Admission Control to Queueing Systems with Arrival Forecast Information

Xiaoshan Peng*

Kelley School of Business, Indiana University, Bloomington, IN 47405,

Problem definition: This paper examines an admission control problem in a single-class queue using a general arrival forecast model. The arrival process is modeled as a general point process with an associated forecast model. Upon each customer's arrival, the system manager evaluates the current state and all available information to decide whether to admit or reject the incoming customer. Rejected customers leave the system immediately, while admitted customers join the queue and wait for service. The system incurs a waiting cost of h per time unit for each customer in the system, and a one-time penalty of p for each rejected customer. **Methodology/results:** We model the admission control problem with a general forecast model as a multi-stage stochastic programming problem. We derive an optimality condition characterized by an adjoint process, which is the optimal solution to an associated dual problem. **Managerial implications:** The optimality condition offers a unified approach to verify the optimality of admission policies under various information scenarios, including both no information and full information cases. Additionally, we propose a simulation-based heuristic that incorporates the expected remaining congestion time, resulting in admission policies that perform well with both accurate and noisy forecasts in numerical analysis.

Key words: admission control problem, forecast model, queueing system

---

*Email: xp1@iu.edu

1 Introduction

There has been a growing body of literature on developing predictive models to forecast customer arrival volumes in service systems. The arrival volumes can be more accurately forecasted thanks to the development of predictive models and the increasing easiness of collecting and analyzing the arrival data nowadays. By analyzing the datasets from call centers and healthcare systems, recent studies have developed various statistical or machine learning models to forecast the arrival volumes of service systems; see references in the survey paper by Ibrahim et al. (2016). Such predictive models can be utilized to improve capacity planning, staffing, and customer routing decisions in these systems. In this paper, we present an approach to study how to incorporate predictive models in dynamic decision making processes to improve system performance.

This paper studies how to incorporate the predictive models in an admission control problem of a canonical single-server queueing model. The admission control problem has wide applications in various industries, e.g., see Gurvich and Perry (2012) in service systems, Deo and Gurvich (2011) for applications in healthcare systems, and Spencer et al. (2014) for applications in telecommunication. Most papers studying the admission control problem in queueing theory assume that the system is stationary; see Stidham (1985) and Stidham (2002). Specifically, the interarrival and service times are commonly modeled as independent and identically distributed (i.i.d.) random variables. However, this assumption is often inconsistent with observations from the analysis of arrival data. The literature has well documented that the arrival processes in call centers and healthcare systems exhibit predictable seasonality patterns and unpredictable fluctuations; see Ibrahim et al. (2016a). Moreover, the extensive analysis of the data from three different service systems in Oreshkin et al. (2016) shows that none of the predictive models of the arrivals they tested is universally good, implying that the choice of models for the primitive processes should be context-based. Thus, it is critical to use an analytical framework to study the admission control problem that can accommodate general arrival and service models.

Limited attempts have been made to study the admission control problem with general arrival and service models. The structure of the optimal policies is known for two particular information scenarios: no and full future information scenarios. In the no future information scenario, the number of incoming customers and the service capacity in each period are modeled as i.i.d. random variables. In the full information case, the system manager knows the exact number of arrivals and service capacities in the future. In other words, the system is a non-stationary deterministic system. In previous literature, the optimal policies of these two information scenarios are derived using very different techniques. The optimal policy of the no future information case is analyzed by formulating the problem as a Markov decision process (MDP), while the one of the full future information case is derived via a pathwise analysis; see Spencer et al. (2014) and Ata and Peng (2020), etc. A few studies attempt to characterize the optimal policies for a non-stationary system with uncertainty. Lewis et al. (2002) study the admission control problem when the arrival process follows a non-homogeneous Poisson process with a deterministic and time-dependent arrival rate function. The authors prove that the optimal policy is a threshold-type of policy, in which the system manager admits incoming customers only when the number of customers in the system does not exceed a (time-dependent) threshold. However, little was known about the structure of the threshold function for a general non-stationary system, which is commonly seen in practice. Lewis et al. (2002) provide examples and counter-examples showing that the threshold may or may not be monotonic in the (current) arrival rate.

This paper considers a discrete time single-server queue. In each period, new customers arrive in the system, and the system may complete a few service requests if there are any. The arrival counts and the service capacity in each period follow general stochastic processes. At the beginning of each period, customers arrive at the system. The system manager observes the system's current state and collects all available information about the system, including the number of incoming customers and the service capacity in the current period. The system manager then uses all collected information to forecast future arrival volumes. She then decides the number of customers accepted to the system and rejects others based on the forecasts. If an incoming customer is rejected, the customer leaves the system immediately. The system incurs a penalty for each rejected customer. If an incoming customer is admitted to the system, the customer enters service if the server is idle. If the server is busy, the admitted customers join the queue and wait for the service. The system incurs a holding cost for each customer in the system at the end of each period. The system seeks to find the optimal admission policy that minimizes the expected total cost incurred over a finite horizon.

We formulate the admission control problem with the general arrival and service process as a multi-stage stochastic programming problem. The forecasting model is embedded within the information structure of the model. Specifically, we use a filtration associated with the stochastic programming problem to characterize the information available to the system manager for making admission decisions in each period. The admission policy must be non-anticipative with respect to the associated filtration. By appropriately defining the filtration, the model can incorporate a wide range of predictive models and information scenarios.

We use a dual approach to derive an optimality condition of the optimal admission policy. Specifically, we formulate a dual problem of the admission control problem, considered as the primal problem. We then characterize the optimality condition using the complementary slackness conditions jointly satisfied by the optimal admission decisions (primal decision variables) and the optimal adjoint process (dual decision variables). The optimality condition provides a unified approach to verify the optimality of the admission policies for general information scenarios, including the no and full information scenarios, analyzed using two different approaches, MDP and pathwise analysis in the literature. We also propose a lower bound for the performance of any given policies by using the weak duality result as a side product from deriving the optimality condition.

For a general information scenario, it is challenging to characterize the optimal policy analytically. In practice, the system manager may prefer to implement parsimonious policies with reasonably good performance than the optimal policy if its structure is complex. We propose a simulation-based heuristic that flexibly incorporates general forecast models. This heuristic admits incoming customers when the expected remaining congestion time of the system, which can be computed efficiently through simulations, is below a given threshold under a pre-specified admission policy.

We show that the proposed heuristic performs well numerically in various settings. In particular, the proposed heuristic is near optimal with limited but accurate forecasts in stationary systems. It also performs well in non-stationary systems with both accurate and noisy forecasts. In the numerical studies, we observe that both accurate and noisy future information effectively reduces the operational cost. However, the marginal value of future information decreases as the lookahead window expands.

The rest of this paper is organized as follows. Section 2 reviews the related literature. Section 3 introduces the queueing model to study the admission control problem, with Section 3.1 describing the model, Section 3.2 discussing the information structure and Section 3.3 defining the stochastic multi-stage programming model to study the problem. Section 4 derives the optimality condition that the optimal admission control policy needs to satisfy. It also provides useful properties of the optimal (dual) adjoint process and uses these properties to derive the optimal admission policies of special arrival models. Section 4.4 proposes a simulation-based heuristic. Section 5 presents the numerical study and Section 6 provides a few concluding remarks. All proofs are provided in the appendix.

2 Literature Review

The literature on developing predictive models for the arrival process in service systems are growing. For example, Shi et al. (2016) studies a discharge policy in an emergency department. Green and Kolesa (1991) propose the pointwise stationary approximations (PSA) to model the time-varying demand. Whitt (2018) provides an extensive review of research on the performance analysis of queueing systems with time-varying arrival rates; see also Gans et al. (2003), Green et al. (2007) and Aksin et al. (2007) and the references therein. In particular, the recent literature on call centers have shown that the within-day call arrival process has shown two important properties: Time dependence of call arrival rates and over-dispersion of arrival counts. Ibrahim et al. (2016a) provides an extensive review of predictive models for analyzing call center arrivals; see also Ibrahim and L'Ecuyer (2013), Brown et al. (2005), Avramidis et al. (2004) for examples in call centers. Sun et al. (2009) and Jones et al. (2009) develop predictive models based on time-series analysis to predict ED workload. The focus of this paper is developing a model framework that incorporates these predictive models rather than complimenting this stream of literature with a new predictive model.

The development of predictive models can be utilized to improve system performance in service and healthcare applications. For example, Chen et al. (2024a) uses arrival information to improve routing decisions in service systems. Chen et al. (2024b) considers an online learning approach to determine the pricing and capacity sizing decisions of an G/G/1 queue. Helm et al. (2015) study how to conduct dynamic forecasting and construct algorithms for clinical decision support. Abouee-Mehrizi et al. (2022) uses a data-driven approach to manage the inventory of blood products in healthcare systems.

Our work considers incorporating the predictive models in a dynamic admission control problem in managing a queueing system. The study of the admission control problem has a long history; see Stidham (1985) and Stidham (2002) for overviews and references therein. Most of the admission control literature formulate the problem as a Markov decision process and studies a stationary system. Examples include Lewis et al. (1999), Ata and Shneorson (2006), Zayas-Cabán and Lewis (2020) and Audina and Ramanan (2011). Two closely related paper are Lewis et al. (2002) and Yoon and Lewis (2004). Both paper study the admission control problem with the time-varying arrival rate. The authors consider a system in which the arrival and service rates are bounded, periodic functions of time. They model the problem as a Markov decision process and show that under the infinite horizon discounted and average reward optimality criteria, for each fixed time, optimal admission control strategies are threshold-type policies, i.e. the optimal policy admits customers when the queue length is below a threshold. They propose a pointwise stationary approximation (PSA) to approximate the optimal policies, suggest a heuristic to improve the implementation of the PSA and verify its usefulness via a numerical study. Lewis et al. (2002) also provide examples to show that the threshold value may or may not decrease in the arrival rate. The monotonicity depends on the assumption on the arrival rate function. In our paper, we use the stochastic programming framework to study the problem, which allows us to incorporate a general information structure. Instead of characterizing the optimal policy using the queue length threshold, we propose an optimality condition which is characterized by the expected remaining congestion time. This new characterization allows us to derive the optimal policy for general forecast models.

Spencer et al. (2014), Xu and Chan (2016) and Ata and Peng (2020) are also closely related papers in the admission control literature. Spencer et al. (2014) study an admission problem for an overloaded $M/M/1$ queue under the assumption that the future information is available. The authors propose the no-job-left-behind policy, which effectively rejects those jobs with "excessive" delay (hence, left behind) by looking into the future. The authors show that the no-job-left-behind policy is asymptotically optimal in heavy traffic. Xu and Chan (2016) consider a similar model to manage the admission into an emergency department using the knowledge of future arrivals. They also enhance their policies using the thresholds on the queue length, which diverts arrivals if either the queue length is large or a high number of patients will arrive in future periods. The authors show that the proposed policy provides delay improvements over standard policies used in practice. Ata and Peng (2020) studies a call center with the callback option. This problem can be studied using the admission control problem. Similar to Spencer et al. (2014) and Xu and Chan (2016), Ata and Peng (2020) assumes that the future information of the arrival counts is available. However, Ata and Peng (2020) allows customers to choose to take or reject the offered callback option. Thus, the optimal lookahead policy takes the decisions of the customers into account. Our model differs from Spencer et al. (2014), Xu and Chan (2016) and Ata and Peng (2020) in several ways.

First, instead of assuming that the exact future arrival counts are known, our model allows a more general information structure. Second, the characterization of the optimal policy is different. Instead of restricted to the lookahead type policy, the optimality condition in this paper is characterized by the expected remaining congestion time of the system. The special case studies in Section 4.3 is closely related to these three papers. The optimal policy derived under the full information case concur the optimality of the lookahead policy found in these three papers. However, we are able to derive the optimal policy under the partial future information case. We show that the optimal policy is a lookahead type policy with threshold. However, this policy differs from the lookahead policy with queue threshold proposed in Xu and Chan (2016).

Our paper is also related to papers constructing lower bounds of a dynamic decision problem. Brown et al. (2010) and their follow-up works consider a general technique for lower bounds on minimal costs in stochastic dynamic programs. They take a dual approach and relax the nonanticipativity constraints; see Brown and Smith (2022) for a review of this approach. Our paper also uses a dual approach, but we relax the constraint characterizing the dynamics of the system instead of the nonanticipativity constraints.

3 Problem Formulation and the Information Structure

This section formulates the admission control problem as a multi-stage stochastic optimization problem. Section 3.1 describes the model setup, and Section 3.2 introduces a filtration associated with the model that describes how the system manager's available information evolves. With the defined filtration, Section 3.3 formulates the problem as a multi-stage stochastic optimization problem.

3.1 The Admission Control Problem

We consider a discrete-time model of a canonical single class queue. The system manager determines the admission control and routing policies over a finite horizon T. In each period, the system manager observes the current state of the system, the number of arrivals and the service capacity of the current period, and collects all available information to forecast the numbers of arrivals and service capacities in future periods. Using all available information collected, she then decides the number of incoming customers that are rejected by the system and admits the rest into the system. The number of customers that arrive to the system is denoted by $\lambda_t \geq 0$ in period t (for $t = 1, \ldots, T$). The system has a service capacity of $\bar{\mu}_t \geq 0$ in each period t. That is, if the system serves the customers at its full capacity, a number $\bar{\mu}_t$ of customers finish the service and leave the system in period t. We assume that the arrival process $\{\lambda_t, t = 1, \ldots, T\}$ and the service capacity $\{\bar{\mu}_t, t = 1, \ldots, T\}$ are two exogenous stochastic processes, i.e., the dynamics of these two primitive processes do not depend on the system manager's decision history. We use $d_t \in [0, \lambda_t]$ to denote the number of incoming customers that are rejected by the system in period t. If an incoming customer is rejected, he leaves the system immediately. If an incoming customer is admitted, he enters the system and stays in the system until he completes the service. The system manager also determines the service rate, denoted by $\mu_t \in [0, \bar{\mu}_t]$, in period t based on all available information[1]. Thus, the number of customers that leaves the system after service completion are μ_t in period t.

The system incurs a penalty of p for each rejected customer. The system incurs a holding cost of h per

[1]Although it is natural to assume that the system follows a non-idling service policy, making the service rate a decision variable ensures that the problem is formulated as a linear stochastic programming problem. In Section 4, we show that the optimal service policy follows a non-idling policy.

period for each customer in the system at the end of the period. The system manager seeks to find the optimal admission policy (that uses the available forecast information) that minimizes the expected total costs incurred over the finite horizon T. In what follows, we first formulate of the forecast model, and defines the set of the admissible policies given the forecast information structure. We then formulate the problem as a multi-stage stochastic programming problem.

3.2 The Information Structure and Admissible Policies

We use a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to characterize the uncertainty, where Ω is the sample space and $\mathcal{F}$ is a compete $\sigma-$algebra defined on the sample space Ω. Each event $\omega \in \Omega$ denotes a fixed sample path. Thus, the arrival process $\{\lambda_t(\omega), t = 1, \ldots, T\}$ and the service capacity process $\{\bar{\mu}_t(\omega), t = 1, \ldots, T\}$ are two stochastic processes defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The system manager's admission and service decisions are affected by the information that she collects in each period. To characterize the evolution of information, we define a filtration $\mathbb{F} = \{\mathcal{F}_t, t = 1, \ldots, T\}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ are sub-$\sigma-$algebra of $\mathcal{F}$ (for $1 \leq s \leq t \leq T$). Each sub-σ-algebra $\mathcal{F}_t$ represents all information available to the system manager up to period t. More discussions about the filtration can be found in Shapiro et al. (2014) and Carpentier et al. (2015). Later in this section, we will provide a few examples to illustrate how various σ-algebra represents different information scenarios and how it affects the set of admissible policies. By using the filtration to characterize the information, we make an implicit assumption that the system manager never forgets the information collected so far.

Since the manager observes and remembers all past arrivals $\lambda_1(\omega), \ldots, \lambda_t(\omega)$ and service capacities $\bar{\mu}_1(\omega), \ldots, \bar{\mu}_t(\omega)$ in period t (for $t = 1, \ldots, T$), the history of past arrivals and service capacities is available information in period t. That is, the history $(\lambda_1(\omega), \cdots, \lambda_t(\omega), \bar{\mu}_1(\omega), \cdots, \bar{\mu}_t(\omega))$ is $\mathcal{F}_t$-measurable, i.e., the arrival and service capacity processes are adapted to the filtration $\mathbb{F}$. In addition, we make a technical assumption that the expected mean of the numbers of arrivals and service capacities in future periods are finite given all the available information so far. These two assumptions are stated formally in Assumption 1.

Assumption 1. *The following hold:*

(i) The arrival process $\{\lambda_t(\omega), t = 1, \ldots, T\}$ and the service capacity process $\{\bar{\mu}_t(\omega), t = 1, \ldots, T\}$ are adapted to the filtration $\mathbb{F}$;

(ii) Moreover, we have that $\mathbb{E}[\lambda_s(\omega)|\mathcal{F}_t] < \infty$ and $\mathbb{E}[\bar{\mu}_s(\omega)|\mathcal{F}_t] < \infty$ for $t = 1, \ldots, T$ and $s = t+1, \ldots, T$.

When the system manager makes the decisions of the number $d_t(\omega)$ of customers and the actual service rate $\mu_t(\omega)$ in period t (for $t = 1, \ldots, T$), she uses all available information up to period t. Thus, we restrict our attentions to nonanticipative policies, i.e., the admission and service decisions that are adapted to the filtration $\mathbb{F}$. To be specific, we denote an admissible policy $\pi = \{(d_t(\omega), \mu_t(\omega) : t = 1, \ldots, T\}$ as a decision rule that maps the current system state and information to the rejection and service decisions. The set of all admissible policies is denoted by Π, which consists of all policies $\pi = \{(d_t(\omega), \mu_t(\omega) : t = 1, \ldots, T\}$ such that $(d_t(\omega), \mu_t(\omega)) \in \mathcal{F}_t$, for $t = 1, \ldots, T$. It is worth mentioning that different information structure, which is characterized by the information set $\mathcal{F}_t$, yields different sets of admissible policies.

One way to interpret the filtration $\mathbb{F}$ is assuming that it is the natural filtration generated by a multi-dimensional stochastic process $\{X_t(\omega) : t = 1, \ldots, T\}$. The choice of the stochastic process $\{X_t(\omega) : t = 1, \ldots, T\}$ affects the information availability of the system manager. For instance, if the system manager only observes the history of arrivals and service capacities in the past, the stochastic process $X_t(\omega)$ that generates the information sets is $(\lambda_t(\omega), \bar{\mu}_t(\omega))$ (for $t = 1, \ldots, T$). In particular, an $M/M/1$ queue with arrival rate λ and service rate μ, can be converted into a discrete time Markov chain via uniformization. The discrete time analogue of the $M/M/1$ queue is characterized by independently and identically distributed $(\lambda_t(\omega), \bar{\mu}_t(\omega))$ with the following distribution: For $t = 1, \ldots, T$,

$$(\lambda_t(\omega), \bar{\mu}_t(\omega)) = \begin{cases} (1, 0), & \text{with probability } \frac{\lambda}{\lambda+\mu}, \\ (0, 1), & \text{with probability } \frac{\mu}{\lambda+\mu}. \end{cases} \tag{1}$$

The nonanticipativity constraint requires that the admission and service decisions only depends on observations of the past arrival counts and service capacities.

Another special information structure is that the system manager has all future information about the arrivals and the service capacities; see Ata and Peng (2020) and Xu and Chan (2016) for example. To be specific, the system manager foresees the exact numbers of arrivals and the service capacities in all future periods. In this case, the information sets $\mathcal{F}_t = \mathcal{F}_T$ (for $t = 1, \ldots, T$), where $\mathcal{F}_T$ is the σ-algebra generated by the entire sample path $(\lambda_1(\omega), \bar{\mu}_1(\omega), \cdots, \lambda_T(\omega), \bar{\mu}_T(\omega))$. The nonanticipativity constraint simply implies that the admission and service decisions depend on the entire sample path. These two examples of information structure are the cases when the system manager has the minimal and maximal available information. We call them the no future information and full future information scenarios.

In practice, the information available for the system manager lies in between these two extreme cases. The system manager may use past data to learn the arrival and service rate patterns. She may also use auxiliary information, such as companies' marketing campaign plans and staffing schedules, to facilitate the forecast of future arrival volume and service capacities. Ibrahim et al. (2016a) and Ibrahim et al. (2016b) have extensive discussions of the forecast models of the arrival process and service times in call center settings. For example, a system manager may use the framework proposed in Ibrahim et al. (2016a) and use a regression model to forecast the arrival volume. To be specific, assume that the system manager uses the following regression model to forecast the numbers of arrivals and the service capacities in future periods: For $t = 1, \ldots, T$,

$$\lambda_t(\omega) = f_t(X_1(\omega), \cdots, X_{t-1}(\omega)) + \epsilon_t(\omega) \quad \text{and} \quad \bar{\mu}_t(\omega) = g_t(X_1(\omega), \cdots, X_{t-1}(\omega)) + \nu_t(\omega). \tag{2}$$

where $X_t(\omega)$ is a multi-dimensional vector taking values in the state space $\mathcal{X}$, $f_t(\cdot)$ and $g_t(\cdot)$ are two functions that characterize the forecast models of the arrival and service processes, and $\epsilon_t(\omega)$, $\nu_t(\omega)$ are i.i.d random noises. In this case, the filtration $\mathbb{F}$ is the natural filtration generated by the variate process X_t. The nonanticipativity constraint requires that the admission and service decisions are functions of the history of the random variable X_t.

3.3 The Multi-stage Stochastic Programming Problem

This section formulates the problem as a multi-stage stochastic programming problem. One natural way to solve the admission control problem with a general forecast model is to formulate it as a dynamic programming problem by including all the variate of the forecast model into the state. However, we may suffer from the curse of dimensionality because the variate of the forecast model can be high-dimensional. Using the filtration to model the information structure helps us separate the states whose evolution is given endogenously from the information states that evolve independent of the admission and service decisions. Since the evolution of the number of arrivals and the service capacity are exogenous, the system state of this admission control problem is merely the number of customers in the system. The dependence of the decisions on the information structure is modeled using the nonanticipativity constraint.

To state the problem, we first fix an admissible policy $\pi = \{(d_t^\pi(\omega), \mu_t^\pi(\omega)), t = 1, \ldots, T\} \in \Pi$. The state of the system at the end of period t is the number of customers in the system, denoted by $q_t^\pi(\omega)$. Thus, the dynamics of the state $q_t^\pi(\omega)$ is characterized by the following equation: For $t = 1, \ldots, T$,

$$q_t^\pi(\omega) = q_{t-1}^\pi(\omega) + \lambda_t(\omega) - d_t^\pi(\omega) - \mu_t^\pi(\omega), \tag{3}$$

where $q_0^\pi(\omega) = q_0$ is the initial number of customers in the system. Since the number of customers in the system never goes below zero, we have to impose the following state constraint: For $t = 1, \ldots, T$,

$$q_t(\omega) \geq 0 \ a.s. \tag{4}$$

The number of the rejected incoming customers cannot exceed the number of total incoming customers in each period. Moreover, the actual service rate cannot exceed the service capacity in each period. Thus, we also include the following control constraints: For $t = 1, \ldots, T$,

$$0 \leq d_t(\omega) \leq \lambda_t(\omega) \ \text{ and } \ 0 \leq \mu_t(\omega) \leq \bar{\mu}_t(\omega) \ a.s. \tag{5}$$

Thus, the evolution of $q_t^\pi(\omega)$ is fully characterized by the arrival and service capacity processes and the control policy π. Moreover, the state process $\{q_t^\pi(\omega), t = 1, \ldots, T\}$ is adaptive to the filtration $\mathbb{F}$.

The system manager's objective is to minimized the expected total cost, which is the sum of the holding cost and the rejection penalty, over the finite horizon T. To account for the ending effect of the problem, we assume that the system has no new arrivals and that the service capacity is a constant $\bar{\mu}_T$ after period T. Under these assumptions, if there are $q_T(\omega)$ customers at the end of period T, the total holding cost incurred is $hq_T^2(\omega)/2\bar{\mu}_T$ after period T [2]. For a fixed policy $\pi \in \Pi$, we denote $C^\pi(q_0)$ as the total expected cost over the finite horizon T if the system starts with q_0 customers in the system. The expected cost $C^\pi(q_0)$ is given as follows: For $q_0 \geq 0$,

$$C^\pi(q_0) = \mathbb{E}\left[\sum_{t=1}^{T}(pd_t^\pi(\omega) + hq_t^\pi(\omega)) + \frac{h\left(q_T^\pi(\omega)\right)^2}{2\bar{\mu}_T}\right].$$

The system manager seeks to minimize the expected total cost that incurs over T periods. To be specific, the system manger solves the following problem:

$$C^*(q_0) = \inf_{\pi \in \Pi} \ C^\pi(q_0) \ \text{ s.t. } \ (3)\text{-}(5). \tag{P}$$

We denote the admission control problem that the system manager solves as problem (P) (primal problem). We let $C^*(q_0)$ denote the optimal objective value of (P) if the initial number of customers in the system is q_0.

[2] We ignore the integrality effect in computing the holding cost after period T. The exact holding cost after time T is $hl(q_T(\omega))/\bar{\mu}_T$, where $l(q) = (\lfloor q \rfloor + 1)q - \frac{\lfloor q \rfloor(\lfloor q \rfloor + 1)}{2}$. We approximate this cost with the quadratic function to simplify the algebra.

4 Characterization of the Optimal Admission Policy

This section characterizes the optimal admission policies that solves problem (P), and provides properties of the optimal policy. We first shows that the optimal policy is a threshold policy. We then construct a dual problem to characterize the optimal policy using the dual process. For the brevity of the notations, we omit the dependence of the decisions and states on the sample path ω for the rest of paper.

4.1 The Threshold Policy

This section shows that the optimal admission policy follows a threshold-type structure. To facilitate the analysis to go, we first define a threshold policy.

Definition 1. *A policy $\pi = \{(d_t^\pi, \mu_t^\pi) : t = 1, \ldots, T\} \in \Pi$ is a threshold policy if there exists a non-negative stochastic process $\{n_t^\pi : t = 1, \ldots, T\}$ adaptive to $\mathbb{F}$ such that the following holds: For $t = 1, \ldots, T$,*

$$d_t^\pi = \begin{cases} \lambda_t, & \text{if } \; n_t^\pi < q_{t-1}^\pi - \bar{\mu}_t, \\ q_{t-1} + \lambda_t - \bar{\mu}_t - n_t^\pi, & \text{if } \; q_{t-1}^\pi - \bar{\mu}_t < n_t^\pi \le q_{t-1}^\pi + \lambda_t - \bar{\mu}_t, \\ 0, & \text{if } \; q_{t-1}^\pi + \lambda_t - \bar{\mu}_t \le n_t^\pi, \end{cases} \tag{6}$$

$$\mu_t^\pi = \begin{cases} \bar{\mu}_t, & \text{if } \; 0 \le q_{t-1}^\pi + \lambda_t - \bar{\mu}_t, \\ q_{t-1}^\pi + \lambda_t, & \text{if } \; q_{t-1}^\pi + \lambda_t - \bar{\mu}_t < 0, \end{cases} \tag{7}$$

where q_t^π follows equation (3).

Definition 1 states that a threshold policy π is characterized by a stochastic process $\{n_t^\pi : t = 1, \ldots, T\}$, in which the value of the threshold $n_t^\pi \in \mathcal{F}_t$ is a random variable depending on the available information and the realization of the realization of the sample path ω. Under the threshold policy, equation (7) states that the system follows a non-idling service policy, i.e., the system only foregoes service capacity if the system is empty (with all incoming customers admitted in that period). If the system is not empty, equation (6) states that the number q_t^π of customers in the system chases the threshold n_t^π by adusting the number of incoming customers admitted to the system. To be specific, if the number of customers in the system is already large enough, all incoming customers are rejected; see case 1 of equation (6). In this case, the number of customer in the system still exceeds the threshold n_t^π after rejecting all incoming customers and serving the customers at its full capacity. In case 2, the number of customers in the system at the end of the period matches the threshold n_t^π. To achieve this, the system serves customers at its full capacity and admits the exact number of incoming of customers to match the threshold. In case 3, the system admits all incoming customers. The number of customers in the system at the end of the period is still less than the threshold n_t^π after admitting all incoming customers. The following theorem states that there exists a threshold policy that is optimal.

Theorem 1. *Problem (P) has an optimal policy π^* characterized by the threshold process $\{n_t^*, t = 1, \ldots, T\}$.*

Although Theorem 1 shows that the optimal admission policy has a threshold-type structure, i.e., the system manager starts to reject incoming customers if the number of customers in the system exceeds given thresholds. Since the threshold depends on the forecast of the future arrivals and service capacities, the characterization of the thresholds remain challenging. In what follows, we provide an optimality condition of the optimal policy, and then use the condition to characterize the optimal thresholds.

4.2 The Optimality Condition

In this section, we derive an optimality condition of the optimal policy by constructing a dual problem of the original admission control problem (P). By showing strong duality and constructing the complementary slackness conditions of the primal and dual problems, we provide the characterization of the optimality condition of the optimal admission policy. In the full information scenario, the optimality condition reduces into the optimality condition derived using the Pontryagin's maximum principle for the (deterministic) optimal control problem. Thus, this optimality condition can be viewed as a generalization for the Pontryagin's maximum principle that incorporates the uncertainty.

In what follows, we first define a dual problem (D) of admission control problem (P). To this end, let $\phi = \{y_t : t = 1, \ldots, T\}$ be a stochastic process adapted to the filtration $\mathbb{F}$. For all $q_0 \geq 0$ and a given ϕ, we define a function $D^\phi(q_0)$ as follows:

$$D^\phi(q_0) = y_1 q_0 + \mathbb{E}\left[\sum_{t=1}^{T}((\lambda_t - \bar{\mu}_t)y_t - (y_t - p)^+\lambda_t - y_t^-\bar{\mu}_t) - \frac{y_{T+1}^2}{2h}\right]. \tag{8}$$

In addition, we impose the following constraint that the process ϕ needs to satisfy: For $t = 0, \ldots, T$,

$$y_t \leq h + \mathbb{E}[y_{t+1}|\mathcal{F}_t] \quad \text{a.s.} \tag{9}$$

Let Φ be the set of all feasible solutions of the dual problem (D), i.e., the set of all stochastic process ϕ adapted to $\mathbb{F}$ that satisfies equation (9). Thus, the dual problem (D) is defined as follows: For $q_0 \geq 0$,

$$D^*(q_0) = \sup_{\phi \in \Phi} D^\phi(q_0). \tag{D}$$

The following theorem states that there is no duality gap between the primal problem (P) and the dual problem (D). It also provides the complementary slackness conditions that the optimal solutions of the primal and dual problems need to satisfy.

Theorem 2. *The following hold: For $t = 1, \ldots, T$,*

(i) There is no duality gap between problems (P) and (D), i.e., $C^*(q_0) = D^*(q_0)$ for $q_0 \geq 0$;

(ii) An admissible policy $\pi^* = \{(d_t^*, \mu_t^*) : t = 1, \ldots, T\} \in \Pi$ is optimal if there exists an admissible solution $\phi^* = \{y_t^* : t = 1, \ldots, T\} \in \Phi$ that satisfies constraint (9) and the following conditions almost surely: For $t = 1, \ldots, T$,

$$\text{if } \ q_t^* > 0, \qquad \text{then } y_t^* = h + \mathbb{E}[y_{t+1}^* | \mathcal{F}_t]; \tag{10}$$

$$\text{if } \ y_t^* > p, \qquad \text{then } \ d_t^* = \lambda_t, \ \mu_t^* = \bar{\mu}_t; \tag{11}$$

$$\text{if } \ y_t^* = p, \qquad \text{then } \ 0 \leq d_t^* \leq \lambda_t, \ \mu_t^* = \bar{\mu}_t; \tag{12}$$

$$\text{if } \ y_t^* \in (0, p), \quad \text{then } \ d_t^* = 0, \ \mu_t^{\pi^*} = \bar{\mu}_t; \tag{13}$$

$$\text{if } \ y_t^* = 0, \qquad \text{then } \ d_t^* = 0, \ 0 \leq \mu_t^* \leq \bar{\mu}_t; \tag{14}$$

$$\text{if } \ y_t^* < 0, \qquad \text{then } \ d_t^* = 0, \ \mu_t^* = 0, \tag{15}$$

where q_t^* is the resulting number of customers in the system in period t under policy π^*.

Theorem 2 provides a condition to verify the optimlaity of a given admission policy of problem (P). Part (i) of Theorem 2 states that to find the optimal policy, it is equivalent to find the optimal stochastic process $\phi \in \Phi$ that maximize the dual problem (D). The stochastic process $\{y_t : t = 1, \ldots, T\}$ is an analogue of the adjoint (or costate) function of the deterministic optimal control problem. We call the process $\{y_t : t = 1, \ldots, T\}$ as the adjoint process of the primal problem (P). Part (ii) of Theorem 2 states the complementary slackness conditions that the optimal solutions of the primal problem (P) and the dual problem (D) jointly satisfy. It characterizes the dynamics of the optimal adjoint process ϕ^* under the optimal policy. Condition (10) states that the inequality constraint (9) in the dual problem becomes an equality when the system is not empty. Conditions (11)-(14) provides the characterization of the optimal policy π^* given the optimal adjoint process ϕ^*. Fixed the time t and the sample path ω, if the value of y_t^* exceeds the one-time rejection penalty p, the system rejects all incoming customers. If y_t^* equals p, then the system rejects some incoming customers and admit the rest. If the value of y_t^* is less then the penalty p but strictly positive, we admit all incoming customers. In all of these three cases, the system use its full service capacity in this period. When the value of y_t^* is less than or equal to zero, the system admits all incoming customer, completes service for all customers in the system, and forgoes the extra service capacity.

Theorem 2 states that the optimal admission policy can be determined easily if the optimal adjoint process is known. However, solving for the optimal adjoint process remains challenging. In what follows, we provide some useful properties of the optimal adjoint process, and use them to solve the optimal admission policies for special cases.

4.3 Characterization of the Optimal Adjoint Process

This subsection characterizes some useful properties of the optimal adjoint process. These properties are then utilized to derive closed-form characterizations of the optimal admission policy for two specific information scenarios: the no future information case and the full future information case.

The following lemma states that there exists an optimal adjoint process that is non-negative.

Lemma 1. *There exists a non-negative solution $\phi^* = \{y_t^* : t = 1, \ldots, T+1\} \in \Phi$ that is optimal to (D), i.e, $y_t^* \geq 0$ almost surely for $t = 1, \ldots, T+1$.*

Thus, we restrict our attention to non-negative adjoint processes by Lemma 1 for the rest of this paper. Next, we fix an optimal admission policy π^* and a nonnegative optimal adjoint process ϕ^* that jointly satisfies equations (10)-(15). Lemma 2 follows immediately from Theorem 2 and Lemma 1. Lemma 2 shows that the system has slackness in the service capacity only when the optimal adjoint process is zero. In this case, the system must be empty. In other words, the optimal routing policy follows a non-idling policy.

Lemma 2. *If $\mu_t^* < \bar{\mu}_t$, then $y_t^* = 0$ and $q_t^* = 0$ for $t = 1, \ldots, T$.*

Next we show the connection between the optimal adjoint process ϕ^* and the expected length of the busy periods under the optimal policy π^*. Note that equation (10) provides a backward characterization the optimal adjoint process when the system is not empty. To be specific, the value of the optimal adjoint process can be decomposed into two parts: The expected number of periods until the system is empty and the expected value of the optimal adjoint process when the system is empty. This key observation provides bounds of the value of the optimal adjoint process.

To facilitate the analysis to follow, we define two stopping times associated with the dynamics of the system under a given admission policy. First, we denote the stopping time $R_t^\pi(q_{t-1})$ as the time until the system is empty after time t given that there are q_t number of customers in the system at the end of period t under the optimal policy for a given policy $\pi \in \Pi$. That is,

$$R_t^{\pi}(q_t) = \inf\Big\{0 \le t' \le T - t : q_{t+t'}^{\pi} = 0\Big\}.$$

Second, we denote $S_t^{\pi}(q_{t-1})$ as the time until the system has slackness in its service capacity under a given policy $\pi \in \Pi$, i.e.,

$$S_t^{\pi}(q_t) = \inf\Big\{0 \le t' \le T - t : \mu_{t'}^{\pi} < \bar{\mu}_{t'}\Big\}.$$

The following proposition provides upper and lower bounds of the optimal adjoint process ϕ^* using the two stopping times defined under the optimal policy π^*.

Proposition 1. *For $t = 1, \ldots, T$ and $\omega \in \Omega$, the following holds:*

$$h\mathbb{E}[R_t^{\pi^*}(q_t^*)|\mathcal{F}_t] \le y_t^* \le h\mathbb{E}[S_t^{\pi^*}(q_t^*)|\mathcal{F}_t].$$

Moreover, if $q_t^* = n_t^*$, then the following holds:

$$\mathbb{E}[R_t^{\pi^*}(n_t^*)|\mathcal{F}_t] \le p/h \le \mathbb{E}[S_t^{\pi^*}(n_t^*)|\mathcal{F}_t].$$

Proposition 1 provides a necessary condition for the optimal policy to satisfy, it can fully characterize the optimal admission threshold using condition (1) for special cases. In particular, we show that condition (1) provides the closed-form characterization of the optimal admission thresholds for the two special information structure: full and no future information scenarios.

Full information. In this case, $\mathcal{F}_t = \mathcal{F}_T$ for all $t = 1, \ldots, T$. If we fix the initial number q_0 of customers in the system, both the dynamic of the system and the optimal adjoint process are deterministic. In what follows, we are interested in finding the threshold n_0 such that the system starts to reject incoming customers in the first period when $q_0 = n_0$.

The constraint (9) and optimality condition (10) characterize the dynamics of the optimal adjoint process. It follows from (10) that the optimal adjoint process $\{y_t^* : t = 0, \ldots, T+1\}$ decreases at a rate of h if the system is not empty. When the system is empty, it follows from the condition (9) that the y_t^* may jump down. Thus, it follows from (11)-(13) that the optimal admission policy that in each busy period, once the system stops rejecting incoming customers, it admits all incoming customers until the end of the current busy period. If the initial number of customers equal to the threshold, i.e, $q_0 = n_0$, then $y_0 = p/h$ and the system does not reject any incoming customers afterwards. Therefore, it follows from Proposition 1 that the optimal threshold n_0 satisfies following condition:

$$\inf\left\{t \geq 0 : n_0 + \sum_{s=1}^{t}(\lambda_s - \bar{\mu}_s) \ \leq 0\right\} \leq p/h < \inf\left\{t \geq 0 : n_0 + \sum_{s=1}^{t}(\lambda_s - \bar{\mu}_s) \ < 0\right\}.$$

The left-hand side of the inequality is $R_1(n_0)$, i.e., the first time when the system is empty, while the right-hand side is $S_1(n_0)$, the first time when the system has excess service capacity. Equivalently, the following inequality holds:

$$\sup\left\{q \geq 0 : q + \sum_{s=1}^{t}(\lambda_s - \bar{\mu}_s) < 0, \exists\ t \in T_{p/h}\right\} \leq n_0 \leq \inf\left\{q \geq 0 : q + \sum_{s=1}^{t}(\lambda_s - \bar{\mu}_s) > 0, \forall\ t \in T_{p/h}\right\},$$

where $T_{p/h} = \{0, 1, \ldots, \lfloor p/h \rfloor\}$. The left-hand side of the inequality is the largest initial number of customers in the system so that the system has excess service (without declining any incoming customers) within next $\lfloor p/h \rfloor$ periods. The right-side denotes the smallest initial number of customers such that the system never becomes empty (without declining any incoming customers) within next $\lfloor p/h \rfloor$ periods. Note that the left-hand and right-hand sides of the inequalities have the same values. Therefore, the threshold value n_0 is uniquely determined by either the left-hand side or the right-hand side of the inequalities.

No Future Information. We consider a stationary system, i.e., the number of arrivals and the service capacity in each period are both i.i.d. random variables. for $t = 1, \ldots, T$, we assume that $\lambda_t \in \mathbb{N}_+$ and $\bar{\mu}_t \in \mathbb{N}_+$ are non-negative discrete random variables with probability density functions $g_\lambda(\cdot)$ and $g_\mu(\cdot)$, respectively. In addition, we use $g_d(\cdot) \in \mathbb{N}$ to denote the probability density function of the net number of incoming customers, i.e., $\lambda_t - \bar{\mu}_t$, in each period.

We are interested in finding a stationary optimal policy when the total length T of the finite horizon goes to infinity. Since the system manager cannot infer future arrival volume and service capacity from past observation, the admission policy depends only on the current number of customers in the system. Thus, the optimal threshold policy is characterized by a constant $n^* \in \mathbb{N}_+$. We use π_n denote an admission policy that admits incoming customers until the number of customers in the system at the end of the period is n. Under a policy π_n, the number of customers in the system only goes down if it exceeds the threshold n. Once the number of customers in the system is below the threshold, it never exceeds n again. In a stationary system, we can ignore the case when the number of customers in the system exceeds the threshold n. To apply Proposition 1, we are interested in computing the values of $\mathbb{E}[R^{\pi_n}(q)]$ and $\mathbb{E}[S^{\pi_n}(q)]$ for $q = 0, 1, \ldots, n$. Under a threshold policy π_n, the expected length of the busy period starting with a number q of customers in the system is given as follows: For $q = 1, \dots, n$,

$$\mathbb{E}[R^{\pi_n}(q)] = 1 + \sum_{i=-q+1}^{n-q-1} \mathbb{E}[R^{\pi_n}(q+i)] g_d(i) + \mathbb{E}[R^{\pi_n}(n)] \sum_{i=n-q}^{\infty} g_d(i),$$

with $\mathbb{E}[R^{\pi_n}(0)] = 0$. Moreover, we observe that $\mathbb{E}[S^{\pi_n}(q)] = \mathbb{E}[R^{\pi_{n+1}}(q+1)]$ for $q \geq 0$ and $\mathbb{E}[S^{\pi_n}(-1)] = 0$. Thus, Proposition 1 implies that with the optimal threshold $n^* \in \mathbb{N}_+$, the following holds:

$$\mathbb{E}[R^{\pi_{n^*}}(n^*)] \leq \frac{p}{h} < \mathbb{E}[R^{\pi_{n^*+1}}(n^*+1)]. \tag{16}$$

Note that there is only one unique integer n^* that satisfies equation (16). Thus, it gives the optimality condition of the optimal threshold n^*. A special case is a discrete time $M/M/1$ queue, in which the number of arrivals and the service capacity follow equation (1). In this case, we have that for $n \geq 0$,

$$\mathbb{E}[R^{\pi_n}(n)] = \frac{1}{\mu(1-\rho)} \left[n - \rho \frac{1-\rho^n}{1-\rho} \right].$$

Thus, equation (16) implies that the optimal threshold $n^* \in \mathbb{N}_+$ satisfies the following:

$$\frac{1}{\mu(1-\rho)} \left[n^* - \rho \frac{1-\rho^{n^*}}{1-\rho} \right] \leq \frac{p}{h} < \frac{1}{\mu(1-\rho)} \left[(n^*+1) - \rho \frac{1-\rho^{n^*+1}}{1-\rho} \right].$$

The two special information scenarios show that the structure of the optimal policy depends on the information structure. The optimal policy derived in the full future information scenario is consistent with the optimal lookahead policy in Ata and Peng (2020), which only proves the optimality of the policy in the under-loaded regime via a pathwise approach. The optimal policy in the no future information scenario is consistent with the optimal threshold provided in Naor (1969) and Wang (2016), which provide the same optimal threshold by analyzing a Markov decision process or analyzing the queueing system via a carefully designed service discipline. Theorem 2 not only provides a unified approach of the optimal policy in both information scenarios, but also characterizes the optimal policy for more general information structure.

Lastly, we present a corollary that provides a lower bound of the optimal objective value. For any given feasible adjoint process $\phi \in \Phi$, the weak duality result provides a lower bound of the optimal objective value of the primal admission control problem (P).

Corollary 1. *For any feasible adjoint process $\phi = \{y_t : t = 1, \dots, T\} \in \Phi$, the following holds: For $q_0 \geq 0$,*

$$y_1 q_0 + \mathbb{E}\left[\sum_{t=1}^{T-1} ((\lambda_t - \bar{\mu}_t) y_t - (y_t - p)^+ \lambda_t - \frac{(y_T)^2}{2h} \right] \leq C^*(q_0).$$

Although Corollary 1 provides a lower bound of the optimal value of the primal problem (P), it does not ensure that this lower bound is tight for a given adjoint process. The tightness of the bound depends on the choice of the adjoint process ϕ.

4.4 A Simulation-based Heuristic for General Forecast Models

For a general information structure, solving for the optimal policy is challenging. Even if the optimal policy is solved, its potentially complicated structure, dependent on forecast information, can make it difficult to implement in practice. Therefore, this section proposes a simulation-based heuristic that is easy to implement and can flexibly incorporate general forecast models in practice.

Proposition 1 provides an upper bound of the expected remaining congestion time $\mathbb{E}[R_t^{\pi^*}(n_t^*)|\mathcal{F}_t]$ under the optimal admission policy. It suggests that the optimal admission control policy aims to manage the expected remaining congestion time within a reasonable range. However, this expectation is computed under the optimal admission policy, which is typically unknown. To address this, we use a pre-specified policy π and compute the expected remaining congestion time under the given policy π as a proxy for the expected remaining congestion time under the optimal policy. The following section proposes a heuristic that bounds the expected remaining congestion time under the specified policy π.

Definition 2. *A bounded congestion time policy under policy π admits an incoming customers in period t if $E[R_t^{\pi}(q_t)|\mathcal{F}_t] \leq \bar{t}$, where q_t is the number of customers in the system at the beginning of period t.*

The bounded congestion time policy is a simulation-based heuristic because the expectation can be computed easily via simulations. In particular, one can easily simulate expected remaining congestion time based on the information available for a specified policy π using discrete event simulations. If the expected remaining congestion time is smaller than a given threshold $\bar{t}$, then the incoming customer is admitted. Otherwise, the incoming customer is rejected. Since the pre-specified policy may not be optimal, the expectation $E[R_t^{\pi}|\mathcal{F}_t]$ is used as a proxy as how long the system will remain congested under the available information. Thus, the threshold $\bar{t}$ serves as a tuning parameter in the heuristic, controling how congested the system is depending on the values of the holding cost h and the penalty p.

Although the choice of the pre-specified policy π affects the performance of the proposed heuristic, we show that a naive policy may generate a policy that is near optimal in the numerical analysis. Specifically, we consider a no-rejection policy, i.e., the system manager admits all incoming customers. In case of an overloaded system, the number of customers in the system accumulates under the no-rejection policy. To stablize the system, she rejects all incoming customers if the congestion time exceeds a large value.

5 Numerical Analysis

This section examines the performance of the proposed heuristic under various arrival and forecast models. The first study considers the scenario where forecasts of future arrivals and service completions are accurate in both stationary and non-stationary systems. The second study considers the scenario where forecasts of future arrivals and service completions are noisy.

5.1 Accurate Forecast Models

This section considers the case when the system manager has accurate future but limited information about the arrival and service completion counts. This scenario was also studied in Xu and Chan (2016). Specifically, the system manager collects accurate future arrival and service completion counts for a finite number of periods to make admission decisions.

We first study the performance of the simulation-based heuristic, i.e, the bounded congestion time policy, for a stationary system. In each period, the joint distribution of the arrival and service completion counts follows equation (1). At time t, the system manager knows the exact realization of $(\lambda_t, \mu_t, \cdots, \lambda_{t+w}, \mu_{t+w})$ for next w periods, and she assumes that the number of arrivals and service completion in each period follows equation (1) afterwards. Recall from the analysis in Section 4.3 that if the system manager has the future information for next p/h periods, she is able to construct an optimal policy that achieves the same performance as if she knows the information of all future periods. Thus, it suffices to consider the lookahead period w that is less than p/h.

As mentioned in Section 4.4, we pick a simple policy π that admits all incoming customers to estimate the expected remaining congestion time in the simulation-based heuristic. Note that under this no-rejection policy π, the simulated-based heuristic reduces to a simple threshold policy. If the system becomes empty within next w period, then the expected remaining congestion time is known and deterministic. Since the joint distribution of (λ, μ) follows an i.i.d. distribution, the expected remaining congestion time only depends on the estimated number of customers in the system in period $t+w$, if the system is not empty under the no-rejection π. Thus, the bounded congestion time policy is equivalent a simple threshold policy. An incoming customers is rejected if both conditions hold:

1. The number of customers in the system never becomes zero in next w periods;

2. The number of customers in the system in period $t+w$ under the no-rejection policy π is greater than

or equal to the threshold $\bar{q}$, i.e.,

$$q_t + \sum_{s=t}^{t+w} (\lambda_s - \mu_s) \geq \bar{q}.$$

We compare the average cost per period under the bounded congestion time policy with that under the optimal policy and under the lookahead policy $PA_w(0, l)$ proposed in Xu and Chan (2016). The optimal policy is computed by solving a Markov decision process (MDP), where the state of the system includes the w dimensional vector representing the realization of (λ, μ) for next w periods. As the state grows exponentially as the number w of lookahead periods grows, the time to solve the associated MDP also grows exponentially in w due to curse of dimensionality. Thus, we only compute the optimal policy for $w \leq 18$. The lookahead policy $PA_w(0, l)$ declines an incoming customer in period i if both of the following conditions are satisfied:

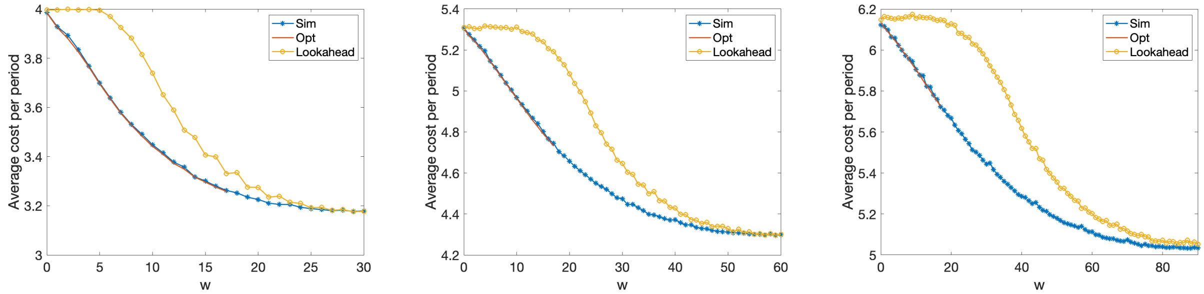
1. The number of customers in the system never hits zero in next w periods;

2. The number of customers in the system in period t is greater than or equal to the threshold l.

The threshold l is a tuning parameter in the lookahead policy $PA_w(0, l)$. Therefore, we compute the average cost for various thresholds l and choose the optimal one. The average costs under the bounded congestion time policy and the w-lookahead policy are estimated via simulations. The average cost under the optimal cost is obtained via solving the associated MDP.

Figure 1 considers a under-loaded system with the traffic intensity $\rho = \lambda/\mu = 0.9 < 1$ under three different p/h values. In an under-loaded system, only a small fraction of incoming customers are declined when the system manager anticipates the system will be congested. Figure 1 shows that the bounded congestion time policy is near optimal. The maximum cost difference between the bounded congestion time policy and the optimal policy is less 0.4% in all cases. Although we are unable to compare the cost difference for large lookeahead periods, the full information analysis in Section 4.3 implies that the bounded congestion time policy is optimal when the lookahead period w equals the value of p/h. Since the cost only reduces marginally when w is large as shown in Figure 1, the bounded congestion time policy is also near-optimal when w is large. One advantage of the bounded congestion time policy is that its computation time grows linearly in the lookeahd period w, making it efficient to construct a new-optimal policy for large lookahead period.

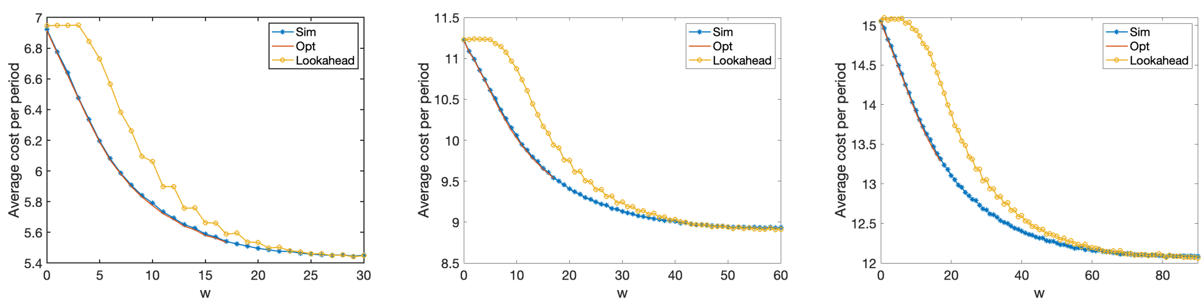
Figure 1 also shows that future information effectively reduces the operational cost . The average cost per period reduces by 20.2%, 19.0% and 17.7% for $p/h = 30, 60, 90$, repsectively. However, the marginal value of future information decreases as the lookahead window expands. One managerial implication from this observation is that even information about only the near future can help reduce the operational cost significantly. For example, the average cost reduces by 7% with a lookahead period $w = 6$ for the case of $p/h = 30$.



(a) $p/h = 30$ (b) $p/h = 60$ (c) $p/h = 90$

Figure 1: Average cost per period under the simulation-based heuristic (the bounded congestion time policy), the optimal policy and the lookahead policy for $\rho = 0.9$

Additionally, Figure 1 shows that the lookahead policy performs well only when the lookahead period is very large. When the lookahead period w is small, the lookahead policy perform similarly to a fixed threshold policy when there is no future information available. The main difference between the proposed heurstic in this paper and the lookahead is that the queue length threshold is imposed on the current queue length in the lookahead policy. Specifically, if the system does not empty the queue within next w periods, the incoming customers are rejected only when the current queue length is high. However, the bounded congestion time policy utilizes the information in next w periods to estimate the queue length at the end of the lookahead window and determines whether or not to admit the incoming customers. When the lookahead period w is large, the admission/rejection decisions is mainly driven by whether the system is emptied within next w periods or not. In this case, the two policies result in similar decisions and performance.



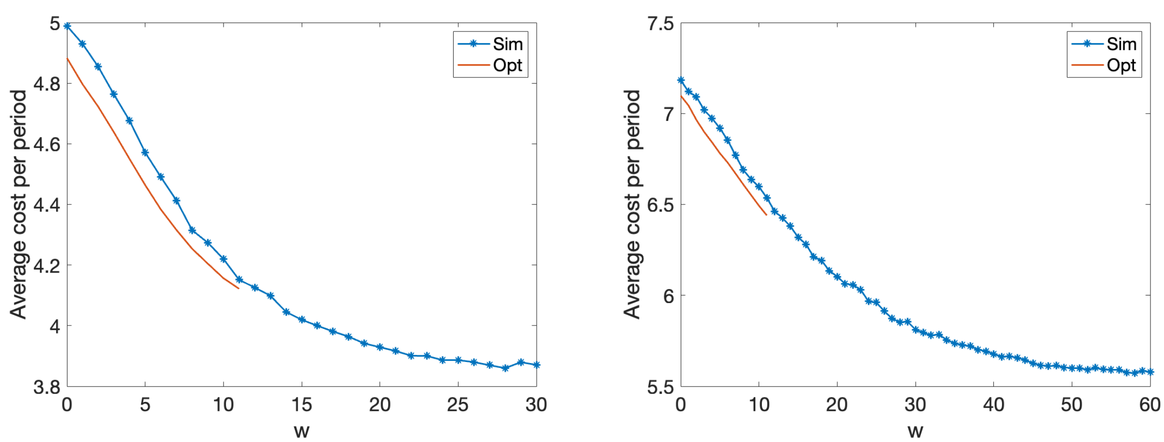
(a) $p/h = 30$ (b) $p/h = 60$ (c) $p/h = 90$

Figure 2: Average cost per period under the simulation-based heuristic (the bounded congestion time policy), the optimal policy and the lookahead policy for $\rho = 1.2$

Figure 2 considers an overloaded system, in which the travel intensity $\rho = 1.2$. A large fraction of incoming customers have to be rejected to balance the load of the system. Figure 2 shows that the same observations discussed above are also valid for an overloaded system. The value of the near future information is more prominent in an over-loaded system. The average cost reduces by more than 12% with a lookahead period $w = 6$ for the case of $p/h = 30$.

Next, we compare the performance of the bounded congestion time policy and the optimal policy in a non-stationary system. Specifically, we consider a system with periodic arrival and service rates with a period of 60. The system is underloaded with traffic intensity $\rho = 0.9$ for 40 periods and then switches to an overloaded state with traffic intensity $\rho = 1.2$ for next 20 periods. We still consider the case when the system manager knows the exact numbers of arrivals and service completion for w-periods. She has no additional information other than knowing the arrival and service rates afterwards.

The estimated remaining congestion time in the simulation-based heuristic is also computed under the no-rejection policy π. Since the system manager has the accurate information for next w periods, the estimated remaining congestion time is deterministic if the system is empty within next w periods. If the system is not empty for all next w periods, we generate 20 sample paths to estimate the remaining congestion time. The threshold $\bar{t}$ is set to be $c_t p/h$, where the coefficient c_t is enumerated from 1.0 to 2.0 with an increment of 0.1. The optimal policy is computed via an associated MDP with an additional state for the periodic arrival and service rates. Thus, we are only able to compute the optimal policy up to $w = 11$.



(a) $p/h = 30$

(b) $p/h = 60$

Figure 3: Average cost per period under the simulation-based heuristic (the bounded congestion time policy) and the optimal policy.

Figure 3 compares the performance of the bounded congestion time policy and the optimal policy. The maximum optimality gap is 2.8% for all cases considered. The maximum optimality gap occurs when the lookahead period w is close to zero. This is because the bounded congestion time policy reduces to a fixed threshold policy when $w = 0$ and ignores the non-stationarity in the arrival and service rates. As the lookahead period w increases, the future information captures the time-dependence in the arrival and service rates. Since the bounded congestion time policy reduces to the lookahead policy in the full information case, it is optimal when the lookahead window w equals p/h. Thus, it is expected that the performance gap between the bounded congestion time policy and the optimal policy shinks as the lookahead period w increases.

5.2 A Noisy Forecast Model

In practice, it is hard to collect accurate information about the future arrivals and service completion, but the system manager may be able to obtain noisy predictions. We consider a scenario that is similar to the setting studied in Figure 1. However, the system manager obtains noisy predictions of the number of arrivals and service completion for w period with 90% of accuracy. Specifically, in period t, the system manager obtains a w-dimensional signal vector $(s_{t+1}, \cdots, s_{t+w})$ for next w periods in the future, where s_t indicates whether an arrival or service completion event happens in period t. The signal s_t for period t is observed in period $t - w$ and is not updated in periods $t - w + 1, \ldots, t - 1$. In period t, the system manager observes the actual realization of the event. If s_t indicates an arrival in period, then the actual realization of the event is an arrival with probability 0.9, i.e., $(\lambda_t, \mu_t) = (1, 0)$ with probability 0.9 and vise versa.

Similarly, we assume a no-rejection policy π in the bounded congestion time policy. In each period, we generate 20 sample paths to estimate the expected remaining congestion time using the noisy signal vector $(s_{t+1}, \cdots, s_{t+w})$. The threshold $\bar{t}$ is set to be $c_t p/h$, where the coefficient c_t is enumerated from 1.0 to 2.0 with an increment of 0.1. The optimal policy is computed via an associated MDP with the state of the system being the signal vector $(s_{t+1}, \cdots, s_{t+w})$.

Figure 4 shows that the bounded congestion time policy performs well. The maximum optimality gap is 3% for all cases considered. Another observation from Figure 4 is that the marginal value of additional signals reduces significantly when the lookahead window w exceeds 10 periods, whereas this happens when the lookahead window w is close to 30 if the forecast is accurate. With noisy forecasts, the cost saving from a 10-period lookahead window is 5.8%, and the saving reduces to merely 6.8% for a 30-period lookahead window. Thus, the value of near future information is still prominent even with noisy predictions.

However, the same amount of cost reduction achieved with noisy predictions for a lookahead window

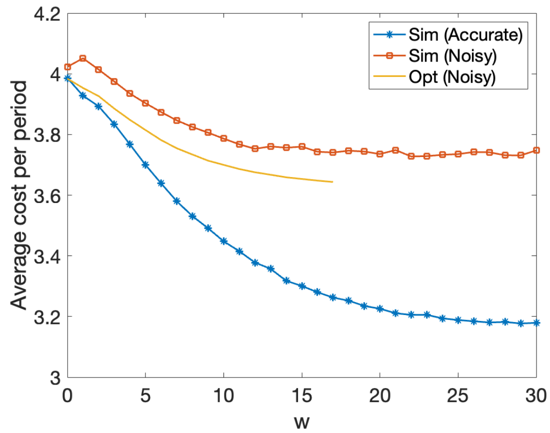


Figure 4: Average cost per period under the simulation-based heuristic (the bounded congestion time policy) and the optimal policy with noisy forecast for $p/h = 30$ and $\rho = 0.9$.

of $w = 30$ periods can be attained with accurate forecasts for a lookahead window of only $w = 4$ periods. Although the forecast for each period may have high accuracy, the accumulation of noises quickly degrades the quality of the prediction as the lookahead window expands. This results in a noisy estimate of the expected remaining congestion time, even for a relatively small lookahead window. Therefore, instead of relying on noisy forecasts for a large lookahead window, it may be more beneficial to invest in obtaining more accurate forecasts for a smaller lookahead window, as suggested by our numerical study.

6 Concluding Remarks

This paper studies the optimal admission policy for a system when either noisy or accurate future arrival and service information is available. We propose a framework to analyze this problem, incorporating general forecast models, and we provide an optimality condition for the admission policy. Additionally, we introduce a simulation-based heuristic that can be easily implemented with general forecast models and demonstrates strong numerical performance in various settings.

There are many possible avenues for future work. Given that many service and healthcare systems involve multiple types of customers, one possible direction is to extend the analysis to multi-class systems where customers exhibit heterogeneous delay sensitivity. Another direction would be to extend the framework to study other critical dynamic decisions in service and healthcare systems, such as scheduling and routing. It would be interesting to analyze and construct simple heuristics for these decisions that incorporate general forecast information. One limitation of this paper is that the service process is assumed to be exogenously given. In practice, service capacities, and thus the service process, are also determined based on forecasts of arrivals. Therefore, another potential extension is to allow for an endogenous service process.

References

Abouee-Mehrizi, H., M. Mirjalili, V. Sarhangian. 2022. Data-driven platelet inventory management under uncertainty in the remaining shelf life of units. *Production and Operations Management* **31**(10) 3914–3932.

Aksin, Z., M. Armony, V. Mehrotra. 2007. The modern call-center: A multi-disciplinary perspective on operations management research. *Production Oper. Management* **16** 665–688.

Ata, B., X. Peng. 2020. An optimal callback policy for general arrival processes: A pathwise analysis. *Operations Research* **68**(2) 327–347.

Ata, B., S. Shneorson. 2006. Dynamic control of an m/m/1 service system with adjustable arrival and service rates. *Management Science* **52**(11) 1778–1791.

Audina, M., K. Ramanan. 2011. Asymptotically optimal controls for time-inhomogeneous networks. *SIAM Journal on Control and Optimization* **49**(2) 611–645.

Avramidis, A. N., A. Deslauriers, P. L'Ecuyer. 2004. Modeling daily arrivals to a telephone call center. *Management Science* **50**(7) 896–908.

Brown, D. B., J. E. Smith. 2022. *Information Relaxations and Duality in Stochastic Dynamic Programs: A Review and Tutorial*. Now Foundations and Trends.

Brown, D. B., J. E. Smith, P. Sun. 2010. Information relaxations and duality in stochastic dynamic programs. *Operations Research* **58**(4-part-1) 785–801.

Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, L. Zhao. 2005. Statistical analysis of a telephone call center. *Journal of the American Statistical Association* **100**(469) 36–50.

Carpentier, P., J.-P. Chancelier, G. Cohen, M. De Lara. 2015. *Stochstic multi-stage optimization*. Springer.

Chen, J., J. Dong, P. Shi. 2024a. Optimal routing under demand surges: The value of future arrival rates. *Operations Research* **Forthcoming**.

Chen, X., Y. Liu, G. Hong. 2024b. An online learning approach to dynamic pricing and capacity sizing in service systems. *Working Paper* .

Deo, S., I. Gurvich. 2011. Centralized vs. decentralized ambulance diversion: A network perspective. *Management Sci.* **57**(7) 1300–1319.

Gans, N., G. Koole, A. Mandelbaum. 2003. Telephone call centers: Tutorial, review and research prospects. *Manufacturing Service Oper. Management* **5** 73–141.

Green, L. V., P. Kolesa. 1991. The pointwise stationary approximation for queues with nonstationary arrivals. *Management Sci.* **37**(1) 84–97.

Green, L. V., P. Kolesa, W. Whitt. 2007. The pointwise stationary approximation for queues with nonstationary arrivals. *Production and Oper. Management* **16**(1) 13–39.

Gurvich, I., O. Perry. 2012. Overflow networks: Approximations and implications to call-center outsourcing. *Oper. Res.* **60**(4) 996–1009.

Helm, J. E., M. S. Lavieri, M. P. Van Oyen, J. D. Stein, D. C. Musch. 2015. Dynamic forecasting and control algorithms of glaucoma progression for clinician decision support. *Operations Research* **63**(5) 979–999.

Ibrahim, R., H. Ye, P. L'Ecuyer, H. Shen. 2016a. Modeling and forecasting call center arrivals: A literature survey and a case study. *International Journal of Forecasting* **32**(3) 865–874.

Ibrahim, R., P. L'Ecuyer. 2013. Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models. *Manufacturing & Service Operations Management* **15**(1) 72–85.

Ibrahim, R., P. L'Ecuyer, H. Shen, M. Thiongane. 2016b. Inter-dependent, heterogeneous, and time-varying service-time distributions in call centers. *Eur. J. Oper. Res.* **250** 480–492.

Jones, S., R. Evans, T. Allen, P. Haug, S. Welch, G. Snow. 2009. A multivariate time series approach to modeling and forecasting demand in the emergency department. *J Biomed Inform.* **42**(1) 123–139.

Lewis, M. E., H. Ayhan, R. D. Foley. 1999. Bias optimality in a queue with admission control. *Probability in the Engineering and Informational Sciences* **13**(3) 309–327. doi:10.1017/S0269964899133047.

Lewis, M., H. Ayhan, R. Foley. 2002. Bias optimal admission policies for a nonstationary multiclass queueing system. *Journal of Applied Probability* **39**(1) 20–37.

Naor, P. 1969. The regulation of queeu size by levying tolls. *Econometrica* **37**(1) 15–24.

Oreshkin, B. N., N. Réegnard, P. L'Ecuyer. 2016. Rate-based daily arrival process models with application to call centers. *Oper. Res.* **64**(2) 510–527.

Shapiro, A., D. Dentcheva, R. A. 2014. *Lectures on Stochastic Programming: Modeling and Theory*. 2nd ed. SIAM, Philadelphia, NY.

Shi, P., M. Chou, J. Dai, D. Ding, J. Sim. 2016. Models and insights for hospital inpatient operations: Time-dependent ed boarding time. *Management Sci.* **62**(1) 1–28.

Spencer, J., M. Sudan, K. Xu. 2014. Queueing with future information. *Ann. Appl. Probab.* **24**(5) 2091–2142.

Stidham, S. 1985. Optimal control of admission to a queueing system. *IEEE Transactions on Automatic Control* **30**(8) 705–713. doi:10.1109/TAC.1985.1104054.

Stidham, S. 2002. Analysis, design, and control of queueing systems. *Operations Research* **50**(1) 197–216. doi:10.1287/opre.50.1.197.17783. URL `https://doi.org/10.1287/opre.50.1.197.17783`.

Sun, Y., B. Heng, Y. Seow, E. Seow. 2009. Forecasting daily attendances at an emergency department to aid resource planning. *BMC Emerg Med* 9:1.

Wang, C. 2016. On socially optimal optimal queue length. *Management Sci.* **62**(3) 899–903.

Whitt, W. 2018. Queues with time-varying arrival rates: A bibliography. *Working paper* .

Xu, K., C. Chan. 2016. Using future information to reduce waiting times in the emergency department via diversion. *Manufacturing Service Oper. Management* **18**(3) 314–331.

Yoon, S., M. Lewis. 2004. Optimal pricing and admission control in a queueing system with periodically varying parameters. *Queueing Systems* **47** 177–199.

Zayas-Cabán, G., M. E. Lewis. 2020. Admission control in a two-class loss system with periodically varying parameters and abandonments. *Queueing Systems* **94**(1–2).

Appendix

A Proofs for Results in Section 4

Proof of Theorem 1:

Let $V(q_{t-1}, \omega_t)$ denote the value-to-go function under the optimal policy. Thus, the one-step recursive equation, i.e., the Bellman equation, to solve the optimal value-to-go function $V_t(q_{t-1}, \omega) \in \mathcal{F}_t$ is given as follows: For $t = 1, \ldots, T-1$,

$$V_t(q_{t-1}, \omega) = \min_{d_t, \mu_t} \; hq_t + pd_t + \mathcal{V}_{t+1}(q_t, \omega), \tag{A.1}$$

$$\text{subject to } -q_{t-1} + q_t + d_t + \mu_t = \lambda_t, \tag{A.2}$$

$$q_t \geq 0, \; d_t \in [0, \lambda_t], \; \mu_t \in [0, \bar{\mu}_t] \tag{A.3}$$

where $\mathcal{V}_{t+1}(q_t, \omega) = \mathbb{E}[V^*_{t+1}(q_t, \omega)|\mathcal{F}_t]$ and $V^*_{T+1}(q) = hq^2/2 - hq$. Note that the function $\mathcal{V}_t(q_{t-1}, \omega)$ is the expectation of the value-to-go function V_t before observing the realization of the arrivals and service completion in period t. Therefore, we have that $\mathcal{V}_t(q_{t-1}, \omega) \in \mathcal{F}_{t-1}$. For any given period t, we call the problem in equations (A.1)-(A.2) as problem (P).

We denote the optimal policy that solve this problem by π^*.

We prove both parts together by backward induction. We first verify that $V_T(q, \omega)$ is increasing and convex in q. Moreover, there exists a threshold policy characterized by a constant $n_T(\omega)$ is optimal. Note that in period T, the problem (P) becomes a deterministic convex optimization, stated as follows:

$$V^*_T(q_{T-1}, \omega) = \min_{q_T, d_T, \mu_T} \; \frac{1}{2}hq_T^2 + pd_T,$$

$$\text{subject to } q_T + d_T + \mu_T = \lambda_T + q_{T-1},$$

$$q_T \geq 0, \; d_t \in [0, \lambda_T], \; \mu_T \in [0, \bar{\mu}_T]$$

It follows from the Karush–Kuhn–Tucker conditions that the optimal solution of problem (P) is given as follows:

$$(d_T^*, \mu_T^*) = \begin{cases} (\lambda_T, \bar{\mu}_T), & \text{if } q_{T-1} - \bar{\mu_T} > p/h, \\ (q_{T-1} + \lambda_T - \bar{\mu}_T - p/h, \bar{\mu}_T), \bar{\mu}_T), & \text{if } q_{T-1} - \bar{\mu_T} \leq p/h \leq q_{T-1} + \lambda_T - \bar{\mu}_T, \\ (0, \bar{\mu}_T), & \text{if } 0 \leq q_{T-1} + \lambda_T - \bar{\mu}_T < p/h, \\ (0, q_{T-1} + \lambda_T), & \text{if } q_{T-1} + \lambda_T - \bar{\mu}_T < 0, \end{cases}$$

and $q_T^* = q_{T-1} + (\lambda_T - d_T^*) - \mu_T^*$. Substituting the optimal solution into the objective function gives the optimal value function $V_T^*(q_{T-1}, \omega)$ as follows:

$$V_T^*(q_{T-1}, \omega) = \begin{cases} \frac{1}{2}h(q_{T-1} - \bar{\mu}_T)^2 + p\lambda_T, & \text{if } q_{T-1} - \bar{\mu_T} > p/h, \\ \frac{p^2}{2h} + p(q_{T-1} + \lambda_T - \bar{\mu}_T - p/h), & \text{if } q_{T-1} - \bar{\mu_T} \leq p/h \leq q_{T-1} + \lambda_T - \bar{\mu}_T, \\ \frac{1}{2}h(q_{T-1} + \lambda_T - \bar{\mu}_T)^2, & \text{if } 0 \leq q_{T-1} + \lambda_T - \bar{\mu}_T < p/h, \\ 0, & \text{if } q_{T-1} + \lambda_T - \bar{\mu}_T < 0. \end{cases}$$

It is easy to see that $V_T(q, \omega)$ is increasing and convex in q. Moreover, the optimal solution is characterized by the threshold $n_T^*(\omega) = p/h$.

Now we assume that $V_{t+1}(q_t, \omega)$ is increasing and convex in q. Moreover, there exists an optimal threshold policy $n_{t+1}^*(\omega)$. By the induction assumptions $\mathcal{V}_t(q_t, \omega) = \mathbb{E}[V_{t+1}(q_t, \omega|\mathcal{F}_t]$ is increasing and convex in q_t.

First, we verify that $V_t(q_{t-1}, \omega)$ is increasing and convex in q_{t-1}. Consider $q_{t-1}^1 \geq q_{t-1}^2 \geq 0$ and fix ω_t. Let (d_1, μ_1, q_t^1) be an optimal solution of the problem starting with q_{t-1}^1. Note that $q_t^1 = q_{t-1}^1 + \lambda_t - d_1 - \mu_1$. We can construct a feasible solution $(\tilde{d}, \tilde{\mu}, \tilde{q}_t^2)$ for the problem with the initial value of q_{t-1}^2. To be specific, let $\tilde{d} = d_1$, $\tilde{\mu} = \min(q_{t-1}^2 + \lambda_t - \tilde{d}, \mu_1)$ and $\tilde{q}_t^2 = q_{t-1}^2 + \lambda_t - \tilde{d} - \tilde{\mu}$. Note that

$$\tilde{q}_t^2 \leq q_{t-1}^2 + \lambda_t - d_1 - \mu_1 \leq q_t^1.$$

Therefore, the following holds:

$$V_t(q_{t-1}^2, \omega) \leq h\tilde{q}_t^2 + p\tilde{d} + \mathbb{E}[\mathcal{V}_t(\tilde{q}_t^2, \omega_{t+1})|\mathcal{F}_t] \leq hq_t^1 + pd_1 + \mathbb{E}[\mathcal{V}_t(q_t^1, \omega_{t+1})|\mathcal{F}_t] = V_t(q_{t-1}^1, \omega_t).$$

Thus, the value-to-go function $V_t(q_{t-1}, \omega)$ is increasing in q_{t-1}.

Now let the optimal solution of the problem starting with q_{t-1}^2 be (d_2, μ_2, q_t^2). Then $(\tilde{d}_3, \tilde{\mu}_3, \tilde{q}_t^3)$ is a feasible solution of the problem with initial value q_{t-1}^3, where $q_{t-1}^3 = \alpha q_{t-1}^1 + (1-\alpha)q_{t-1}^2$ for some $\alpha \in [0, 1]$. Now consider $\tilde{d}_3 = \alpha d_1 + (1-\alpha)d_2$ and $\tilde{\mu}_3 = \alpha\mu_1 + (1-\alpha)d_2$ and $\tilde{q}_t^3 = q_{t-1}^3 + \lambda_t - \tilde{d}_3 - \tilde{\mu}_3 = \alpha q_t^1 + (1-\alpha)q_t^2$. Note that $(\tilde{d}_3, \tilde{\mu}_3, \tilde{q}_t^3)$ is a feasible solution to the problem with initial value q_{t-1}^3. Therefore, the following holds:

$$\begin{aligned}V_t(q^3_{t-1},\omega) &\le h\tilde{q}^3_t + p\tilde{d}_3 + V_t(\tilde{q}^3_t,\omega)\\ &= \alpha(hq^1 + pd_1) + (1-\alpha)(hq^2 + pd_2) + V_t(\alpha q^1_t + (1-\alpha)q^2_t,\omega)\\ &\le \alpha V_t(q^1_{t-1},\omega) + (1-\alpha)V_t(q^2_{t-1},\omega).\end{aligned}$$

Thus, $V_t(q_{t-1},\omega)$ is convex in q_{t-1}.

Next we show that the optimal policy can be characterized by a simple threshold $n^*_t(\omega) \in \mathcal{F}_t$. To facilitate the analysis to follow, define a function $f_t(q,d,\mu)$ for $q \ge 0$, $d \in [0,\lambda_t]$ and $\mu \in [0,\bar{\mu}_t]$ such that $d + \mu \le q_{t-1} + \lambda_t$ as follows:

$$f_t(q_{t-1},d,\mu) = h(q_{t-1} + \lambda_t - d - \mu) + pd + \mathcal{V}_t(q_{t-1} + \lambda_t - d - \mu,\omega).$$

Therefore, we have that

$$V_t(q_{t-1},\omega) = \min_{d\in[0,\lambda_t],\mu\in[0,\bar{\mu}_t]} f_t(q_{t-1},d,\mu).$$

It follows the convexity of the function $\mathcal{V}_t$ that $f(q_{t-1},d,\mu)$ is convex in d and μ. Moreover, it follows from the monotonicity of the function $\mathcal{V}_t$ that the function $f(q_{t-1},d,\mu)$ decreases in μ. Therefore, to find the optimal values of d and μ, we can first fix d and optimize μ. Note that given a value of $d \in [0,\lambda]$, the optimal μ^* is given by $\mu^*(d) = \min(\bar{\mu}_t, q_{t-1} + \lambda_t - d))$. Substituting the optimal $\mu^*(d)$ info the function f_t, we obtain that

$$f_t(q_{t-1},d,\mu^*(d)) = h(q_{t-1} + \lambda_t - d - \bar{\mu}_t)^+ + pd + \mathcal{V}_t((q_{t-1} + \lambda_t - d - \bar{\mu}_t)^+,\omega).$$

If $q_{t-1} + \lambda_t - \bar{\mu}_t \le 0$, then $f_t(q_{t-1},d,\mu^*(d))$ strictly increases in d. Thus, $d^* = 0$ and $\mu^*(0) = q_{t-1} + \lambda_t$.

If $q_{t-1} + \lambda_t - \bar{\mu}_t > 0$, then the function $f_t(q_{t-1},d,\mu^*(d))$ strictly increases in d if $d > q_{t-1} + \lambda_t - \bar{\mu}_t$. Thus, $d^* \le q_{t-1} + \lambda_t - \bar{\mu}_t$ and $\mu^*(d) = \bar{\mu}_t$. If we ignore the constraints that $d \in [0,\lambda_t]$, then the optimal $d^* \in [0, q_{t-1} + \lambda_t - \bar{\mu}_t]$ satisfies

$$0 \in (p-h) - \partial\mathcal{V}_t(q_{t-1} + \lambda_t - d - \bar{\mu}_t,\omega),$$

where $\partial\mathcal{V}_t$ is the subdifferential of the function $\mathcal{V}_t$. Now let $n^*_t(\omega)$ be the threshold that $p-h \in \partial\mathcal{V}_t(n^*_t(\omega),\omega)$. Since $\mathcal{V}_t(q_t,\omega) \in \mathcal{F}_t$, the optimal threshold $n^*_t(\omega) \in \mathcal{F}_t$.

If $q_{t-1} + \lambda_t - \bar{\mu}_t \le n^*_t(\omega)$, then $f(q_{t-1},d,\bar{\mu}_t)$ is increasing in d for $d \in [0,\lambda_t]$. We have that $d^* = 0$. If $q_{t-1} + \lambda_t - \bar{\mu}_t > n^*_t(\omega) > q_{t-1} - \bar{\mu}_t$, we set $d^* = q_{t-1} + \lambda_t - \bar{\mu}_t - n^*_t(\omega)$. If $n^*_t(\omega) \le q_{t-1} - \bar{\mu}_t$, then $f(q_{t-1},d,\bar{\mu}_t)$ is decreasing in d for $d \in [0,\lambda_t]$. Therefore, $d^* = \lambda_t$. This completes the proof that there exists a threshold policy that is optimal. □

Proof of Theorem 2: We start by proving part (ii) of the lemma because we can directly apply Theorem 3.7 of Shapiro et al. (2014) to prove the theorem. First, we set up the problem using the framework of Chapter 3.2.1 of Shapiro et al. (2014). First we fix the time t and re-write the constraints of problem (P) as follows:

$$
\begin{aligned}
-q_{t-1} + q_t + d_t + \mu_t &= \lambda_t, \\
d_t + u_t &= \lambda_t, \\
\mu_t + w_t &= \bar{\mu}_t, \\
q_t, d_t, \mu_t, u_t, w_t &\geq 0.
\end{aligned}
$$

We define the vectors $x_{t-1} = q_{t-1}$, $x_t = (q_t, d_t, \mu_t, u_t, w_t)$, $b_t = (\lambda_t, \lambda_t, \bar{\mu}_t)$. In addition, let

$$
B_t = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad A_t = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}.
$$

Then the constraints can be written as $B_t q_{t-1} + A_t x_t = b_t$. In addition, define the function $f_t(q, d, \omega_t)$ as follows: For $t = 1, \ldots, T-1$,

$$
f_t(x_t, \omega_t) = \begin{cases} hq_t + pd_t, & \text{if } x_t \geq 0, \\ +\infty, & \text{otherwise.} \end{cases}
$$

and

$$
f_T(x_T, \omega_T) = \begin{cases} \frac{hq_T^2}{2} + pd_T, & \text{if } x_T \geq 0, \\ +\infty, & \text{otherwise.} \end{cases}
$$

Note that $f_t(x_t, \omega_t)$ is convex in x_t a.s. Thus, the functions Q_t and $\mathcal{D}_{t+1}$ in Shapiro et al. (2014) are the value-to-go function V_t and the function $\mathcal{V}_{t+1}$ in our problem.

To apply Theorem 3.7 in Shapiro et al. (2014), we verify that the two conditions (D2) and (D3) are satisfied. (D2) states that for any small perturbation of b_t, the value-to-go function $V_t(q_{t-1}, \omega_t)$ is finite for all q_{t-1} for all $q_{t-1} \geq 0$ for $t = 1, \ldots, T$. To verify (D2), consider b'_t that is a small perturbation of b_t. Note that the proof of Theorem 3.7 in Shapiro et al. (2014) relies on the fact that the function $\theta(y)$ is continuous and subdifferntiable at point y. In our problem, $y = [q_{t-1} + \lambda_t, \lambda_t, \bar{\mu}_t] \geq 0$ for all $q_{t-1} \geq 0$ and ω_t. Thus, we only need to consider the perturbation b' such that $\lambda'_t \geq 0$, $\mu'_t \geq 0$ and $q_{t-1} + \lambda'_t \geq 0$. (D3) states that $\mathcal{V}_t(q_{t+1}, \omega_{t+1})$ is finite for $t = 1, \ldots, T-1$. In particular, we provide an explicit bound of $\mathcal{V}_t(q_{t+1}, \omega_{t+1})$ and want to show that

$$\mathcal{V}_t(q_{t-1},\omega_{t-1}) \le hq_{t-1}(T-t) + \frac{1}{2}hq_{t-1}^2 + p\mathbb{E}[\lambda_t|\mathcal{F}_{t-1}] + p\sum_{s=t}^{T}\mathbb{E}[\lambda_s|\mathcal{F}_{t-1}].$$

We can show the two conditions by backward induction. First, we check that both conditions (D2) and (D3) hold for $t = T$. Note that (q_T, d_T, μ_T) is a feasible solution to the problem with the perturbation b', where $q_T = q_{T-1} + \lambda'_T \ge 0$ and $d_T = \bar{\mu}_T = 0$. Thus, the value-to-go function with the perturbation b' is less than or equal to $\le h(q_{T-1} + \lambda_T)^2/2$, which is finite. In other words, (D2) holds for $t = T$. Next, we check condition (D3). Note that for any q_{T-1} and ω_T, we have that (q'_T, d'_T, μ'_T) is feasible, where $q'_T = q_{T-1}$, $d'_T = \lambda'_T$ and $\mu'_T = 0$. Therefore, the following holds:

$$\mathcal{V}_T(q_{T-1},\omega_{T-1}) = \mathbb{E}[V_T(q_T,\omega(T))|\mathcal{F}_{T-1}] \le \mathbb{E}\left[\frac{1}{2}hq_{T-1}^2 + p\lambda_T\middle|\mathcal{F}_{T-1}\right] = \frac{1}{2}hq_{T-1}^2 + p\mathbb{E}\left[\lambda_T\middle|\mathcal{F}_{T-1}\right].$$

Now suppose that (D2) and (D3) hold for $t+1,\ldots,T$ and we want to verify that it holds for period t using similar arguments. Consider b'_t that is a small perturbation of b_t. Note that $q_t = q_{t-1} + \lambda'_t$ and $d_t = \bar{\mu}_t = 0$ is a feasible solution. Thus, the value-to-go function $V_t(q_{t-1},\omega_t)$ with the perturbation b'_t is always less than or equal to

$$h(q_{t-1} + \lambda'_t)^2/2 + \mathcal{V}_t(q_{t-1} + \lambda'_t, \omega_t) < \infty,$$

by the induction assumption. Thus, condition (D2) holds for period t.

Note that for any q_{t-1} and ω_t, we have that $q_t = q_{t-1}$, $d_t = \lambda_t$ and $\mu_t = 0$ is a feasible solution to problem (P). Therefore, the following holds:

$$\begin{aligned}
\mathcal{V}_t(q_{t-1},\omega_{t-1}) &= \mathbb{E}[V_t(q_t,\omega_t)|\mathcal{F}_{t-1}] \\
&\le \mathbb{E}\left[hq_t + pd_t + \mathcal{V}_{t+1}(q_t,\omega_t)\middle|\mathcal{F}_{t-1}\right] \\
&= \mathbb{E}\left[hq_{t-1} + p\lambda_t + \mathcal{V}_{t+1}(q_{t-1},\omega_t)\middle|\mathcal{F}_{t-1}\right] \\
&\le hq_{t-1} + p\mathbb{E}[\lambda_t|\mathcal{F}_{t-1}] + \mathbb{E}\left[hq_{t-1}(T-t-1) + \frac{1}{2}hq_{t-1}^2 + p\sum_{s=t+1}^{T}\mathbb{E}[\lambda_s|\mathcal{F}_t]\middle|\mathcal{F}_{t-1}\right] \\
&= hq_{t-1}(T-t) + \frac{1}{2}hq_{t-1}^2 + p\sum_{s=t}^{T}\mathbb{E}[\lambda_s|\mathcal{F}_{t-1}].
\end{aligned}$$

The inequality on the third line follows from the induction assumption. This completes the verification of conditions (D2) and (D3).

Thus, it follows from Theorem 3.7 of Shapiro et al. (2014) states that $(q_t^{\pi^*}, d_t^{\pi^*}, \mu_t^{\pi^*})$ is optimal if there exist a process y_t such that $y_t \in \mathcal{F}_t$ and satisfies the following:

$$0 \in \mathcal{N}_{\mathbb{R}_+}(q_t^{\pi^*}) + h - y_t + \mathbb{E}[y_{t+1}|\mathcal{F}_t], \tag{A.4}$$

$$0 \in \mathcal{N}_{[0,\lambda_t]}(d_t^{\pi^*}) + p - y_t, \tag{A.5}$$

$$0 \in \mathcal{N}_{[0,\bar{\mu}_t]}(\mu_t^{\pi^*}) - y_t, \tag{A.6}$$

where $\mathcal{N}_C(x)$ is the normal cone to set C at a point $x \in C$. Note that $\mathcal{N}_{\mathbb{R}_+}(q_t^{\pi^*}) = \{0\}$ if $q_t^{\pi^*} > 0$ and $\mathcal{N}_{\mathbb{R}_+}(q_t^{\pi^*}) = \mathbb{R}_-$ if $q_t^{\pi^*} = 0$. Thus, equation (A.4) implies that the measurable process y_t satisfies the following:

$$y_t = h + \mathbb{E}[y_{t+1}|\mathcal{F}_t] \text{ if } q_t^{\pi^*} > 0,$$

$$y_t \leq h + \mathbb{E}[y_{t+1}|\mathcal{F}_t] \text{ if } q_t^{\pi^*} = 0.$$

Note that

$$\mathcal{N}_{[0,\lambda_t]}(x) = \begin{cases} \mathbb{R}_-, & \text{if } x = 0, \\ \{0\}, & \text{if } x \in (0, \lambda_t), \\ \mathbb{R}_+, & \text{if } x = \lambda_t. \end{cases}$$

Equation (A.5) implies that

$$\begin{aligned} &\text{if } y_t^* > p, && \text{then } d_t^{\pi^*} = \lambda_t; \\ &\text{if } y_t^* = p, && \text{then } d_t^{\pi^*} \in [0, \lambda_t]; \\ &\text{if } y_t^* < p, && \text{then } d_t^{\pi^*} = 0. \end{aligned}$$

Similarly, it follows from (A.6) and $y_t \geq 0$ that

$$\begin{aligned} &\text{if } y_t^* > 0, && \text{then } \mu_t^{\pi^*} = \bar{\mu}_t; \\ &\text{if } y_t^* = 0, && \text{then } \mu_t^{\pi^*} \in [0, \bar{\mu}_t]; \\ &\text{if } y_t^* < 0, && \text{then } \mu_t^{\pi^*} = \bar{\mu}_t. \end{aligned}$$

Next, we show part (i) in two steps. Fixing t, q_{t-1} and ω_t, we want to show that any for any $\phi \in \Phi$, $V_t(q_{t-1}, \omega_t) \geq D_t^{\phi}(q_{t-1}, \omega_t)$. Second, let $\phi^* = (y_t^*, \dots, y_T^*)$ where y_s^* is the stochastic process that satisfies the condition in part (ii), then $V_t(q_{t-1}, \omega_t) = D_t^{\phi^*}(q_{t-1}, \omega_t)$.

We will show that $V_t(q_{t-1}, \omega_t) \geq D_t^{\phi}(q_{t-1}, \omega_t)$ for any $\phi \in \Phi$ by backward induction. For $t = T$, we want to show that $V_T(q_{T-1}, \omega_T) = D_T(q_{T-1}, \omega_T)$. The dual problem (D) is a deterministic convex optimization problem stated as follows:

$$D_T(q_{T-1},\omega_T) = \max_{y_T}\ y_T(q_{T-1}+\lambda_T-\bar{\mu}_T) - (y_T-p)^+\lambda_T - y_T^-\bar{\mu}_T - \frac{y_T^2}{2h}.$$

Since the objective function is a piece-wise linear function, the optimal solution of this problem is

$$y_T^* = \begin{cases} h(q_{T-1}-\bar{\mu}_T), & \text{if } \ q_{T-1}-\bar{\mu_T} > p/h, \\ p, & \text{if } \ q_{T-1}-\bar{\mu_T} \le p/h \le q_{T-1}+\lambda_T-\bar{\mu}_T, \\ h(q_{T-1}+\lambda_T-\bar{\mu}_T), & \text{if } \ 0 \le q_{T-1}+\lambda_T-\bar{\mu}_T < p/h, \\ 0, & \text{if } \ q_{T-1}+\lambda_T-\bar{\mu}_T < 0. \end{cases}$$

Then it follows from the proof of Theorem 1 that $V_T(q_{T-1},\omega_T) = D_T(q_{T-1},\omega_T)$ for all $q_{T-1} \ge 0$ and ω_T.

Now we suppose that in period $t+1$, $V_{t+1}(q_t,\omega_{t+1}) \ge D_{t+1}(q_t,\omega_{t+1})$. We want to show that $V_t(q_{t-1},\omega_t) \ge D_t(q_{t-1},\omega_t)$ for all $q_{t-1} \ge 0$ and ω_t. To this end, we define the associated Lagrangian of the problem (P) in period t by relaxing the constraint (A.2) as follows:

$$L_t(q_t,d_t,\mu_t,y_t) = hq_t + pd_t + \mathcal{V}_{t+1}(q_t,\omega_t) + y_t(\lambda_t + q_{t-1} - q_t - d_t - \mu_t). \tag{A.7}$$

By assumption that $V_{t+1}(q_t,\omega_{t+1}) \ge D_{t+1}(q_t,\omega_{t+1})$, we have that

$$\mathcal{V}_{t+1}(q_t,\omega_{t+1}) \ge \mathbb{E}[D_{t+1}(q_t,\omega_t)|\mathcal{F}_t] \ge q_t\mathbb{E}[y_{t+1}|\mathcal{F}_t] + \mathbb{E}\left[\sum_{s=t+1}^{T} (\lambda_s-\bar{\mu}_s)y_s - \lambda_s(y_s-p)^+ - y_s^-\bar{\mu}_s - \frac{y_T^2}{2h}\middle| \mathcal{F}_t\right],$$

for any feasible solution $\{y_t,\ldots,y_T\} \in \Phi$ to problem (D). Substituting this inequality into (A.7), we obtain that

$$\begin{aligned} L_t(q_t,d_t,\mu_t,y_t) \ge{}& hq_t + pd_t + y_t(\lambda_t + q_{t-1} - q_t - d_t - \mu_t) + q_t\mathbb{E}[y_{t+1}|\mathcal{F}_t] \\ &+ \mathbb{E}\left[\sum_{s=t+1}^{T} (\lambda_s-\bar{\mu}_s)y_s - \lambda_s(y_s-p)^+ - y_s^-\bar{\mu}_s - \frac{y_T^2}{2h}\middle| \mathcal{F}_t\right]. \end{aligned}$$

Let $\mathcal{X} = \mathbb{R}_+ \times [0,\lambda_t] \times [0,\bar{\mu_t}]$. Thus, we have that for all $\{y_t,\ldots,y_T\} \in \Phi$,

$$\begin{aligned} V_t(q_{t-1},\omega_t) \ge{}& \inf_{(q_t,d_t,\mu_t)\in\mathcal{X}} L_t(q_t,d_t,\mu_t,y_t) \\ \ge{}& y_t(q_{t-1}+\lambda_t) + \inf_{(q_t,d_t,\mu_t)\in\mathcal{X}} \{(h - y_t + \mathbb{E}[y_{t+1}|\mathcal{F}_t])q_t + (p-y_t)d_t - y_t\mu_t\} \\ &+ \mathbb{E}\left[\sum_{s=t+1}^{T} (\lambda_s-\bar{\mu}_s)y_s - \lambda_s(y_s-p)^+ - y_s^-\bar{\mu}_s - \frac{y_T^2}{2h}\middle| \mathcal{F}_t\right]. \end{aligned} \tag{A.8}$$

Since y_t satisfies $y_t \le h + \mathbb{E}[y_{t+1}|\mathcal{F}_t]$, the mininimization problem on the right-hand side of equation (A.8) has an optimal solution, which is given as follows:

$$(q_t, d_t, \mu_t) = \begin{cases} (0,0,0), & \text{if } y_t < 0, \\ (0,0,\bar{\mu}_t), & \text{if } 0 \le y_t < p, \\ (0,\lambda_t,\bar{\mu}_t), & \text{if } y_t \ge p. \end{cases}$$

Substituting this solution into the infimum term in (A.8) gives the following:

$$\begin{aligned} &\inf_{(q_t,d_t,\mu_t)\in\mathcal{X}} \{(h - y_t + \mathbb{E}[y_{t+1}|\mathcal{F}_t])q_t + (p - hy_t)d_t - hy_t\mu_t\} \\ &= -(y_t - p)^+\lambda_t - y_t^+\bar{\mu}_t \\ &= -y_t\bar{\mu}_t - (y_t - p/h)^+\lambda_t - y_t^-\bar{\mu}_t. \end{aligned}$$

Substituting this equation into (A.8), we obtain that

$$\begin{aligned} &V_t(q_{t-1},\omega_t) \\ &\ge y_t(q_{t-1}+\lambda_t) - y_t\bar{\mu}_t - (y_t-p)^+\lambda_t - y_t^-\bar{\mu}_t + \mathbb{E}\left[\sum_{s=t+1}^{T}(\lambda_s-\bar{\mu}_s)y_s - \lambda_s(y_s-p)^+ - y_s^-\bar{\mu}_s - \frac{y_T^2}{2h}\middle|\mathcal{F}_t\right] \\ &= y_t q_{t-1} + \mathbb{E}\left[\sum_{s=t}^{T}(\lambda_s-\bar{\mu}_s)y_s - \lambda_s(y_s-p)^+ - y_s^-\bar{\mu}_s - \frac{y_T^2}{2h}\middle|\mathcal{F}_t\right]. \end{aligned}$$

Then, taking the supremum on the right-hand side yields $V_t(q_{t-1},\omega_t) \ge D_t(q_{t-1},\omega_t)$.

We will end the proof by showing that the inequalities becomes qualities if we select $\phi = \phi^*$ which satisfies the conditions in part (ii). It follows from Proposition 3.5 of Shapiro et al. (2014) that

$$\begin{aligned} &V_t(q_{t-1},\omega_t) \\ &= L_t(q_t^*, d_t^*, \mu_t^*, y_t^*) \\ &= y_t^* q_{t-1} + y_t^*(\lambda_t - \bar{\mu}_t) - (y_t^* - p)^+\lambda_t - (y_t^*)^-\bar{\mu}_t + (h - y_t^*)q_t^* + \mathbb{E}\left[V_{t+1}(q_t^*,\omega_{t+1})|\mathcal{F}_t\right] \\ &= y_t^* q_{t-1} + y_t^*(\lambda_t - \bar{\mu}_t) - (y_t^* - p)^+\lambda_t - (y_t^*)^-\bar{\mu}_t + (h - y_t^*)q_t^* + \mathbb{E}\left[L_{t+1}(q_{t+1}^*, d_{t+1}^*, \mu_{t+1}^*, y_t^*)|\mathcal{F}_t\right] \\ &= y_t^* q_{t-1} + y_t^*(\lambda_t - \bar{\mu}_t) - (y_t^* - p)^+\lambda_t - (y_t^*)^-\bar{\mu}_t + (h - y_t^*)q_t^* + \\ &+ \mathbb{E}\left[y_{t+1}^* q_t + y_{t+1}^*(\lambda_{t+1} - \bar{\mu}_{t+1}) - (y_{t+1}^* - p)^+\lambda_{t+1} - (y_{t+1}^*)^-\bar{\mu}_{t+1} + (h - y_{t+1}^*)q_{t+1}^*|\mathcal{F}_t\right] \\ &+ \mathbb{E}[\mathbb{E}[V_{t+1}(q_{t+1}^*, w_{t+2}|\mathcal{F}_{t+1}]|\mathcal{F}_t] \\ &= y_t^* q_{t-1} + \mathbb{E}\left[\sum_{s=t}^{t+1} y_s^*(\lambda_s - \bar{\mu}_s) - (y_s^* - p)^+\lambda_s - (y_s^*)^-\bar{\mu}_s|\mathcal{F}_t\right] \\ &\quad + \left(h - y_t^* + \mathbb{E}[y_{t+1}^*]\right) q_t^* + \mathbb{E}\left[V_{t+1}(q_{t+1}^*, w_{t+2})|\mathcal{F}_t\right] \\ &= y_t^* q_{t-1} + \mathbb{E}\left[\sum_{s=t}^{t+1} y_s^*(\lambda_s - \bar{\mu}_s) - (y_s^* - p)^+\lambda_s - (y_s^*)^-\bar{\mu}_s|\mathcal{F}_t\right] + \mathbb{E}\left[V_{t+1}(q_{t+1}^*, w_{t+2})|\mathcal{F}_t\right] \end{aligned}$$

The equality in last line follow from part (ii) which implies that

$$(h - y_t^* + \mathbb{E}[y_{t+1}^* | \mathcal{F}_t]) q_t^* = 0 \ \text{ a.s.}$$

We can continue the iterations and get that $V_t(q_{t-1}, \omega_t) = D_t(q_{t-1}, \omega_t)$. □

Proof of Lemma 1: We prove the lemma by construction. Let $\phi = \{y_t : t = 1, \ldots, T+1\}$ be an optimal solution of the dual problem (D). Define $\tilde{\phi} = \{\tilde{y}_t : t = 1, \ldots, T+1\}$ such that $\tilde{y}_t = y_t^+$. It is easy to verify that condition (9) holds. Moreover, we have that if $y_t < 0$,

$$(\lambda_t - \bar{\mu}_t) y_t - (y_t - p)^+ \lambda_t - y_t^- \bar{\mu}_t = \lambda_t y_t \leq 0 = (\lambda_t - \bar{\mu}_t) \tilde{y}_t - (\tilde{y}_t - p)^+ \lambda_t - \tilde{y}_t^- \bar{\mu}_t,$$

where the last equation follows from the fact that $\tilde{y}_t = 0$ if $y_t < 0$. Therefore, $D^{\tilde{\phi}}(q_0) \geq D^{\phi}(q_0)$. Given the optimality of ϕ, it must be that $\tilde{\phi}$ is optimal as well. □

Proof of Proposition 1: It follows from equation (10) that

$$y_t^* = h\mathbb{E}[R_t^{\pi^*}(q_t^*) | \mathcal{F}_t] + \mathbb{E}[y_{t_1}^* | \mathcal{F}_t] \geq h\mathbb{E}[R_t^{\pi^*}(q_t^*) | \mathcal{F}_t],$$

where $t_1 = R_t^{\pi^*}(q_t^*)$. The last inequality follows from $\mathbb{E}[y_{t_1}^* | \mathcal{F}_t] \geq 0$ by Lemma 1. Moreover, the following holds:

$$y_t^* \leq h\mathbb{E}[S_t^{\pi^*}(q_t^*) | \mathcal{F}_t] + \mathbb{E}[y_{t_2}^* | \mathcal{F}_t] \leq h\mathbb{E}[S_t^{\pi^*}(q_t^*) | \mathcal{F}_t],$$

where $t_2 = S_t^{\pi^*}(q_t^*)$. The last inequality follows from $\mathbb{E}[y_{t_2}^* | \mathcal{F}_t] = 0$ by Lemma 2. When $q_t^* = n_t^*$, it follows from condition (12) that $y_t^* = p$. □