □

■ □ ■

B.3 Convergence of Beliefs

In this section, we will provide a proof of Theorem 2. This proof builds on the theory of convergence of Banach space-valued martingales, so we begin with some definitions. Let $L_1(\Omega, \mathcal{F}, P; \mathcal{B})$ be the space of L_1-bounded $\mathcal{B}$-valued strongly P-measurable random functions, where the norm $\|\nu\|_1$ is defined for $\nu\colon \Omega \to \mathcal{B}$ according to

$$\|\nu\|_1 := \int_{\Omega} \|\nu(\omega)\|_{\mathsf{TV}}\, P(d\omega).$$

Note that if $\nu(\omega) \in \mathcal{B}$ is a probability measure for almost every $\omega \in \Omega$, $\|\nu\|_1 = 1$.

Define μ_t to be the random measure corresponding to the posterior distribution of demand at the time t, i.e.,

$$\mu_t(A, \omega) := P\left(D_{t+1} \in A | \mathcal{F}_t\right),$$

for all $A \in \Sigma$ (we will sometimes suppress the dependence of μ_t on the sample path ω). We have the following lemma:

Lemma B1 (Martingale Convergence of Beliefs.)**.** *Define*

$$\mathcal{F}_\infty := \sigma\left(\bigcup_{t \geq 1} \mathcal{F}_t\right), \quad \mu_\infty := \mathbb{E}\left[\mu_\theta | \mathcal{F}_\infty\right] \in L_1(\Omega, \mathcal{F}, P; \mathcal{B}).$$

As $t \to \infty$, $\mu_t \to \mu_\infty$ pointwise in $\mathcal{B}$ almost surely. In other words, for almost every $\omega \in \Omega$, we have that

$$\lim_{t \to \infty} \sup_{A \in \Sigma} |\mu_t(A, \omega) - \mu_\infty(A, \omega)| = 0.$$

Proof. The result is an immediate consequence of the martingale convergence theorem for Banach spaces (Theorem 3.3.2, Hytönen et al., 2016). To see this, first observe that since μ_θ is a (random) probability measure and is strongly P-measurable by assumption, we have $\mu_\theta \in L_1(\Omega, \mathcal{F}, P; \mathcal{B})$. By Proposition 2.6.3 and Theorem 2.6.23 of Hytönen et al. (2016), the conditional expectation $\hat{\mu}_t := \mathbb{E}[\mu_\theta|\mathcal{F}_t]$ exists and $\hat{\mu}_t \in L_1(\Omega, \mathcal{F}, P; \mathcal{B})$. However, Now, observe that using the tower property of expectation, for $A \in \Sigma$,

$$\mu_t(A) = P\left(D_{t+1} \in A|\mathcal{F}_t\right) = \mathbb{E}\left[\mathbb{E}\left[\mathbb{I}_{\{D_{t+1}\in A\}}\middle|\mathcal{F}_t, \theta\right]\middle|\mathcal{F}_t\right] = \mathbb{E}\left[\mu_\theta(A)|\mathcal{F}_t\right] = \hat{\mu}_t(A).$$

Therefore, $\mu_t = \mathbb{E}[\mu_\theta|\mathcal{F}_t]$. By applying Theorem 3.3.2 of Hytönen et al. (2016), we have, almost surely, the pointwise convergence

$$\mu_\infty := \mathbb{E}\left[\mu_\theta|\mathcal{F}_\infty\right] = \lim_{t\to\infty} \mathbb{E}\left[\mu_\theta|\mathcal{F}_t\right].$$

□

■ □ ■

Note that a standard, scalar martingale convergence theorem would establish that $\mu_t(A) \to \mu_\infty(A)$ for any set $A \in \Sigma$. Lemma B1 leverages a Banach space martingale convergence theorem to establish that this convergence occurs uniformly over all sets A according to the total variation metric. This will be important in what follows.

While Lemma B1 establishes convergence of the posterior random measures μ_t to a random measure μ_∞, it may be the case that $\mu_\infty \neq \mu_\theta$. However, as the following lemma establishes, convergence to the limiting measure in the total variation metric guarantees convergence of the myopic order quantity.

Lemma B2. *Almost surely, there exists a (random) order quantity $y_\infty \in \mathbb{R}_+$ such that*

$$\lim_{t\to\infty} y_t^m = y_\infty.$$

Proof. For $x \geq 0$, define the cumulative demand distribution

$$F_\infty(x) := \mu_\infty([0, x]) = \mathbb{E}[\mu_\theta([0, x])|\mathcal{F}_\infty] = \mathbb{E}[F(x|\theta)|\mathcal{F}_\infty].$$

Applying the same argument as in the proof of Proposition 1, there exists a unique ($\mathcal{F}_\infty$-measurable random variable) $y_\infty \in \mathbb{R}_+$ satisfying $F_\infty(y_\infty) = r$.

We would like to establish that $y_t^m \to y_\infty$. Define the conditional cumulative demand distribution after time t by $F_t(x) := \mu_t([0,x])$, for $x \geq 0$. Suppose that for some $\omega \in \Omega$,

$$\limsup_{t\to\infty} y_t^m > y_\infty. \tag{B.2}$$

Define

$$\epsilon := \frac{1}{2}\left(\limsup_{t\to\infty} y_t^m - y_\infty\right) > 0.$$

Then, for infinitely many t, we have that $y_t^m > y_\infty + \epsilon$, so

$$r = F_{t-1}(y_t^m) \geq F_{t-1}(y_\infty + \epsilon),$$

where we have applied Proposition 1. For such t,

$$\begin{aligned} r &\geq F_{t-1}(y_\infty + \epsilon) \\ &\geq F_\infty(y_\infty + \epsilon) - \left|F_{t-1}(y_\infty + \epsilon) - F_\infty(y_\infty + \epsilon)\right| \\ &\geq F_\infty(y_\infty + \epsilon) - \sup_{A\in\Sigma} \left|\mu_{t-1}(A) - \mu_\infty(A)\right|. \end{aligned}$$

Taking a limit along this subsequence and applying Lemma B1, we have that

$$r \geq F_\infty(y_\infty + \epsilon).$$

Now, note that $F_\infty(\cdot)$ is a non-decreasing function, and that y_∞ is the unique solution of $F_\infty(y) = r$. Then,

$$r \geq F_\infty(y_\infty + \epsilon) > F_\infty(y_\infty) = r.$$

By contradiction, the set of $\omega \in \Omega$ for which (B.2) holds must be of measure zero. By a symmetric argument for the lim inf case, it must be that $y_t^m \to y_\infty$ almost surely. □

■ □ ■

Now, we are ready to prove the main result.

Proof of Theorem 2. First, observe that conditional on θ, the demands $\{D_t\}$ are i.i.d. Applying the Glivenko-Cantelli theorem, we have that

$$\lim_{T\to\infty} \sup_{x\in\mathbb{R}_+} \left|\frac{1}{T}\sum_{t=1}^{T} \mathbb{I}_{\{D_t \leq x\}} - F(x|\theta)\right| = 0, \quad \text{almost surely.} \tag{B.3}$$

Now, consider a sample path $\omega \in \Omega$ such that Theorem 1, Lemma B2, and (B.3) hold (such sample paths are of measure 1). Fix $\epsilon > 0$. Along this sample path, there exists a time $\tau \geq 1$ such that, for all $t \geq \tau$, $|y_t^m - y_\infty| < \epsilon$. Then, for $T \geq \tau$,

$$\frac{1}{T}\sum_{t=1}^{T} \mathbb{I}_{\{D_t \leq y_t^m\}} \leq \frac{1}{T}\sum_{t=1}^{\tau-1} \mathbb{I}_{\{D_t \leq y_t^m\}} + \frac{1}{T}\sum_{t=\tau}^{T} \mathbb{I}_{\{D_t \leq y_\infty + \epsilon\}}.$$

Taking limits as $T \to \infty$,

$$\begin{aligned}
r &= \frac{1}{T}\sum_{t=1}^{T} \mathbb{I}_{\{D_t \leq y_t^m\}} \\
&\leq \lim_{T\to\infty} \frac{1}{T}\sum_{t=\tau}^{T} \mathbb{I}_{\{D_t \leq y_\infty + \epsilon\}} \\
&\leq F(y_\infty + \epsilon|\theta) + \lim_{T\to\infty} \left| \frac{1}{T}\sum_{t=\tau}^{T} \mathbb{I}_{\{D_t \leq y_\infty + \epsilon\}} - F(y_\infty + \epsilon|\theta) \right| \\
&\leq F(y_\infty + \epsilon|\theta) + \lim_{T\to\infty} \sup_{x \in \mathbb{R}_+} \left| \frac{1}{T}\sum_{t=\tau}^{T} \mathbb{I}_{\{D_t \leq x\}} - F(x|\theta) \right| \\
&= F(y_\infty + \epsilon|\theta).
\end{aligned}$$

Here, the first equality follows from Theorem 1, while the last equality follows from (B.3).

By a symmetric argument, we can establish that for this ω, $r \geq F(y_\infty - \epsilon)$. So, for this sample path, we have that

$$F(y_\infty - \epsilon|\theta) \leq r \leq F(y_\infty + \epsilon|\theta),$$

for all $\epsilon > 0$. Since $F(\cdot|\theta)$ is continuous, we must have that $F(y_\infty|\theta) = r$. By the uniqueness of $y^i(\theta)$ in Proposition 1, we must have that $y_\infty = y^i(\theta)$. □

■ □ ■

C Parametric Bounds and Structural Results

In this section, we will provide proofs for the results in Sections 5 and 6. In what follows, we will use the following notation conventions:

- $\mathbb{E}[\cdot]$ will denote expectation with respect to the pair (θ, D), where $\theta \sim \Gamma(a, S)$ is the unknown parameter of demand, and $D \sim F(\cdot|\theta)$. In some cases, in order to avoid ambiguities, we will specify the hyperparameters a and S in the expectation as $\mathbb{E}_{a,S}[\,\cdot\,]$.

- $C(a,S) := \min_{y\geq 0} \mathbb{E}\big[L(y,D)\big]$ denotes the single period optimal cost.
- $C^o_t(a,S)$ represents the future expected one period cost, $t+1$ periods in the future, when demands are observable. That is, $C^o_t(a,S) := \mathbb{E}\left[C(a+t, S+D^\ell_1+\ldots+D^\ell_t)\right]$.
- $C(\theta) := \min_{y\geq 0} \mathbb{E}\big[L(y,D)\big|\theta\big]$ denotes the single period optimal cost when θ is known. We also will denote $C^o_\infty(a,S) := \mathbb{E}\left[C(\theta)\right]$.
- The myopic order quantity $y^m(a,S)$ is defined as $y^m(a,S) := M^{-1}(r|a,S)$. The myopic order quantity when θ is known is denoted as $y^m(\theta) := F^{-1}(r|\theta)$.
  In the Weibull case these can be explicitly written as $y^m(a,S) := S^{1/\ell}\left((1-r)^{-1/a}-1\right)^{1/\ell}$ and $y^m(\theta) := \theta^{-1/\ell}(-\log(1-r))^{1/\ell}$ respectively.

The proofs rely on a set of shorter technical lemmas presented in Section E.

C.1 Proofs for Section 5

Proof of Lemma 1. We will establish the result of Lemma 1 for general newsvendor distributions. Define, for this proof only,

$$\begin{aligned} F(x|\theta) &:= 1-e^{-\theta d(x)}, \\ \pi(\theta|a,S) &:= \frac{S^a\theta^{a-1}e^{-S\theta}}{\Gamma(a)}, \end{aligned}$$

where $d\colon [0,\infty)\to[0,\infty)$ is a differentiable, increasing and unbounded function with $d(0)=0$.[1]

In the general newsvendor distribution case, one has

$$V^m_T(a,S) = C(a,S) + \mathbb{E}_{a,S}[V^m_{T-1}(a+1, S+d(D))] + \Gamma_{T-1}(a,S), \tag{C.1}$$

where

$$\Gamma_{T-1}(a,S) := (1-r)V^m_{T-1}(a, S+d(y^m)) - \mathbb{E}_{a,S}[V^m_{T-1}(a+1,S+d(D))\mathbb{I}_{\{D\geq y^m\}}],$$

and y^m represents the myopic order quantity. Here, with some slight abuse of notation, we keep the same notation as in the Weibull case in the main text. To lighten notation, we omit the dependency of y^m on a and S; in what follows, y^m always represents the myopic order quantity with respect to a and S, that is $y^m = d^{-1}(S[(1-r)^{-1/a}-1])$.

[1] Note that $d(\cdot)$ is defined to be *nondecreasing* in Section 4.1. To make the exposition more clear in this proof, and that of Lemma E3, we assume $d(\cdot)$ to be strictly increasing, though the proof can be extended to the more general case by defining, for example, $d^{-1}(z) := \inf\{y|d(z)=y\}$.

First note that, by Lemma E3,

$$\mathbb{E}_{a,S}\left[V_T^m(a,S+d(D))|D\geq y^m\right]=\mathbb{E}_{a,S+d(y^m)}\left[V_T^m\big(a,S+d(y^m)+d(D)\big)\right]. \tag{C.2}$$

Next, note that

$$\begin{aligned}
\Gamma_T(a,S) &= (1-r)V_T^m(a,S+d(y^m))-(1-r)\mathbb{E}_{a,S}[V_T^m(a+1,S+d(D))|D\geq y^m] \\
&= (1-r)\Big[C(a,S+d(y^m))+\mathbb{E}_{a,S+d(y^m)}[V_{T-1}^m(a+1,S+d(y^m)+d(D))]+ \\
&\qquad \Gamma_{T-1}(a,S+d(y^m))-\mathbb{E}_{a,S}[V_T^m(a+1,S+d(D))|D\geq y^m]\Big] \\
&= (1-r)\Big[C(a,S+d(y^m))+\mathbb{E}_{a,S+d(y^m)}[V_{T-1}^m(a+1,S+d(y^m)+d(D))]+ \\
&\qquad \Gamma_{T-1}(a,S+d(y^m))-\mathbb{E}_{a,S+d(y^m)}[V_T^m\big(a+1,S+d(y^m)+d(D)\big)]\Big],
\end{aligned} \tag{C.3}$$

where the second equality follows from expanding $V_T^m(a,S+d(y^m))$ according to (C.1) and the third inequality follows from applying equation (C.2) to the last term in the second equation. Define

$$C_t^m(a,S):=V_{t+1}^m(a,S)-V_t^m(a,S), \quad \text{for} \;\; t=0,\dots,T.$$

$C_t^m(a,S)$ represents the future expected cost $t+1$ periods in the future, if the myopic policy is applied, and we start with hyperparamters a and S. One can then rewrite (C.3) as

$$\begin{aligned}
\Gamma_T(a,S) &= (1-r)\Big[C(a,S+d(y^m))+\Gamma_{T-1}(a,S+d(y^m)) \\
&\qquad -\mathbb{E}_{a,S+d(y^m)}[C_{T-1}^m(a+1,S+d(y^m)+d(D))]\Big] \\
&= (1-r)\Big[C(a,S(1-r)^{-1/a})+\Gamma_{T-1}(a,S(1-r)^{-1/a}) \\
&\qquad -\mathbb{E}_{a,S(1-r)^{-1/a}}[C_{T-1}^m(a+1,S(1-r)^{-1/a}+d(D))]\Big],
\end{aligned}$$

where the last equality comes from the fact that $S+d(y^m)=S(1-r)^{-1/a}$. By repeating the argument, one obtains

$$\begin{aligned}\Gamma_T(a,S) &= \sum_{k=1}^{T}(1-r)^k\left[C\left(a,S(1-r)^{-k/a}\right)-\mathbb{E}_{a,S(1-r)^{-k/a}}\left[C^m_{T-k}\left(a+1,S(1-r)^{-k/a}+d(D)\right)\right]\right]\\ &\le \sum_{k=1}^{T}(1-r)^k\left[C\left(a,S(1-r)^{-k/a}\right)-\mathbb{E}_{a,S(1-r)^{-k/a}}\left[C^o_{T-k}\left(a+1,S(1-r)^{-k/a}+d(D)\right)\right]\right]\\ &= \sum_{k=1}^{T}(1-r)^k\left[C\left(a,S(1-r)^{-k/a}\right)-C^o_{T-k+1}\left(a,S(1-r)^{-k/a}\right)\right],\end{aligned}$$

The inequality is due to the fact that for any $t \ge 1$, $C^o_t(a,S) \le C^m_t(a,S)$, a fact we formally prove in Lemma E4. This completes the proof.

□

■ □ ■

Proof of Theorem 3. The proof is based on first bounding the difference $V^m_T(a,S) - V^o_T(a,S)$ by the difference $V^o_{T-1}(a,S) - (T-1)C^o_\infty(a,S)$ properly scaled. We then derive a bound on the latter difference.

Specializing the bound on $\Gamma_T(a,S)$ given in Lemma 1 for the Weibull case, one obtains

$$\begin{aligned}\Gamma_T(a,S) &\le \sum_{k=1}^{T}(1-r)^k\left[C(a,S(1-r)^{-k/a})-C^o_{T-k+1}(a,S(1-r)^{-k/a})\right]\\ &\overset{(a)}{=} S^{1/\ell}\sum_{k=1}^{T}(1-r)^{k\frac{a\ell-1}{a\ell}}\left[C(a,1)-C^o_{T-k+1}(a,1)\right]\\ &\overset{(b)}{\le} \left[\sum_{k=1}^{T}(1-r)^{k\frac{a\ell-1}{a\ell}}\right]\left[C(a,S)-C^o_\infty(a,S)\right], \qquad \text{(C.4)}\end{aligned}$$

where (a) follows from the scalability property and (b) follows from the fact that $C^o_t(a,S) \ge C^o_\infty(a,S)$ for any $t = 0,\ldots,T$.[2]

By equation (C.1), page 9, specialized to the Weibull case, we have

$$V^m_T(a,S) - V^o_T(a,S) = \mathbb{E}_{a,S}\left[V^m_{T-1}(a+1,S+D^\ell) - V^o_{T-1}(a+1,S+D^\ell)\right] + \Gamma_{T-1}(a,S).$$

[2]It is not hard to derive this fact from the definition of $C^o_t(a,S)$ and $C^o_\infty(a,S)$. Intuitively, if θ is known there is always less demand uncertainty than if θ is unknown (even after any number of demand observations) and hence $C^o_\infty(a,S)$ is smaller than $C^o_t(a,S)$.

If we denote $\widehat{D}_t^\ell := D_1^\ell + \ldots + D_t^\ell$ and proceed recursively, we obtain

$$
\begin{aligned}
V_T^m(a,S) - V_T^o(a,S) &= \sum_{t=0}^{T-2} \mathbb{E}_{a,S}\left[\Gamma_{T-1-t}\left(a+t, S+\widehat{D}_t^\ell\right)\right] \\
&\leq \left[\sum_{k=1}^{T}(1-r)^{k\frac{a\ell-1}{a\ell}}\right] \sum_{t=0}^{T-2} \mathbb{E}_{a,S}\left[C(a+t, S+\widehat{D}_t^\ell) - C_\infty^o(a+t, S+\widehat{D}_t^\ell)\right] \\
&= \left[\sum_{k=1}^{T}(1-r)^{k\frac{a\ell-1}{a\ell}}\right] \sum_{t=0}^{T-2} \left(C_t^o(a,S) - C_\infty^o(a,S)\right) \\
&= \left[\sum_{k=1}^{T}(1-r)^{k\frac{a\ell-1}{a\ell}}\right] \left(V_{T-1}^o(a,S) - (T-1)C_\infty^o(a,S)\right) \\
&= \frac{\lambda - \lambda^{T+1}}{1-\lambda}\left(V_{T-1}^o(a,S) - (T-1)C_\infty^o(a,S)\right), \qquad \text{(C.5)}
\end{aligned}
$$

where the inequality follows by (C.4).

We now bound the difference $V_{T-1}^o(a,S) - (T-1)C_\infty^o(a,S)$. For that we rely on lemma E5 that bounds the difference $C_t^o(a,S) - C_\infty^o(a,S)$ of costs in the observable case.

$$
\begin{aligned}
&V_{T-1}^o(a,S) - (T-1)C_\infty^o(a,S) = \sum_{t=0}^{T-2}\left[C_t^o(a,S) - C_\infty^o(a,S)\right] \\
&\leq \; S^{1/\ell}K(r,\ell)\left[\frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell}\right]^{\frac{1}{\ell}} \sum_{t=0}^{T-2}\frac{1}{a+t-1/\ell} \\
&\leq \; S^{1/\ell}K(r,\ell)\left[\frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell}\right]^{\frac{1}{\ell}} \left[\log(a-1/\ell+T-2) - \log(a-1/\ell) + (a-1/l)^{-1}\right] \\
&\leq \; S^{1/\ell}K(r,\ell)\left[\frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell}\right]^{\frac{1}{\ell}} \left[\log(1+\frac{T}{a-1/\ell}) + (a-1/l)^{-1}\right]. \qquad \text{(C.6)}
\end{aligned}
$$

Let us now define

$$
Q(a,r,\ell) := [K(r,\ell)]^\ell \frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell}.
$$

Note that, because $\exp\{x\} \leq 1 + (e-1)x$ for any $x \in (0,1)$, we have that

$$
\frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell} = O\left(\frac{1}{a-1/\ell} + \frac{e-1}{(a-1/\ell)^2}\right) = O\left(\frac{1}{a}\right) \quad \text{as } a \to \infty.
$$

Therefore, $Q(a,r,\ell) = O(1/a)$ when r and ℓ are fixed.

By substituting $Q(a,r,\ell)$ in (C.6) and combining the inequality with (C.5), we obtain the result of the theorem and the proof is complete.

□

■ □ ■

Proof of Proposition 2. As stated in the proof of Theorem 3, it is not hard to show that $C_t^o(a,S) \geq C_\infty^o(a,S)$ for all t. Therefore,

$$V_T^o(a,S) := \sum_{t=0}^{T-1} C_t^o(a,S) \geq TC_\infty^o(a,S) := T\mathbb{E}_{a,S}\big[C(\theta)\big].$$

Combining the latter with the result of Theorem 3, one obtains

$$\frac{V_T^m(a,S) - V_T^o(a,S)}{V_T^o(a,S)} \leq S^{1/\ell}\left[Q(a,r,\ell)\right]^{1/\ell} \frac{\lambda - \lambda^{T+1}}{1-\lambda} \frac{\log\left(1 + T/(a-1/\ell)\right) + (a-1/l)^{-1}}{T\mathbb{E}_{a,S}\big[C(\theta)\big]}.$$

Since the right side is $O(T^{-1}\log(T))$ as $T \uparrow \infty$, the proof is complete. □

■ □ ■

Proof of Proposition 3. Note that, since the myopic value of information is independent of S (cf. the scalability property), the statement is equivalent to

$$\frac{V_T^m(a,a) - V_T^o(a,a)}{V_T^o(a,a)} = O\left(\frac{1}{a}\right) \text{ as } a \to \infty.$$

We start by showing that $V_T^o(a,a)$ is lower bounded by a positive constant for any a. Since $V_T^o(a,S) := \sum_{t=0}^{T-1} C_t^o(a,S)$, it suffices to show that each term $C_t^o(a,a)$ is lower bounded by a positive constant itself,

$$\begin{aligned}
C_t^o(a,a) &\geq C_\infty^o(a,a) \\
&= \mathbb{E}_{a,a}\left[\frac{1}{1-r}\int_{y^m(\theta)}^\infty xf(x|\theta)dx - \int_0^\infty \overline{F}(x|\theta)dx\right] \\
&= \mathbb{E}_{a,a}\left[\theta^{-1/\ell}\right]\left[\frac{1}{1-r}\int_{y^m(1)}^\infty xf(x|1)dx - \int_0^\infty \overline{F}(x|1)dx\right] \\
&\geq \mathbb{E}_{a,a}\left[\theta\right]^{-1/\ell}\left[\frac{1}{1-r}\int_{y^m(1)}^\infty xf(x|1)dx - \int_0^\infty \overline{F}(x|1)dx\right] \\
&= \frac{1}{1-r}\int_{y^m(1)}^\infty xf(x|1)dx - \int_0^\infty \overline{F}(x|1)dx > 0,
\end{aligned}$$

where the first equality follows from Lemma E1 and the second inequality is a result of Jensen's inequality applied to the expectation term. The last equality follows from the fact that $\mathbb{E}_{a,a}[\theta] = 1$. We then have shown that

$$V_T^o(a,a) \geq TC_\infty^o(a,a) \geq T\left[\frac{1}{1-r}\int_{y^m(1)}^{\infty} xf(x|1)dx - \int_0^\infty \overline{F}(x|1)dx\right] =: \underline{m} > 0.$$

Using Theorem 3 one obtains

$$\frac{V_T^m(a,a) - V_T^o(a,a)}{V_T^o(a,a)} \leq \frac{\lambda - \lambda^{T+1}}{\underline{m}(1-\lambda)}\left[aQ(a,r,\ell)\right]^{1/\ell}\left[\log(1+\frac{T}{a-1/\ell}) + \frac{1}{a-1/l}\right],$$

and since $aQ(a,r,\ell) = O(1)$, the right side is $O(a^{-1})$ as $a \uparrow \infty$ and the result is established. □

■ □ ■

Proof of Proposition 4. Following the inequality given in (C.5), in the proof of Theorem 3 (page 12), it suffices to show that

$$\frac{\lambda - \lambda^{T+1}}{1-\lambda}\frac{V_{T-1}^o(a,S) - (T-1)\mathbb{E}\big[C(\theta)\big]}{V_{T-1}^o(a,S)} = \begin{cases} O\left((1-r)^{1-1/a\ell}\right) & \text{as} \quad r \to 1^-, \\ O\left(r^{1/\ell}\right) & \text{as} \quad r \to 0^+. \end{cases} \tag{C.7}$$

where $\lambda := (1-r)^{1-\frac{1}{a\ell}}$.

Note: for this proof, all expectations are taken with respect to hyperparameters a and S, that is, $\mathbb{E}[\cdot] \equiv \mathbb{E}_{a,S}[\cdot]$.

i.) We start with the case $r \to 1^-$. Let us start by noting that $(\lambda - \lambda^{T+1})/(1-\lambda) = O((1-r)^{1-1/a\ell})$ as $r \to 1^-$ and hence, it suffices, for example, to show that the ratio involving the value functions in (C.7) converges to a constant. In particular we will show that

$$\frac{(T-1)\mathbb{E}\big[C(\theta)\big]}{V_{T-1}^o(a,S)} \to 0, \quad \text{as} \quad r \to 1^-.$$

It is not hard to see that both numerator and denominator converge to infinity as $r \to 1^-$. We can therefore apply L'Hôpital's rule and differentiate both terms with respect to r. Using the expression for $C(\theta)$ given in Lemma E1 one has

$$\begin{aligned}\mathbb{E}\big[C(\theta)\big] &= \mathbb{E}\left[\frac{1}{1-r}\int_{y^m(\theta)}^{\infty} xf(x|\theta)dx - \int_0^{\infty}\overline{F}(x|\theta)dx\right]\\ &= \mathbb{E}\left[y^m(\theta) + \frac{1}{1-r}\int_{y^m(\theta)}^{\infty}\overline{F}(z|\theta)dz - \int_0^{\infty}\overline{F}(z|\theta)dz\right]\\ &= \mathbb{E}\left[\theta^{-1/\ell}\right](-\log(1-r))^{1/\ell} + \frac{\mathbb{E}\left[\theta^{-1/\ell}\right]}{1-r}\int_{(-\log(1-r))^{1/\ell}}^{\infty} e^{-z^\ell}dz - \mathbb{E}\left[\int_0^{\infty} e^{-\theta z^\ell}dz\right].\end{aligned}$$

By differentiating with respect to r one obtains

$$\begin{aligned}\frac{\partial}{\partial r}\mathbb{E}\big[C(\theta)\big] &= \mathbb{E}\left[\theta^{-1/\ell}\right]\frac{\partial}{\partial r}\left[(-\log(1-r))^{1/\ell} + \frac{1}{1-r}\int_{(-\log(1-r))^{1/\ell}}^{\infty} e^{-z^\ell}dz\right]\\ &= \mathbb{E}\left[\theta^{-1/\ell}\right]\frac{1}{(1-r)^2}\int_{(-\log(1-r))^{1/\ell}}^{\infty} e^{-z^\ell}dz\\ &= \mathbb{E}\left[\frac{1}{(1-r)^2}\int_{y^m(\theta)}^{\infty} e^{-\theta z^\ell}dz\right],\end{aligned}$$

where the second equality is a result of the chain rule and the fundamental theorem of calculus applied to the integral term.

A similar argument applied to $C_t^o(a,S)$ yields that

$$\frac{\partial}{\partial r}\mathbb{E}\big[C_t^o(a,S)\big] = \mathbb{E}\left[\frac{1}{(1-r)^2}\int_{y^m(a_t,S_t)}^{\infty}\overline{M}(z|a_t,S_t)dz\right],$$

where $S_t := S + D_1^\ell + \ldots + D_t^\ell$ and $a_t := a+t$. Then one has that

$$\begin{aligned}\lim_{r\to 1^-}\frac{(T-1)\mathbb{E}\big[C(\theta)\big]}{V_{T-1}^o(a,S)} &\overset{(a)}{=} \lim_{r\to 1^-}\frac{(T-1)\mathbb{E}\left[\int_{y^m(\theta)}^{\infty} e^{-\theta z^\ell}dz\right]}{\sum_{t=0}^{T-2}\mathbb{E}\left[\int_{y^m(a_t,S_t)}^{\infty}\overline{M}(z|a_t,S_t)dz\right]}\\ &\overset{(b)}{=} \lim_{r\to 1^-}\frac{(T-1)(1-r)\frac{\partial}{\partial r}\mathbb{E}\left[y^m(\theta)\right]}{\sum_{t=0}^{T-2}(1-r)\frac{\partial}{\partial r}\mathbb{E}\left[y^m(a_t,S_t)\right]}\\ &\overset{(c)}{=} \lim_{r\to 1^-}\frac{(T-1)\mathbb{E}\left[\theta^{-1/\ell}\right](-\log(1-r))^{1/\ell-1}(1-r)^{-1}}{\sum_{t=0}^{T-2}\mathbb{E}\left[S_t^{1/\ell}\right]\left((1-r)^{-1/a_t}-1\right)^{1/\ell-1}a_t^{-1}(1-r)^{-1/a_t-1}}\end{aligned}$$

where (a) follows from applying L'Hôpital's rule, (b) follows from applying L'Hôpital's rule and interchanging differentiation and expectation[3] and (c) is a result of applying L'Hôpital's to the right side of (b) . Elementary calculus (in particular, repetitively applying L'Hôpital's rule to the ratio) yields that the last limit is equal to 0, and hence the $r \to 1^-$ case is complete.

ii.) We now analyze the case $r \to 0^+$. Let us start by noting that, following (C.7) and noting that $(\lambda - \lambda^{T+1})/(1-\lambda) \to T$ as $r \to 0^+$, it suffices to show that

$$\lim_{r \to 0^+} \frac{V^o_{T-1}(a,S) - (T-1)\mathbb{E}\big[C(\theta)\big]}{r^{1/\ell} V^o_{T-1}(a,S)}$$

exists and is finite.

Defining the following notation

$$f(r) \; := \; V^o_{T-1}(a,S), \qquad g(r) \; := \; (T-1)\mathbb{E}\big[C(\theta)\big],$$

we aim to establish

$$\lim_{r \to 0^+} \frac{f(r) - g(r)}{r^{1/\ell} f(r)} < \infty. \tag{C.8}$$

By following similar arguments as the ones used in *i.)* one can show that

a) Both numerator and denominator in (C.8) converge to 0 as $r \to 0^+$.

b) $\lim_{r \to 0^+} f'(r) = \lim_{r \to 0^+} g'(r) = (T-1)\mathbb{E}[D]$.

c) The second derivatives of $f(\cdot)$ and $g(\cdot)$ are given by

$$\begin{aligned} f''(r) &= \frac{2}{(1-r)^3}\widetilde{f}(r) + \frac{1}{(1-r)^2}\widetilde{f}'(r), \\ g''(r) &= \frac{2}{(1-r)^3}\widetilde{g}(r) + \frac{1}{(1-r)^2}\widetilde{g}'(r), \end{aligned}$$

where

$$\begin{aligned} \widetilde{f}(r) &:= \sum_{t=0}^{T-2} \mathbb{E}\left[\int_{y^m(a_t,S_t)}^{\infty} \overline{M}(z|a_t,S_t)dz\right], \\ \widetilde{g}(r) &:= (T-1)\mathbb{E}\left[\int_{y^m(\theta)}^{\infty} e^{-\theta z^\ell} dz\right], \end{aligned}$$

[3] The interchange in the numerator is justified by the Dominated Convergence Theorem: $\int_{y^m(\theta)}^{\infty} e^{-\theta z^\ell} dz \leq \int_0^{\infty} e^{-\theta z^\ell} dz$ for any $r \in (0,1)$ and $\theta > 0$, and $\mathbb{E}\left[\int_0^{\infty} e^{-\theta z^\ell} dz\right] = \mathbb{E}\left[\mathbb{E}[D|\theta]\right] = \mathbb{E}\left[D\right] < \infty$. A similar argument justifies the interchange in the denominator.

and

$$
\begin{aligned}
\widetilde{f}'(r) &:= -\frac{1}{\ell}\sum_{t=0}^{T-2}\frac{\mathbb{E}\left[S_t^{1/\ell}\right]}{a_t}\left((1-r)^{-1/a_t}-1\right)^{1/\ell-1}(1-r)^{-1/a_t},\\
\widetilde{g}'(r) &:= -\frac{(T-1)}{\ell}\mathbb{E}\left[\theta^{-1/\ell}\right](-\log(1-r))^{1/\ell-1}.
\end{aligned}
$$

By item b) above and L'Hôpital's rule, one has that

$$
\lim_{r\to 0^+} r^{-1}f(r) = \lim_{r\to 0^+} r^{-1}g(r) = (T-1)\mathbb{E}[D]. \tag{C.9}
$$

This implies that, if we take derivatives in (C.8) both the numerator and denominator converge to 0. By taking second derivatives one obtains

$$
\lim_{r\to 0^+}\frac{f''(r)-g''(r)}{r^{1/\ell-1}\left[(\frac{1}{\ell}-1)\frac{1}{\ell}r^{-1}f(r)+\frac{2}{l}f'(r)+rf''(r)\right]} \tag{C.10}
$$

Next, we establish that

$$
\lim_{r\to 0^+}\frac{f''(r)-g''(r)}{r^{1/\ell-1}}\in\mathbb{R}. \tag{C.11}
$$

Let us start by noting that, by elementary calculus,

$$
\begin{aligned}
\lim_{r\to 0^+}\frac{(-\log(1-r))^{1/\ell-1}}{r^{1/\ell-1}} &= 1,\\
\lim_{r\to 0^+}\frac{\left((1-r)^{-1/a_t}-1\right)^{1/\ell-1}(1-r)^{-1/a_t}}{r^{1/\ell-1}} &= 1/a_t^{1/\ell-1},
\end{aligned} \tag{C.12}
$$

and therefore

$$
\lim_{r\to 0^+}\frac{\widetilde{f}'(r)-\widetilde{g}'(r)}{r^{1/\ell-1}} = \frac{T-1}{\ell}\mathbb{E}\left[\theta^{-1/\ell}\right]-\sum_{t=0}^{T-2}\mathbb{E}\left[\frac{S_t^{1/\ell}}{a_t^{1/\ell}}\right] =: K\in\mathbb{R}_{++},
$$

where $K>0$ follows from Jensen's inequality and the law of iterated expectations (see page 28 for a detailed derivation of this fact). This implies that

$$
\lim_{r\to 0^+}\frac{f''(r)-g''(r)}{r^{1/\ell-1}} = \lim_{r\to 0^+}2\frac{\widetilde{f}(r)-\widetilde{g}(r)}{r^{1/\ell-1}}+K.
$$

Note that if $\ell\geq 1$, and because $\left[\widetilde{f}(r)-\widetilde{g}(r)\right]\to 0$, the proof is complete. If $\ell<1$, we can apply

L'Hôpital's rule to obtain

$$\lim_{r\to 0^+} \frac{f''(r)-g''(r)}{r^{1/\ell-1}} \quad = \quad \lim_{r\to 0^+} \frac{2}{\frac{1}{\ell}-1} \frac{\widetilde{f}'(r)-\widetilde{g}'(r)}{r^{1/\ell-2}} + K = K,$$

where the second equality follows from applying (C.12) and the expressions for $\widetilde{f}'(r)$ and $\widetilde{g}'(r)$. Hence (C.11) is established.

Now, by combining (C.11) and (C.9) above, one obtains that the limit in (C.10) is finite. This completes the proof. □

■ □ ■

C.2 Proofs for Section 6

Proof of Proposition 5. We proceed by induction. For $T = 1$ we have $V_1^m(a,S) = C(a,S)$ and the result holds by Lemma E2.

Suppose that the result holds for $T-1$. Then

$$\begin{aligned}
V_T^m(a,S) &:= C(a,S)+ \\
&\quad \int_0^{y^m(a,S)} V_{T-1}^m(a+1,S+z^\ell)m(z|a,S)dz + (1-r)V_{T-1}^m\left(a, S+[y^m(a,S)]^\ell\right) \\
&= C(a,S)+ \\
&\quad \int_0^{y^m(a,S)} (S+z^\ell)^{1/\ell} m(z|a,S)dz \; V_{T-1}^m(a+1,1) + (1-r)\left[S+[y^m(a,S)]^\ell\right]^{1/\ell} V_{T-1}^m(a,1) \\
&= C(a,S) + \int_0^{y^m(a,S)} \frac{aS^a \ell z^{\ell-1}}{(S+z^\ell)^{a+1-\frac{1}{\ell}}} dz \; V_{T-1}^m(a+1,1) + (1-r)S^{1/\ell}(1-r)^{-1/a\ell} \; V_{T-1}^m(a,1) \\
&= S^{1/\ell}C(a,1) - a\frac{S^a}{(a-\frac{1}{\ell})} \frac{1}{(S+z^\ell)^{a-\frac{1}{\ell}}}\Bigg|_0^{y^m(a,S)} V_{T-1}^m(a+1,1) + S^{1/\ell}(1-r)^{1-1/a\ell} \; V_{T-1}^m(a,1) \\
&= S^{1/\ell}C(a,1) - a\frac{S^{1/\ell}}{(a-\frac{1}{\ell})}\left[1-(1-r)^{1-1/a\ell}\right] V_{T-1}^m(a+1,1) + S^{1/\ell}(1-r)^{1-1/a\ell} \; V_{T-1}^m(a,1) \\
&= S^{1/\ell}\left[C(a,1) - \frac{a\ell}{a\ell-1}\left[1-(1-r)^{1-1/a\ell}\right] V_{T-1}^m(a+1,1) + (1-r)^{1-1/a\ell} \; V_{T-1}^m(a,1)\right],
\end{aligned}$$

where the second equality follows from the inductive hypothesis applied to $V_{T-1}^m(a+1, S+z^\ell)$ and $V_{T-1}^m(a, S+[y^m(a,S)]^\ell)$, and the third equality follows from the definition of $y^m(a,S)$. We deduce that the result holds for T; this completes the induction argument and the proof. □

■ □ ■

D Lower Bound on the MCC

To complement Theorem 3, we next provide a lower bound on the MCC for the case of exponential demand ($\ell = 1$). While this lower bound does not directly apply to the MOG values, it sheds light on whether the upper bound of Theorem 3 is tight in terms of having the "right" parametric dependence.

Theorem D1. *Suppose demands are exponential (i.e., $\ell = 1$) and $a > 1$. Then, for any $T \geq 2$, $S > 0$ and $r \in (0,1)$,*

$$V_T^m(a,S) - V_T^o(a,S) \geq (1-r)\log^2(1-r)\frac{S}{a-1}\left[\log\left(1+\frac{T-2}{a}\right)+\frac{1}{a}-\frac{T-1}{a+T-1}\right]. \tag{D.1}$$

Note that the structure of the lower bound obtained in Theorem D1 is similar to the upper bound given in Theorem 3. Indeed, by specializing to the $\ell = 1$ case in (21), Theorem 3 implies that

$$V_T^m(a,S) - V_T^o(a,S) \;\leq\; K(r,1)\,\frac{\lambda-\lambda^{T+1}}{1-\lambda}\,S\,Q(a-1)\,\left[\log\left(1+\frac{T}{a-1}\right)+\frac{1}{a-1}\right].$$

In particular, the logarithmic dependence with respect to the time horizon T is the best possible dependence one could obtain for $V_T^m(a,S) - V_T^o(a,S)$.

The proof of Theorem D1 is presented later in this section. We next describe the key idea underlying the proof of the result as it leads to additional insights on the extent of information collection limitation induced by censoring.

The key idea resides in the introduction of an alternative problem with a new information structure that arises from censoring of a different nature. In particular, we define the *random rejection problem* to be one in which, during each time period, either the decision maker fully observes the realized demand or receives no information at all (this can be interpreted as the decision maker having access to full demand observations, but lacking access to a fraction of them). This revelation occurs independently of the order size or the demand realization, and is based on i.i.d. draws of a Bernoulli random variable with success probability equal to r. Note that the probability of obtaining no information in any given period is equal to the probability of observing a censored observation in the original problem when a myopic policy is applied. The main difference is that, in the original problem every demand realization provides some level of information, while in the random rejection problem some periods provide less (no) information and some periods provide more.

We develop a lower bound on the difference $V_T^m(a,S) - V_T^o(a,S)$ as follows. We establish that the optimal cost in the random rejection case is always lower than that achieved by the myopic policy in the original censored problem. Given this, one can lower bound $V_T^m(a,S) - V_T^o(a,S)$ by the difference between the cost in the random rejection system and that in the observable case. The latter two costs are much simpler to characterize as the update rules do not involve censoring. (Note that, since the information collected at each time period is independent of the decision rule, the random rejection problem is similar to the observable demand case in that an optimal policy is myopic and minimizes the expected current single-period cost.) This yields the bound in (D.1).

One way to interpret the fact that the performance of the myopic policy in the censored case is worse than that in the random rejection problem is as follows. In the former, the policy collects some information in each period but high realizations of demand in the top $(1-r)^{th}$ quantile of the predictive distribution are censored. In the latter, the decision-maker does not collect any information for a fraction $1-r$ of the periods but in the other periods observes the realization of demand, including those that fall in the top $(1-r)^{th}$ quantile of the predictive distribution. As a result, in this problem, the decision-maker has fewer observations but those include the high realizations of demand. The fact that one performs better in this setting suggests that high demand realizations are more informative than low demand realizations.

Proof of Theorem D1. Suppose demands are exponential and $a > 1$. The proof relies on an alternative informational system. In the *random rejection* system, at each step the decision maker obtains either a full demand observation, independently of the order quantity y and with probability r, or no observation at all, with probability $1-r$. Because in this case the information collection is independent of the decision process, a myopic policy is optimal, and the optimal cost is given by the solution to the Bellman equation

$$V_T^r(a,S) := C(a,S) + r\mathbb{E}\left[V_{T-1}^r(a+1,S+D)\right] + (1-r)V_{T-1}^r(a,S).$$

It is not hard to show (by, for example, following similar steps to those in the proof of Proposition 5 for $\ell = 1$) that the cost function for the random rejection system also satisfies the scalability property, that is,

$$V_T^r(a,S) = SV_T^r(a,1), \quad \text{for any} \quad T \geq 1, S > 0, a > 1.$$

By using this fact and the fact that, for the exponential case, $\mathbb{E}_{a,S}[D] = S/(a-1)$ one can rewrite the recursive formula for $V_T^r(a,S)$ as

$$V_T^r(a,S) = C(a,S) + r\frac{a}{a-1}V_{T-1}^r(a+1,S) + (1-r)V_{T-1}^r(a,S). \tag{D.2}$$

To establish Theorem D1, we will establish the two following inequalities.

a) $V_T^m(a,S) \geq V_T^r(a,S) \geq V_T^o(a,S)$.

b) $V_T^r(a,S) - V_T^o(a,S) \geq (1-r)(-\log(1-r))^2 \dfrac{S}{a-1}\left[\log\left(1+\dfrac{T-2}{a}\right) - \dfrac{T-2}{a+T-2}\right]$.

a) The second inequality follows directly from the fact that full observations is a more informative system than random rejections. We prove the first inequality, $V_T^m(a,S) \geq V_T^r(a,S)$, in two steps.

Step 1. We first establish the following cost relationship in the rejection system.

$$V_T^r(a,S) \geq \frac{a}{a-1} V_T^r(a+1,S).$$

We proceed by induction in T. The case $T=1$ follows from the fact that $V_1^r(a,S) = C(a,S)$ and (Bisi et al., 2011, Lemma 1). Suppose $V_{T-1}^r(a,S) \geq \dfrac{a}{a-1} V_{T-1}^r(a+1,S)$ for all $a>1$. Using the recursive equation for $V_T^r(a,S)$ one can write

$$\begin{aligned}
V_T^r(a,S) &= C(a,S) + r\frac{a}{a-1}V_{T-1}^r(a+1,S) + (1-r)V_{T-1}^r(a,S) \\
\frac{a}{a-1}V_T^r(a+1,S) &= \frac{a}{a-1}C(a+1,S) + r\frac{a}{a-1}\frac{(a+1)}{(a+1)-1}V_{T-1}^r(a+2,S) \\
&\quad + \frac{a}{a-1}(1-r)V_{T-1}^r(a+1,S)
\end{aligned}$$

The base case implies that the first term on the right side of the first equation dominates the first term in the right side of the second equation. Similarly, by the induction hypothesis, the two last terms in the right side of the first equation dominate the corresponding terms in the right side of the second equation and hence the inequality is established.

Step 2. We now prove the inequality we are after, $V_T^m(a,S) \geq V_T^r(a,S)$, by induction on T. If $T=1$ the result follows from the fact that $V_1^m(a,S) = V_1^r(a,S) = C(a,S)$.
Suppose that $V_{T-1}^m(a,S) \geq V_{T-1}^r(a,S)$. Then

$$\begin{aligned}
V_T^m(a,S) &= C(a,S) + \frac{a}{a-1}\left(1-(1-r)^{1-\frac{1}{a}}\right)V_{T-1}^m(a+1,S) + (1-r)^{1-\frac{1}{a}}V_{T-1}^m(a,S) \\
&\geq C(a,S) + \frac{a}{a-1}\left(1-(1-r)^{1-\frac{1}{a}}\right)V_{T-1}^r(a+1,S) + (1-r)^{1-\frac{1}{a}}V_{T-1}^r(a,S),
\end{aligned}$$

where the first equality follows from Proposition 5, page 23, applied to $\ell = 1$.

Using the inequality above and Bellman's recursion for $V_T^r(a,S)$ in (D.2)

$$\begin{aligned}
& V_T^m(a,S) - V_T^r(a,S) \geq \\
& \quad (1-r)\frac{a}{a-1}\left(1-(1-r)^{-\frac{1}{a}}\right)V_{T-1}^r(a+1,S) + (1-r)\left((1-r)^{-\frac{1}{a}}-1\right)V_{T-1}^r(a,S) \\
\geq & \quad (1-r)\frac{a}{a-1}\left(1-(1-r)^{-\frac{1}{a}}\right)V_{T-1}^r(a+1,S) + (1-r)\frac{a}{a-1}\left((1-r)^{-\frac{1}{a}}-1\right)V_{T-1}^r(a+1,S) \\
= & \quad 0
\end{aligned}$$

where the second inequality follows from step 1 above. This completes the induction step and a) is established.

b) We first rewrite the full observation and random rejection value functions for the exponential case:

$$\begin{aligned}
V_T^o(a,S) &= C(a,S) + \frac{a}{a-1}V_{T-1}^o(a+1,S) \\
V_T^r(a,S) &= C(a,S) + r\frac{a}{a-1}V_{T-1}^r(a+1,S) + (1-r)V_{T-1}^r(a,S) \\
&= C(a,S) + \Gamma_{T-1}^r(a,S) + \frac{a}{a-1}V_{T-1}^r(a+1,S),
\end{aligned}$$

where $\Gamma_{T-1}^r(a,S) := (1-r)\left[V_{T-1}^r(a,S) - \frac{a}{a-1}V_{T-1}^r(a+1,S)\right]$.
Therefore,

$$V_T^r(a,S) - V_T^o(a,S) = \Gamma_{T-1}^r(a,S) + \frac{a}{a-1}\left[V_{T-1}^r(a+1,S) - V_{T-1}^o(a+1,S)\right].$$

By proceeding recursively one obtains

$$V_T^r(a,S) - V_T^o(a,S) = \sum_{k=0}^{T-2}\frac{a+k-1}{a-1}\Gamma_{T-k-1}^r(a+k,S). \tag{D.3}$$

Lemma E8, presented in Section E, states that

$$\Gamma_T^r(a,S) \geq \frac{1-r}{a-1}\left[(a-1)C(a,S) - (a+T-1)C(a+T,S)\right]. \tag{D.4}$$

By combining (D.3) and (D.4) one obtains

$$\begin{aligned}
V_T^r(a,S) - V_T^o(a,S) &\geq \frac{1-r}{a-1}\sum_{k=0}^{T-2}[(a+k-1)C(a+k,S) - (a+T-2)C(a+T-1,S)] \\
&\overset{(a)}{\geq} (1-r)\frac{\log^2(1-r)S}{2(a-1)}\sum_{k=0}^{T-2}\left[\frac{1}{a+k} - \frac{1}{a+T-1}\right] \\
&\geq (1-r)\frac{\log^2(1-r)S}{2(a-1)}\left[\sum_{k=0}^{T-2}\frac{1}{a+k} - \frac{T-1}{a+T-1}\right] \\
&\geq (1-r)\frac{\log^2(1-r)S}{2(a-1)}\left[\log(a+T-2) - \log(a) + \frac{1}{a} - \frac{T-1}{a+T-1}\right],
\end{aligned}$$

where (a) follows from Lemma E9(c) in Section E. This completes the proof. □

■ □ ■

E Technical Lemmas

Lemma E1.

$$\begin{aligned}
C(a,S) &= \frac{1}{1-r}\mathbb{E}[D\mathbb{I}_{\{D\geq y^m(a,S)\}}] - E[D], \\
C(\theta) &= \frac{1}{1-r}\mathbb{E}[D\mathbb{I}_{\{D\geq y^m(\theta)\}}|\theta] - E[D|\theta].
\end{aligned}$$

Proof.

$$\begin{aligned}
C(a,S) &= \mathbb{E}[L(y^m(a,S),D)] \\
&= \mathbb{E}[(y^m(a,S)-D)^+] + \frac{r}{1-r}\mathbb{E}[(D-y^m(a,S))^+] \\
&= (y^m(a,S)-\mathbb{E}[D]) + \left(1+\frac{r}{1-r}\right)\mathbb{E}\left[(D-y^m(a,S))^+\right] \\
&= (y^m(a,S)-\mathbb{E}[D]) + \frac{1}{1-r}\Big[\mathbb{E}\left[D\mathbb{I}_{\{D\geq y^m(a,S)\}}\right] - (1-r)y^m(a,S)\Big] \\
&= \frac{1}{1-r}\mathbb{E}[D\mathbb{I}_{\{D\geq y^m(a,S)\}}] - E[D].
\end{aligned}$$

This completes the proof for $C(a,S)$. The proof for $C(\theta)$ is analogous. □

■ □ ■

The following Lemma has been proven in the literature (see, for example, Azoury (1985)), but we include a proof here for completeness:

Lemma E2 (Scalability)**.** *Suppose demands are Weibull and* $a\ell > 1$*, then*

$$C(a,S) = S^{1/\ell} C(a,1).$$

Proof. By Lemma E1,

$$\begin{aligned} C(a,S) &= \frac{1}{1-r}\mathbb{E}[D\mathbb{I}_{\{D \geq y^m(a,S)\}}] - E[D] \\ &= \frac{1}{1-r}\int_{y^m(a,S)}^{\infty} zm(z|a,S)dz - \int_0^{\infty} \overline{M}(z|a,S)dz. \end{aligned}$$

The myopic order quantity is given by $y^m(a,S) = S^{1/\ell}\left((1-r)^{-1/a} - 1\right)^{1/\ell}$ and hence $y^m(a,S) = S^{1/\ell} y^m(a,1)$. Also, one has that

$$\begin{aligned} \int_{y^m(a,S)}^{\infty} zm(z|a,S)dz &= \int_{y^m(a,S)}^{\infty} \frac{aS^a \ell z^\ell}{(S+z^\ell)^{a+1}} dz \\ &= \int_{S^{1/\ell} y^m(a,1)}^{\infty} \frac{a\ell z^\ell}{S(1+\frac{z^\ell}{S})^{a+1}} dz \\ &= S^{1/\ell} \int_{y^m(a,1)}^{\infty} \frac{a\ell z^\ell}{(1+z^\ell)^a} dz \\ &= S^{1/\ell} \int_{y^m(a,1)}^{\infty} zm(z|a,1)dz. \end{aligned}$$

Similarly, it is not hard to show that

$$\int_0^{\infty} \overline{M}(z|a,S)dz = S^{1/\ell} \int_0^{\infty} \overline{M}(z|a,1)dz,$$

and hence

$$C(a,S) = S^{1/\ell}\left[\frac{1}{1-r}\int_{y^m(a,1)}^{\infty} zm(z|a,1)dz - \int_0^{\infty} \overline{M}(z|a,1)dz\right] = S^{1/\ell} C(a,1).$$

This completes the proof. □

■ □ ■

Lemma E3. *Let* (D,θ) *follow a Newsvendor distribution with parameters* (a,S)*, and* (D',θ) *a*

Newsvendor distribution with parameters $(a, S + d(y))$*. For any* $x, y \geq 0$*,*

$$\mathbb{P}\big(d(D) \geq x | D \geq y\big) = \mathbb{P}\big(d(y) + d(D') \geq x\big).$$

In words, this lemma states that the conditional distribution of $d(D)$, given that D is greater than y, is equivalent to the unconditional distribution of $d(y)+d(D')$, where D' follows the resulting distribution after having one censored observation at y.

Proof. We start by showing the following equivalence:

$$\begin{aligned} & \mathbb{P}\big(d(D) \geq x | D \geq y\big) = \mathbb{P}\big(d(y) + d(D') \geq x\big) \quad \forall x, y \geq 0 & \text{(E.1)} \\ \Leftrightarrow \quad & \mathbb{P}\big(d(D) \geq d(x') | D \geq y'\big) = \mathbb{P}\big(d(y') + d(D') \geq d(x')\big) \quad \forall x', y' \geq 0. & \text{(E.2)} \end{aligned}$$

To show necessity note that, given x, if there exists x' such that $d(x') = x$ the result follows immediately by setting $x' := d^{-1}(x)$ and $y' := y$. If no such x' exists, this means, since $d(\cdot)$ is increasing and unbounded, that $x < d(0)$. But this implies that the events $\{d(D) \geq x\}$ and $\{d(y) + d(D') \geq x\}$ equal the entire sample space, and hence both sides of equation (E.1) equal 1. This completes the proof for the necessity implication; the sufficiency can be proven by analogous arguments.

We now give a proof of equation (E.2). If $x' < y'$ both sides of equation (E.2) are equal to 1 and the result holds. Suppose that $x' \geq y'$. Then

$$\begin{aligned} \mathbb{P}\big(d(D) \geq d(x') | D \geq y'\big) &= \frac{\mathbb{P}(d(D) \geq d(x'))}{\mathbb{P}(D \geq y')} \\ &= \int_\Theta \frac{\mathbb{P}(d(D) \geq d(x') | \theta)}{\mathbb{P}(D \geq y')} \pi(\theta | a, S) d\theta \\ &= \int_\Theta \frac{\mathbb{P}(d(D) \geq d(x') | \theta)}{\mathbb{P}(D \geq y' | \theta)} \frac{\mathbb{P}(D \geq y' | \theta) \pi(\theta | a, S)}{\mathbb{P}(D \geq y')} d\theta, \qquad \text{(E.3)} \end{aligned}$$

where the first equality comes from the fact that the events $\{D \geq y'\}$ and $\{d(D) \geq d(y')\}$ are equivalent, and $x' \geq y'$ implies $d(x') \geq d(y')$.

Note that the second fraction inside the integral in (E.3) is equivalent to the posterior distribution of θ, given a censored observation of $D \geq y$. That is,

$$\frac{\mathbb{P}(D \geq y' | \theta) \pi(\theta | a, S)}{\mathbb{P}(D \geq y')} = \pi(\theta | a, S + d(y')),$$

where the equality follows from the Bayes update rule of a censored observation in the Newsvendor family. By replacing this expression in (E.3) one obtains

$$\begin{aligned}
\int_\Theta \frac{\mathbb{P}(d(D) \geq d(x')|\theta)}{\mathbb{P}(D \geq y'|\theta)} \frac{\mathbb{P}(D \geq y'|\theta)\pi(\theta|a,S)}{\mathbb{P}(D \geq y')} d\theta &= \int_\Theta \frac{\mathbb{P}(d(D) \geq d(x')|\theta)}{\mathbb{P}(D \geq y'|\theta)} \pi(\theta|a, S + d(y'))d\theta \\
&= \int_\Theta e^{-\theta(d(x')-d(y'))}\pi(\theta|a, S + d(y'))d\theta \\
&= \int_\Theta \mathbb{P}\big(d(D) \geq d(x') - d(y')|\theta\big)\pi(\theta|a, S + d(y'))d\theta \\
&= \mathbb{P}\big(d(y') + d(D') \geq d(x')\big),
\end{aligned}$$

where the third equality comes from the fact that the distribution of $d(D)$ conditional on θ is exponential with parameter θ. This completes the proof. □

■ □ ■

Lemma E4. *For any $t \geq 0$ we have*

$$C_t^m(a, S) \geq C_t^o(a, S).$$

Proof. To simplify the exposition, we will write the proof for $t = 1$, that is, one period in the future. The extension to general t follows with similar reasoning.

Suppose D_1 represents the demand realization in the first period. One can write the future costs as

$$\begin{aligned}
C_1^m(a, S) &= \mathbb{E}\left[\min_y \mathbb{E}\left[L(y, D)\big|D_1 \wedge y^m\right]\right] \\
C_1^o(a, S) &= \mathbb{E}\left[\min_y \mathbb{E}\left[L(y, D)\big|D_1\right]\right].
\end{aligned} \tag{E.4}$$

Note that if we condition the outer expectation in (E.4) on the event $\{D_1 < y^m\}$ we obtain

$$\mathbb{E}\left[\min_y \mathbb{E}\left[L(y, D)\big|D_1\right]\bigg| D_1 < y^m\right] = \mathbb{E}\left[\min_y \mathbb{E}\left[L(y, D)\big|D_1 \wedge y^m\right]\bigg| D_1 < y^m\right]. \tag{E.5}$$

Conditioning on the complementary event in (E.4) one obtains

$$\begin{aligned}\mathbb{E}\left[\min_y \mathbb{E}[L(y,D)\big|D_1]\Big|D_1 \geq y^m\right] &\leq \min_y \mathbb{E}\left[\mathbb{E}[L(y,D)\big|D_1]\Big|D_1 \geq y^m\right] \\ &= \min_y \mathbb{E}[L(y,D)|D_1 \geq y^m] \\ &= \mathbb{E}\left[\min_y \mathbb{E}[L(y,D)|D_1 \geq y^m]\Big|D_1 \geq y^m\right] \\ &= \mathbb{E}\left[\min_y \mathbb{E}[L(y,D)|D_1 \wedge y^m]\Bigg|D_1 \geq y^m\right], \qquad \text{(E.6)}\end{aligned}$$

where the first equality follows from the tower property; the second equality follows from the fact $\min_y \mathbb{E}[L(y,D)|D_1 \wedge y^m, D_1 \geq y^m]$ is measurable with respect to the sigma algebra generated by $D_1 \geq y^m$; and the last equality follows from the fact that on the events where $D_1 \geq y^m$, $\mathbb{E}[L(y,D)|D_1 \wedge y^m] = \mathbb{E}[L(y,D)|D_1 \geq y^m]$. By combining (E.5) and (E.6), one obtains

$$C_1^o(a,S) = \mathbb{E}\left[\min_y \mathbb{E}\left[L(y,D)\big|D_1\right]\right] \leq \mathbb{E}\left[\min_y \mathbb{E}\left[L(y,D)\big|D_1 \wedge y^m\right]\right] = C_1^m(a,S).$$

This completes the proof. □

■ □ ■

Lemma E5. *Suppose Demand is Weibull, with $a\ell > 1$. For any a, S, l, r and $t = 1, \ldots, T$*

$$C_t^o(a,S) - C_\infty^o(a,S) \leq S^{1/\ell} K(r,\ell) \left[\frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell}\right]^{1/\ell} \frac{1}{a+t-1/\ell}.$$

where $K(r,\ell)$ is a constant depending only on r and l.

Proof. Let $(D_1, \ldots, D_T)$ denote the vector of demands, over the time horizon, and let $S_t := S + D_1^\ell + \ldots + D_t^\ell$ and $a_t := a + t$.
By Lemma E1 one has

$$\begin{aligned}C_t^o(a,S) &= \frac{1}{1-r}\mathbb{E}_{a,S}\left[\int_{y^m(a_t,S_t)}^\infty x m(x|a_t,S_t)dx\right] - \mathbb{E}_{a,S}\left[\mathbb{E}_{a_t,S_t}[D]\right] \\ C_\infty^o(a,S) &= \frac{1}{1-r}\mathbb{E}_{a,S}\Big[\int_{y^m(\theta)}^\infty x f(x|\theta)dx\Big] - \mathbb{E}_{a,S}\left[\mathbb{E}[D|\theta]\right].\end{aligned}$$

Note that the last terms on the right side of both equations above are equal to $\mathbb{E}_{a,S}[D]$ by the law of total expectation. Therefore, by subtracting both equations one obtains

$$\begin{aligned}
&(1-r)[C_t^o(a,S)-C_\infty^o(a,S)]\\
&= \mathbb{E}_{a,S}\left[\int_{y^m(a_t,S_t)}^{\infty} xm(x|a_t,S_t)dx\right]-\mathbb{E}_{a,S}\left[\int_{y^m(\theta)}^{\infty} xf(x|\theta)dx\right]\\
&= \mathbb{E}_{a,S}\left[\int_{y^m(a_t,S_t)}^{\infty} x\frac{a_tS_t^{a_t}\ell x^{\ell-1}}{(S_t+x^\ell)^{a_t+1}}dx\right]-\mathbb{E}_{a,S}\left[\int_{y^m(\theta)}^{\infty} x\theta\ell x^{\ell-1}e^{-\theta x^\ell}dx\right]\\
&= a_t\mathbb{E}_{a,S}\left[S_t^{1/\ell}\right]\int_{(1-r)^{-\frac{1}{a_t}}-1}^{\infty}\frac{u^{1/\ell}}{(1+u)^{a_t+1}}du-\mathbb{E}_{a,S}\left[\theta^{-1/\ell}\right]\int_{-\log(1-r)}^{\infty}u^{1/\ell}e^{-u}du, \qquad \text{(E.7)}
\end{aligned}$$

where the last equality follows from the change of variable $u := x^\ell/S_t$ and $u := \theta x^\ell$ in the first and second integrals, respectively.

Note the law of iterated expectations and an application of Jensen's inequality yield that

$$\mathbb{E}_{a,S}\left[\theta^{-1/\ell}\right] \;=\; \mathbb{E}_{a,S}\left[\mathbb{E}_{a_t,S_t}\left[\theta^{-1/\ell}\right]\right] \geq \mathbb{E}_{a,S}\left[\mathbb{E}_{a_t,S_t}\left[\theta\right]^{-1/\ell}\right] = \mathbb{E}_{a,S}\left[\frac{S_t^{1/\ell}}{a_t^{1/\ell}}\right].$$

Returning to (E.7), one obtains

$$\begin{aligned}
&(1-r)\left[C_t^o(a,S)-C_\infty^o(a,S)\right]\leq\\
&\quad \mathbb{E}_{a,S}\left[\left(\frac{S_t}{a_t}\right)^{1/\ell}\right]\left[a_t^{1/\ell+1}\int_{(1-r)^{-\frac{1}{a_t}}-1}^{\infty}\frac{u^{1/\ell}}{(1+u)^{a_t+1}}du-\int_{-\log(1-r)}^{\infty}u^{1/\ell}e^{-u}du\right]\\
&= \mathbb{E}_{a,S}\left[\left(\frac{S_t}{a_t}\right)^{1/\ell}\right]\left[\left(\frac{a_t}{a_t+1}\right)^{1/\ell+1}\int_{(a_t+1)\left((1-r)^{-\frac{1}{a_t}}-1\right)}^{\infty}\frac{u^{1/\ell}}{(1+\frac{u}{a_t+1})^{a_t+1}}du-\int_{-\log(1-r)}^{\infty}u^{1/\ell}e^{-u}du\right]\\
&\leq \mathbb{E}_{a,S}\left[\left(\frac{S_t}{a_t}\right)^{1/\ell}\right]\left[\int_{-\log(1-r)}^{\infty}\frac{u^{1/\ell}}{(1+\frac{u}{a_t+1})^{a_t+1}}du-\int_{-\log(1-r)}^{\infty}u^{1/\ell}e^{-u}du\right]\\
&= \mathbb{E}_{a,S}\left[\left(\frac{S_t}{a_t}\right)^{1/\ell}\right]\left[\int_{-\log(1-r)}^{\infty}u^{1/\ell}\left(\left(1+\frac{u}{a_t+1}\right)^{-(a_t+1)}-e^{-u}\right)du\right]\\
&\leq \mathbb{E}_{a,S}\left[\left(\frac{S_t}{a_t}\right)^{1/\ell}\right]\left[\int_{0}^{\infty}u^{1/\ell}\left(\left(1+\frac{u}{a_t+1}\right)^{-(a_t+1)}-e^{-u}\right)du\right]\\
&= \mathbb{E}_{a,S}\left[\left(\frac{S_t}{a_t}\right)^{1/\ell}\right]\left[(a_t+1)^{1/\ell+1}\int_{0}^{\infty}\frac{u^{1/\ell}}{(1+u)^{a_t+1}}du-\int_{0}^{\infty}u^{1/\ell}e^{-u}du\right]. \qquad \text{(E.8)}
\end{aligned}$$

The second inequality follows from the fact that $a_t/(a_t+1) < 1$ and the fact that $(a_t+1)\left[(1-r)^{-\frac{1}{a_t}}-1\right] > a_t\left[(1-r)^{-\frac{1}{a_t}}-1\right] > -\log(1-r) > 0$ (see, for example, Lemma E9.a). The last inequality follows from the fact that the integrand is non-negative (noting that $e^{-u} \leq e^{-(a_t+1)\ln(1+u/(a_t+1))}$ using that

$\ln(1+x) \le x$ for all $x > 0$), and the last equality from a change of variable.

Consider now the last term in square braces. One can rewrite the integrals using the Gamma function as follows

$$\begin{aligned}(a_t+1)^{1/\ell+1}\int_0^\infty \frac{u^{1/\ell}}{(1+u)^{a_t+1}}du - \int_0^\infty u^{1/\ell}e^{-u}du &= (a_t+1)^{1/\ell+1}\frac{\Gamma(1/\ell+1)\Gamma(a_t-1/\ell)}{\Gamma(a_t+1)} - \Gamma(1/\ell+1) \\ &= \Gamma(1/\ell+1)\left[\frac{(a_t+1)^{1/\ell+1}\Gamma(a_t-1/\ell)}{\Gamma(a_t+1)} - 1\right] \\ &\le \Gamma(1/\ell+1)\widehat{K}(1/\ell+1)\frac{1}{a_t-1/\ell}, \qquad \text{(E.9)}\end{aligned}$$

where the first equality follows from basic properties of the Gamma and Beta functions[4] (see, for example, Abramowitz et al. (1964), Chapter 6) and the last inequality follows from Lemma E6. In particular, $\widehat{K}(\cdot)$ is defined in (E.11).

Therefore (E.8) and (E.9) yield

$$C_t^o(a,S) - C_\infty^o(a,S) \le \mathbb{E}_{a,S}\left[\left(\frac{S_t}{a_t}\right)^{1/\ell}\right]\frac{\Gamma(1/\ell+1)\widehat{K}(1/\ell+1)}{1-r}\frac{1}{a_t-1/\ell}. \qquad \text{(E.10)}$$

Furthermore, it is possible to show (see Lemma E7) that

$$\mathbb{E}_{a,S}\left[\left(\frac{S_t}{a_t}\right)^{1/\ell}\right] \le \left[S\frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell}\right]^{\frac{1}{\ell}}$$

and hence, by setting

$$K(r,\ell) := \frac{\Gamma(1/\ell+1)\widehat{K}(1/\ell+1)}{1-r},$$

one obtains

$$C_t^o(a,S) - C_\infty^o(a,S) \le S^{1/\ell}K(r,\ell)\left[\frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell}\right]^{1/\ell}\frac{1}{a_t-1/\ell}.$$

This completes the proof. □

■ □ ■

[4] In particular we are using the definition of the Gamma function: $\Gamma(x) := \int_0^\infty u^{x-1}e^{-u}du$, the following properties of the beta function: $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = \int_0^\infty \frac{u^{x-1}}{(u+1)^{x+y}}du$, and letting $x := 1/\ell + 1$ and $y := a_t - 1/\ell$.

Lemma E6. *Let a, b be positive real numbers such that $a > b$. Then*

$$\frac{a^b\Gamma(a-b)}{\Gamma(a)} - 1 \leq \frac{\widehat{K}(b)}{a-b},$$

where $\widehat{K}(\cdot)$ is given by

$$\widehat{K}(x) = \begin{cases} 0 & \text{if } x = 0, \\ x\left(\widehat{K}(x-1)+1\right) & \text{if } x \in \mathbb{N}, \\ \widehat{K}(\lfloor x \rfloor) + x - \lfloor x \rfloor & \text{if } x \in \mathbb{R} \setminus \mathbb{N}. \end{cases} \tag{E.11}$$

Proof. *i.*) If $0 \leq b < 1$ the results follows directly from a traditional bound on the gamma function, first developed by Wendel (1948) (a proof can be found in Qi and Losonczi (2010)).[5]

ii.) Suppose $b = n \in \mathbb{N}$, the result follows from the inequality (E.12) we establish next and the factorial property of the gamma function: $\Gamma(x) = (x-1)\Gamma(x-1) \ \ \forall x \geq 1$.

Let $n \in \mathbb{N}$ and $a \in \mathbb{R}_+$ such that $a > n$. We establish the following inequality by induction on n:

$$\frac{a^n}{(a-1)(a-2)\dots(a-n)} - 1 \leq \frac{\widehat{K}(n)}{a-n}, \tag{E.12}$$

where $\widehat{K}(n)$ was defined in (E.11). The base case, $n = 1$, is trivial. Suppose the inequality holds for $n-1$, then

$$\begin{aligned}
\frac{a^n}{(a-1)(a-2)\dots(a-n)} - 1 &= \frac{a^{n-1}}{(a-1)(a-2)\dots(a-(n-1))}\frac{a}{a-n} - 1 \\
&\overset{(a)}{\leq} \left(\frac{\widehat{K}(n-1)}{a-n+1} + 1\right)\frac{a}{a-n} - 1 \\
&= \frac{\left(\widehat{K}(n-1) + a - n + 1\right)a - (a-n+1)(a-n)}{(a-n+1)(a-n)} \\
&= \frac{a\widehat{K}(n-1) + (a-n+1)n}{(a-n+1)(a-n)} \\
&\overset{(b)}{=} \frac{1}{a-n}\left[\frac{a/n}{a-n+1}(\widehat{K}(n) - n) + n\right] \\
&\overset{(c)}{\leq} \frac{1}{a-n}\widehat{K}(n),
\end{aligned}$$

[5] The original bound is written as: $\left(\frac{x}{x+s}\right)^{(1-s)} \leq \frac{\Gamma(x+s)}{x^s\Gamma(x)}$ for any $0 < s < 1$ and $x > 0$. Case *i.*) can be derived from this bound by letting $x := a - b$ and $s := b$.

where (a) is a consequence of the induction hypothesis, (b) follows by the recursive definition of $\widehat{K}(n)$ and (c) is a result of the fact that $a/n \leq a - n + 1$ for all $a \geq n \geq 0$. This concludes the induction argument.

iii.) Finally, suppose $b \in \mathbb{R} \setminus \mathbb{N}$ and $b > 1$. Let $n := \lfloor b \rfloor$ and $\epsilon := b - \lfloor b \rfloor$. Then

$$
\begin{aligned}
\frac{a^b \Gamma(a-b)}{\Gamma(a)} &= \frac{a^{n+\epsilon}\Gamma(a-n-\epsilon)}{\Gamma(a)} \\
&= \frac{a^n}{(a-1)(a-2)\dots(a-n)} \frac{a^\epsilon \Gamma(a-n-\epsilon)}{\Gamma(a-n)} \\
&\leq \left(\frac{\widehat{K}(n)}{a-n} + 1\right) \frac{a^\epsilon \Gamma(a-n-\epsilon)}{\Gamma(a-n)} \\
&\leq \left(\frac{\widehat{K}(n)}{a-n} + 1\right)\left(1 + \frac{\epsilon}{a-n-\epsilon}\right) \\
&= \frac{\widehat{K}(n)}{a-n-\epsilon} + \frac{a-n}{a-n-\epsilon} \\
&= \frac{\widehat{K}(n)+\epsilon}{a-n-\epsilon} + 1 \\
&= \frac{\widehat{K}(n) + b - \lfloor b \rfloor}{a-b} + 1 \\
&= \frac{\widehat{K}(b)}{a-b} + 1
\end{aligned}
$$

where in the first inequality, we have use (E.12) and in the second inequality, we have used the fact that $\epsilon < 1$. This completes the proof. □

■ □ ■

Lemma E7. *Fix $S > 0$ and $a > 1/\ell$. Let $D_1, \dots, D_t$ be i.i.d. random variables, Weibull distributed with parameter $\theta \sim Gamma(a, S)$. Then*

$$
\mathbb{E}\left[\left(\frac{S + D_1^\ell + \dots + D_t^\ell}{a+t}\right)^{1/\ell}\right] \leq \left[S \frac{\exp\{1/(a-1/\ell)\}}{a - 1/\ell}\right]^{\frac{1}{\ell}}.
$$

Proof. First, recall that

$$
\mathbb{E}_D[(S + D^\ell)^{1/\ell}] = \int_0^\infty (S + z^\ell) m(z|a,S) dz = S^{1/\ell} \int_0^\infty (1 + z^\ell) m(z|a,1) dz = S^{1/\ell} \frac{a\ell}{a\ell - 1}.
$$

Therefore, one has that

$$\begin{aligned}\mathbb{E}\left[\left(S+D_1^\ell+\ldots+D_t^\ell\right)^{1/\ell}\right] &= \mathbb{E}\left[\mathbb{E}\left[\left(S+D_1^\ell+\ldots+D_t^\ell\right)^{1/\ell}\Big|D_1,\ldots,D_{t-1}\right]\right] \\ &= \mathbb{E}\left[\left(S+D_1^\ell+\ldots+D_{t-1}^\ell\right)^{1/\ell}\right]\int_0^\infty(1+z^\ell)m(z|a+t-1,1)dz \\ &= \mathbb{E}\left[\left(S+D_1^\ell+\ldots+D_{t-1}^\ell\right)^{1/\ell}\right]\frac{(a+t-1)\ell}{(a+t-1)\ell-1}.\end{aligned}$$

Continuing recursively one obtains

$$\begin{aligned}\mathbb{E}\left[\left(S+D_1^\ell+\ldots+D_t^\ell\right)^{1/\ell}\right] &= S^{\frac{1}{\ell}}\frac{a\ell}{a\ell-1}\frac{(a+1)\ell}{(a+1)\ell-1}\cdots\frac{(a+t-1)\ell}{(a+t-1)\ell-1} \\ &= S^{\frac{1}{\ell}}\frac{a}{a-\frac{1}{\ell}}\frac{a+1}{a+1-\frac{1}{\ell}}\cdots\frac{a+t-1}{a+t-1-\frac{1}{\ell}} \\ &= S^{\frac{1}{\ell}}\prod_{k=0}^{t-1}\left(1+\frac{1}{\ell a-1+\ell k}\right) \\ &\le S^{\frac{1}{\ell}}\exp\left\{\sum_{k=0}^{t-1}\frac{1}{\ell a-1+\ell k}\right\} \\ &\le S^{\frac{1}{\ell}}\exp\left\{\frac{1}{\ell}\left[\log(\ell a-1+\ell(t-1))-\log(\ell a-1)+\frac{1}{a-1/\ell}\right]\right\} \\ &= S^{\frac{1}{\ell}}\left[1+\frac{(t-1)}{a-1/\ell}\right]^{\frac{1}{\ell}}\left[e^{\frac{1}{a-1/\ell}}\right]^{\frac{1}{\ell}}.\end{aligned}$$

And hence

$$\mathbb{E}_{\hat{D}}\left[\left(\frac{S+D_1^\ell+\ldots+D_t^\ell}{a+t}\right)^{1/\ell}\right]\le S^{\frac{1}{\ell}}\left[\frac{a+t-1-1/\ell}{(a-1/\ell)(a+t)}\right]^{\frac{1}{\ell}}\left[e^{\frac{1}{a-1/\ell}}\right]^{\frac{1}{\ell}}\le\left[S\frac{\exp\{1/(a-1/\ell)\}}{a-1/\ell}\right]^{\frac{1}{\ell}},$$

which concludes the proof. □

■ □ ■

Lemma E8. *Suppose demand are exponential. Let*

$$\Gamma_T^r(a,S):=(1-r)\left[V_T^r(a,S)-\frac{a}{a-1}V_T^r(a+1,S)\right].$$

Then, for any $a>1$, $T\ge1$, $S>0$ and $r\in(0,1)$

$$\Gamma_T^r(a,S)\ge\frac{1-r}{a-1}\left[(a-1)C(a,S)-(a+T-1)C(a+T,S)\right].$$

Proof. We proceed by induction on T. The base case, $T = 1$, follows directly from the definition of $\Gamma_1^r(a, S)$. Suppose the result holds for $T - 1$. Then

$$
\begin{aligned}
\frac{\Gamma_T^r(a,S)}{1-r} &= V_T^r(a,S) - \frac{a}{a-1} V_T^r(a+1,S) \\
&= C(a,S) + r\frac{a}{a-1} V_{T-1}^r(a+1,S) + (1-r)V_{T-1}^r(a,S) - \\
&\quad \frac{a}{a-1}\left[C(a+1,S) + r\frac{a+1}{a} V_{T-1}^r(a+2,S) + (1-r)V_{T-1}^r(a+1,S)\right] \\
&= \frac{\Gamma_1^r(a,S)}{1-r} + \frac{r}{1-r}\frac{a}{a-1}\Gamma_{T-1}^r(a+1,S) + \Gamma_{T-1}^r(a,S) \\
&\geq \frac{1}{a-1}\left[(a-1)C(a,S) - aC(a+1,S)\right] + \frac{r}{a-1}\left[aC(a+1,S) - (a+T-1)C(a+T,S)\right] + \\
&\quad \frac{(1-r)}{a-1}\left[(a-1)C(a,S) - (a+T-2)C(a+T-1,S)\right] \\
&= \frac{1}{a-1}\left[(a-1)C(a,S) - (a+T-1)C(a+T,S)\right] + \\
&\quad \frac{1-r}{a-1}\Big[(a-1)C(a,S) - aC(a+1,S) + (a+T-1)C(a+T,S) - (a+T-2)C(a+T-1,S)\Big] \\
&\geq \frac{1}{a-1}\left[(a-1)C(a,S) - (a+T-1)C(a+T,S)\right],
\end{aligned}
$$

where the first inequality follows from the inductive hypothesis and the second inequality follows from the fact that the function $f(x) := (x-1)C(x,S)$ is a decreasing convex function (see Lemma E9 below). This completes the induction step. □

■ □ ■

Lemma E9. *Let* $f : (1, +\infty) \to \mathbb{R}$ *be defined as* $f(a) = a\left[(1-r)^{-1/a} - 1\right]$ *and note that* $f(a) = (a-1)C(a,1)$ *when demand is exponential. Then, for every* $a > 1$

a) $f(a) + \log(1-r) = \sum_{k=2}^{\infty} \frac{(-\log(1-r))^k}{k!}\frac{1}{a^{k-1}}$ *for any* $a > 1$.

b) $f(\cdot)$ *is decreasing and convex.*

c) $f(a) - f(b) \geq \frac{\log^2(1-r)}{2}\left[\frac{1}{a} - \frac{1}{b}\right]$, *for any* $b \geq a$.

Proof. **a)**

$$\begin{aligned}
f(a) + \log(1-r) &= a\big[(1-r)^{-1/a} - 1\big] + \log(1-r) \\
&= a\big[e^{-\log(1-r)/a} - 1 - \frac{-\log(1-r)}{a}\big] \\
&= a\sum_{k=2}^{\infty}\left[\frac{(-\log(1-r))^k}{k!}\frac{1}{a^k}\right] \\
&= \sum_{k=2}^{\infty}\left[\frac{(-\log(1-r))^k}{k!}\frac{1}{a^{k-1}}\right]
\end{aligned}$$

where the third equality follows from the Taylor exp of e^x.

b) Follows directly from a), since the terms in the sum are positive, decreasing and convex functions.

c) Follows from a) by subtracting the two series and discarding all terms except $k = 2$.

□

■ □ ■