Dynamic Control of a Make-to-Order System Under Model Uncertainty

Xu Sun, Xiaohan Zhu
Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611

In the area of dynamic control of make-to-order manufacturing systems, optimal control policies are typically derived on the premise that a precise probabilistic model is given. In many situations, however, the underlying model is a simplification of the real-world scenario that arises from tractability considerations and/or due to lack of operational data to calibrate the model. Thus, policies derived from certain simplifications may perform suboptimally if the resulting model does not capture reality well. This paper advances an approach that accounts for probable modeling misspecification in controlling a multi-product make-to-order manufacturing system with an outsourcing mechanism. In particular, it focuses on model uncertainty in demand for different products and introduces a robust control formulation where model uncertainty is represented using the notion of Rényi divergence. Then, using results concerning the change of measure for Poisson point processes, we obtain a more tractable reformulation. Because the reformulation is still complex, we approximate it with a stochastic differential game in which the underlying state processes are driven by Brownian motions. We illustrate how the stochastic differential game can be solved via dynamic programming. Interpreting the solution to the stochastic differential game in the context of the original make-to-order system, we propose a set of implementable control policies. We also present a data-driven approach for selecting the “uncertainty set”. A numerical experiment is conducted to demonstrate the value of building “robustness” into the decision model.

Key words: make-to-order manufacturing; model uncertainty; ambiguity; robust control; heavy-traffic approximations; stochastic differential games
1. Introduction

This paper considers distributionally robust control of a make-to-order manufacturing system producing multiple types of products with a shared capacity. Most items are supposed to be produced in-house, but the system faces the option of subcontracting or outsourcing its manufacturing needs for certain products at a fixed plus proportional outsourcing cost in response to unexpected upticks in demand for the products. This problem can arise in many settings. One example is additive manufacturing, or 3D printing, which utilizes 3D printers to manufacture a variety of products on-demand. A printable product can be printed directly from a computer model, meaning that 3D printing operations can not build inventory to decouple demand variability from production. When there is a surge in demand for a particular product, a subcontractor, such as a professional 3D printing service, can be called upon to supplement in-house manufacturing efforts (Berman 2012). Orders for 3D-printed products may be prioritized based on their relative urgency, and delays in order delivery may incur costs to the manufacturing system.

Make-to-order systems can be treated as multiclass queuing systems, and operational decisions such as order sequencing and outsourcing can be optimized via either stochastic dynamic programming (Carr and Duenyas 2000, Öner-Közen and Minner 2017) or optimal control theory (Plambeck et al. 2001, Ata 2006, Çelik and Maglaras 2008, Rubino and Ata 2009). Although the problem of controlling a make-to-order system has been widely investigated by many scholars, two aspects merit further exploration. First, when subtracting/outsourcing is present as an option, it has often been assumed that only proportional costs will be incurred. In reality, fixed costs can arise with subtracting/outsourcing decisions (Atamtürk and Hochbaum 2001). To our knowledge, few studies have considered fixed costs in such settings, especially in the multi-product setting. Second, the majority of papers pertaining to managing make-to-order systems assume that the probabilistic law governing the realized demand is precisely known; as a result, the real-world performance of the derived policies would depend on how faithful the model is to reality. However, it is a well-known fact that many parametric models used in theory contain significant uncertainties in parameter estimates and
there may be insufficient historical data to calibrate the model. In addition, it is not uncommon that simplifying assumptions such as "Markovian" and "stationarity" are imposed for tractability considerations despite the fact that the real-world situation may be much more complex. For example, real-world arrivals may exhibit time-dependent or autoregressive features (Sobel and Babich 2012), ruling out the use of a Poisson process as a demand model. All the aforementioned scenarios could potentially lead to model misspecification.

A Bayesian framework may be proven useful when exact knowledge of the model parameters is lacking. This approach starts by assuming a parametric family of distributions which the true distribution belongs to without specifying the values of the parameters. The belief concerning the parameter uncertainty is then updated through prior and posterior distributions based on observations (Chen and Plambeck 2008, Bisi et al. 2011). However, the effectiveness of Bayesian models depends on the choice of the prior; hence, they can be effective if one is certain about how to choose the priors. In reality, no one can be so certain about the prior, so wrong values of the parameters may prevail.

Another approach to handling modeling errors is robust optimization (RO). The idea is to assume that the uncertain model parameters belong to a (often prespecified) uncertainty set. We refer to the excellent survey by Bertsimas and Thiele (2006) for early works in this field; see also Mamani et al. (2017), Bandi et al. (2019), Sun and Van Mieghem (2019) for some recent contributions. Although the classical RO approach has proven its value in a variety of applications, it is inherently a static approach, and for this reason, it can sometimes lead to overly conservative solutions for sequential decision problems where not only here-and-now but also wait-and-see decisions are present. At this point, we wish to point out that our representation of model uncertainty differs from a typical RO setting in that our set of alternative models is only vaguely specified and obtained by perturbing the decision maker’s belief to allow random shocks to feed back on state variables in an arbitrary yet time-consistent fashion.

In a nutshell, this paper aims at presenting a framework that seeks good policies for controlling a make-to-order manufacturing system when (a) the decision maker is not perfectly sure about the
probabilistic model to be assumed or (b) the correct model is known but is difficult to describe or calibrated using the available data. As such, the decision maker faces model uncertainty in contemplating control strategies. In this context, a natural approach is to extract a nominal model from the available data, and then penalize other models based on their deviation from the nominal model. To what extent these alternative models are penalized depends on how ambiguity-averse the decision maker is. Early works that incorporate ambiguity aversion into adaptive decision making include Petersen et al. (2000), Hansen and Sargent (2001), Chen and Epstein (2002), Hansen et al. (2006).

In greater detail, we assume that demands for a product arrive according to a non-homogeneous Poisson process. Unlike the standard case, however, it is postulated that the decision maker has no idea of the “true” value of the arrival intensity at any given point in time, beyond knowing that it tends to fluctuate around some long-term average. We capture ambiguity by considering a set of alternative models from which the real-world model is picked by a malevolent agent (“nature”). The decision maker is risk-neutral and ambiguity-averse in the sense that it calculates the expected long-run average cost by assuming the worst-case scenario in the set of alternative models and seeks a joint sequencing and outsourcing strategy to minimize it (minimax criterion). This gives rise to a two-player game in which the decision maker chooses the best response to nature who can alter the underlying probabilistic law within prescribed limits. In this regard, nature’s malevolence is the decision maker’s tool for analyzing the fragility of alternative decision rules.

Because the original formulation is generally complex, we develop an approximating procedure, referred to as the stochastic differential game (SDG), due to direct analysis of the original formulation being analytically intractable. The SDG is derived based on the hypothesis that both the demand and service capacity are large and that server utilization is close to one. The resulting SDG can then be converted to an equivalent one-dimensional differential game, whose state descriptor is a one-dimensional process that tracks the amount of work in the system over time. The conversion reduces the dimension of the problem while dictating how sequencing decisions ought to be made to achieve the lowest holding cost possible. The solution to the one-dimensional problem, comprises of a
control band policy for the decision maker and a drift-rate control for nature. Whenever the workload exceeds an upper barrier, it is pulled back instantly by the decision maker (through outsourcing the manufacturing needs of a particular product) to a lower threshold level. Between two consecutive outsourcing operations, a (state-dependent) drift-rate control is used by nature to resolve ambiguity-aversion.

*The remainder of this paper.* Section 2 reviews related literature. Section 3 introduces the nominal model and formulates the corresponding optimization problem. In Section 4 we introduce the ambiguity aversion model along with a more tractable reformulation obtained via a change of measure. Section 5.1 derives the SDG that serves to approximate the robust control problem. An equivalent workload formulation of the SDG is formulated and solved in sections 5.2 and 5.3 respectively. Section 6 presents numerical results and discusses some qualitative insights. Section 7 extends the Rényi-type ambiguity to a broader class of uncertainty sets. In Section 8 we conclude. Proofs and additional numerical results can be found in the e-companion of this paper.

2. Literature Review

The present research pertains to three different research streams.

*Controlled queues in heavy traffic.* The present study draws on the vast literature on controlling queues in heavy traffic. In the case of Poisson arrivals and linear delay cost rates, the $c\mu$ rule, which assigns static priority levels to jobs in increasing order of their index $c_i\mu_i$, is known to minimize the delay cost (Cox and Smith 1991). Dynamic versions of the $c\mu$ rule are introduced by Van Mieghem (1995) in the context of convex delay costs and Akan et al. (2012) for convex-concave delay costs. Extensions accounting for job abandonments from the queue(s) include Rubino and Ata (2009), Ata and Tongarlak (2013), Kim and Ward (2013), Ghamami and Ward (2013). In the context of managing make-to-order systems, a few papers (Çelik and Maglaras 2008, Ata and Olsen 2013) have considered combining economic levers (e.g., pricing) with operational decisions; an asymptotic analysis of this model class would give rise to a drift-rate term subject to control. From a modeling perspective, our paper differs from the foregoing research in two key aspects. (i) The aforementioned papers assume
the availability of an accurate probabilistic model in optimization, whereas we consider a scenario where model misspecification is not only possible but permissible, hence lending to a min-max optimal control problem. (ii) Existing works that incorporate outsourcing (or job admission) decisions tend to assume proportional costs only, whereas we also consider fixed costs that make up one of the two components of total cost in many practical settings.

Decision making with model perturbations. Our approach of modeling ambiguity pertains to the strand of literature that allows for ambiguous beliefs in sequential decision making by adding a set of perturbed models surrounding a nominal model and a malevolent agent (i.e., nature) who promotes robustness. Particularly, the nominal model is believed to represent the best description of the real-world scenario while perturbations of it serve to accommodate the possibility of model misspecification. This concept can be traced back to the early works of Petersen et al. (2000), Hansen and Sargent (2001), Hansen et al. (2006) and has found applications in a wide range of problem domains, including portfolio optimization (Maenhout 2004), dynamic pricing (Lim and Shanthikumar 2007), corporate investment (Nishimura and Ozaki 2007), probability of lifetime ruin (Bayraktar and Zhang 2015), among others. Usually the problem can be translated to a two-player stochastic game which can be solved through the so-called Isaacs’ equation. The Isaacs’ equation is in essence a nonlinear differential equation, and addressing certain fundamental questions associated with this equation often makes up the major technical hurdles in the analysis. Finally, an optimal solution can be identified based on the solution of this equation. The majority of the papers use the notion of relative entropy to construct the uncertainty set. Few papers have gone beyond entropic constraint or penalty in describing model uncertainty and ambiguity aversion. To the best of our knowledge, this paper is one of the first to apply Rényi divergence to characterize model uncertainty in stochastic dynamic programming problems. Even though Rényi divergence itself is a special family of statistical distance measures, it nonetheless covers or closely links to important divergence measures such the Bhattacharyya distance, Kullback-Leibler divergence, $\chi^2$-divergence, total variation distance, among others. In this respect, we feel that the present paper signifies an important step forward in expanding this methodological framework. A recent paper by Cohen (2019) studies a Brownian control
problem that arises from the heavy-traffic approximation of a multiclass $M/M/1$ queue under model uncertainty. We would like to draw three distinctions between Cohen (2019) and the present paper. (i) Cohen (2019) uses an infinite-horizon discounted cost formulation whereas we adopt a long-run average cost criterion. (ii) In Cohen (2019) the control on the workload process only incurs “proportional cost”; hence the workload control is of a barrier type; by contrast, our control pertaining to the workload involves both fixed and proportional cost, which leads to a control band policy. (iii) Cohen (2019) uses the entropic approach that leads to a quadratic nonlinear term in the Bellman equation. Our Bellman-Isaacs equation, due to the assumed Rényi-type penalties (see also their further extensions in §7), belongs to a more general category — the nonlinear term is not necessarily quadratic. Each of the three distinctions prevents us from adopting the methodological tools developed in Cohen (2019).

Impulse control. From the methodological point of view, the paper is related to impulse control of Brownian systems. Two-sided impulse control of a Brownian motion having constant drift rate has been widely studied in the literature; see, e.g., Constantinides and Richard (1978), Harrison et al. (1983), Dai and Yao (2013b) for discounted cost formulations and Ormeci et al. (2008), Dai and Yao (2013a) for average cost problems. These works establish the optimality of a control band policy $(a_1, a_2, b_1, b_2)$ with $a_1 < a_2 < b_1 < b_2$. Because the drift rate is constant, the optimality proof can be done based on explicit or semi-explicit solutions of optimality equations. However, such a method does not apply to our problem because it is impossible to obtain the explicit solution of the optimality equation in the presence of a (stochastically) varying drift subject to control by nature.

We accomplish the optimality proof by adopting a very different method that does not require an explicit solution to the value function. In studying joint pricing and inventory control problems, a few recent papers (Yao 2017, Cao and Yao 2018) have considered joint drift-rate control and impulse control for Brownian models. Our paper differs from these works in two aspects: (i) Their problems belong to the class of cost minimization problems whereas our problem adopts a min-max criterion. (ii) Their problems are of a single-class nature whereas ours is inherently a multiclass model.
Contributions. Our paper makes three main contributions to the literature.

- First, it presents a decision framework accounting for model misspecification and fixed outsourcing costs for controlling a make-to-order system with an outsourcing mechanism. Model misspecification is captured by using the notion of Rényi divergence; this method, together with its further generalizations, leads to a robust control formulation that extends the well-established entropic approach used in the literature.

- Second, by assuming the manufacturing resources are adequately utilized, we derive and solve an approximate control problem. We demonstrate that a candidate optimal policy for the approximate problem can be found by repetitively solving a Bellman-Isaacs equation a finite number of times, each time using a different set of boundary conditions. For each fixed set of boundary conditions, we prove the existence and uniqueness of the Bellman-Isaacs equation. The optimality of the candidate policy is established via a novel level-set argument. The optimality proof highlights an interesting observation: In order for control band policies to be optimal, the first-order derivative of the relative value function does not have to be concave. In fact, quasi-concavity would suffice.

- Third, it puts forward a data-driven method synergizing an analytical modeling framework with real-world data in a meaningful way to yield good decision strategies. In a nutshell, this method allows a management to rely on a wrong model to make right decisions while offering the management an opportunity to assess the value of the wrong model.

3. Nominal Model

Let $(\Omega, \mathcal{F})$ be a measurable space endowed with a filtration $(\mathcal{F}_t)$ contained in $\mathcal{F}$. We consider a make-to-order manufacturing system that offers $I$ different products, labeled by $i = 1, \ldots, I$. The manufacturing facility is modeled as a multiclass single server queue. Requests for product $i$, interchangeably referred to as class $i$ orders, occur in a random fashion. Specifically, for each $i = 1, \ldots, I$, we use $A_i(t)$ to model the number of class $i$ orders placed up to time $t$ and assume it to follow a Poisson process with some arrival rate $\bar{\lambda}_i$ under a reference measure $\mathbb{P}$; henceforth we regard $\mathbb{P}$ as our nominal model and denote by $\bar{\lambda} := (\bar{\lambda}_i)$ the associated arrival rate vector. We model the number
of class $i$ products manufactured over time if the server were to be continuously working on class $i$ products as a renewal process $S_i(t)$ whose consecutive renewals have a mean $m_i$ and coefficient of variation $\nu_i$ under the measure $\mathbb{P}$. We shall refer to $\mu_i := 1/m_i$ as the production rate of class $i$ orders. The decision maker has discretion over sequencing of orders, but will adhere to the head-of-line sequencing principle within each queue, orders are processed in a first-in-first-out fashion, so the sequencing decisions can be described by an $I$-dimensional time allocation process $T := (T_i)$, where $T_i(t)$ represents the amount of time spent by the server on producing product $i$. The cumulative amount of idle time up to time $t$ can thus be computed as $(t - \sum_i T_i(t))$. The decision maker may also outsource manufacturing needs at a fixed plus proportional cost. In particular, outsourcing a batch size of $x$ class $i$ orders would incur a cost of

$$\phi_i(x) := (L_i + \ell_i x) \cdot 1_{\{x > 0\}} + 0 \cdot 1_{\{x = 0\}};$$

by doing so, the decision maker is able to reduce its backlog of class $i$ orders by $x$ units instantly; hence we shall refer to this type of outsourcing activities as type $i$ outsourcing operations. To proceed, let $\Psi := (\Psi_i)$ whose $i$th component is specified by

$$\Psi_i := (\tau_i(0), \tau_i(1), \tau_i(2), \ldots, \tau_i(k), \ldots; \xi_i(0), \xi_i(1), \xi_i(2), \ldots, \xi_i(k), \ldots),$$

where $0 = \tau_i(0) < \tau_i(1) < \tau_i(2) < \cdots < \tau_i(k) < \cdots$ is a sequence of increasing stopping times at which a type $i$ outsourcing operation is performed, and $\{\xi_i(k); k \geq 0\}$ represents the batch sizes of the consecutive outsourcing operations. We restrict attention to outsourcing policies that are non-anticipating, meaning that the policy can not require knowledge of the future. Therefore, the decision of removing $\xi_i(k)$ orders from queue $i$ must be based solely on historical records up to time $\tau_i(k)$. For notational convenience, we introduce $I$ counting processes given as $N_i(t) := \sup\{k \geq 0 : \tau_i(k) \leq t\}$ for $i = 1, \ldots, I$. Hence $N_i(t)$ represents the number of type $i$ outsourcing operations performed up to $t$.

Let $Q_i(t)$ denote the number of outstanding orders of class $i$ in the system at time $t$, and write $Q(t) := (Q_i(t))$. Assuming that there are initially $Q_i(0)$ class $i$ orders in the system, we can describe the dynamics of $Q_i(t)$ as

$$Q_i(t) = Q_i(0) + A_i(t) - S_i(T_i(t)) - \sum_{k=0}^{N_i(t)} \xi_i(k) \quad t \geq 0,$$
for $i = 1, \ldots, I$. To specify the decision maker’s objective, assume that penalty is incurred for holding a backlog of class $i$ orders at the rate of $c_i(Q_i(t))$ where $c_i(\cdot)$ is a known function that is continuous and non-decreasing, so the total backlog penalty is incurred at the rate of $\sum_i c_i(Q_i(t))$. Adopting the long-run average cost criterion, the decision maker seeks a policy $(T, \Psi)$ to minimize
\[
\limsup_{t \to \infty} \frac{1}{t} \mathbb{E}_P^P \left[ \sum_{i=1}^I \int_0^t c_i(Q_i(u)) du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) \right],
\]
where $\mathbb{E}_P^P$ denotes the expectation taken respect to the measure $P$. The primary objective of this paper is to formulate and solve a robust version of (1).

4. Ambiguity Aversion Model
This section introduces the notion of Rényi divergence and shows that it can be used to develop a robust control framework. The material presented in this section also serves to motivate the more general framework to be laid out in §7.

4.1. Some Background and Notation
In general, a probability measure $\tilde{P}$ is said to be absolutely continuous with respect to another measure $P$ if $\tilde{P}(A) = 0$ whenever $P(A) = 0$. Each $P$-absolutely continuous measure $\tilde{P}$ can be identified by the Radon-Nikodym derivative, denoted as $d\tilde{P} / dP$. Conversely, a nonnegative $\mathcal{F}$-measurable random variable $\psi$ with $\mathbb{E}_P^P \psi = 1$ induces a $P$-absolutely continuous measure $\tilde{P}$ on $\mathcal{F}$:
\[
\tilde{P}(A) := \mathbb{E}_P^P [\psi 1_A] \quad \text{for all} \quad A \in \mathcal{F}.
\]

Rényi divergence is referred to as an indexed family of statistical distances. Precisely, Rényi divergence of positive order $\alpha \neq 1$ of a measure $\tilde{P}$ from another measure $P$ is defined as
\[
\mathcal{R}_\alpha(\tilde{P} \parallel P) := \frac{1}{\alpha - 1} \ln \int \left( \frac{d\tilde{P}}{dP} \right)^{\alpha - 1} d\tilde{P}.
\]
As shown by Van Erven and Harremos (2014), given $\tilde{P}$ and $P$ fixed, $\mathcal{R}_\alpha$ is continuous in $\alpha$ on $(0, 1) \cup (1, \infty)$, and it tends to the well-known Kullback-Leibler (KL) divergence as $\alpha$ approaches 1. This leads to defining Rényi divergence of order $\alpha = 1$ as the KL divergence, so one could let $\alpha \in (0, \infty)$. It has also been shown that $\mathcal{R}_\alpha(\tilde{P} \parallel P)$ is convex in $\tilde{P}$ when $\alpha \in (0, 1]$ and quasi-convex in $\tilde{P}$ when $\alpha \in (1, \infty)$ (Van Erven and Harremos 2014).
4.2. Robust Control Problem

In reality, the demand rate for product $i$ may not be $\bar{\lambda}_i$ but rather some function of time $\lambda_i := \{\lambda_i(t); t \geq 0\}$. This leads to seeking control policies that are robust against uncertainty concerning the accuracy of the nominal model $\mathbb{P}$. Let $q$ denote nature’s strategy and $Q_i$ be the marginal law governing the arrival process $A_i$ realized under the strategy $q$. Also, denote by $P_i$ the marginal law governing $A_i$ under the nominal measure $\mathbb{P}$. Through the remainder of this paper we shall assume that the real-world demand rate vector $\lambda(t) := (\lambda_i(t))$, albeit not precisely known, must be chosen in such a way that each $Q_i$ is locally absolutely continuous with respect to its nominal counterpart $P_i$, so that $Q_i$ is absolutely continuously with respect to $P_i$ on $\mathcal{F}_t$ for any finite $t \geq 0$. Restricting attention to such perturbations of $P_i$ is reasonable in that it is tantamount to considering perturbations $Q_i$ that are difficult to discern statistically from $P_i$ over any finite time horizon, an idea exploited by Hansen et al. (2006) as well. To proceed, we introduce $\theta_i(t) := (\lambda_i(t) - \bar{\lambda}_i)/\bar{\lambda}_i$ to denote the relative deviation of $\lambda_i(t)$ from its “mean” $\bar{\lambda}_i$. We shall further stipulate that $Q_i$ is such that $\theta_i(t) \in \Theta_i := [a_i, b_i]$ for $-1 < a_i < 0 < b_i < \infty$. This effectively puts a positive lower bound and a finite upper bound for the process $\lambda_i$.

We assume that nature is adversarial in choosing the real-world model; that is, nature chooses $q$ with the goal of inflating the decision maker’s cost to the greatest extent possible. To put restrictions on nature’s actions, we inflict (for each $i$) a penalty on nature that reflects the extent of deviation of $Q_i$ from $P_i$ measured via Rényi divergence; specifically, the penalty received over $[0, t]$ is proportional to the Rényi divergence of order $\alpha$ of $Q_i$ from $P_i$ on $\mathcal{F}_t$, denoted as $R^\alpha_i(t)$. This leads to the following robust control problem: The decision maker seeks an adaptive strategy $(T, \Psi)$ to minimize

$$\max_q \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^q \left[ \sum_{i=1}^I \int_0^t c_i(Q_i(u))du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) - \sum_{i=1}^I \gamma_i R^\alpha_i(t) \right],$$

where each constant $\gamma_i$ quantifies the decision maker’s confidence level for the marginal law $P_i$. To briefly explain, a large value of $\gamma_i$ puts a large penalty on nature for deviations from $P_i$. In contrast, a small value of $\gamma_i$ entails a small penalty on nature for deviations of $Q_i$ from its nominal counterpart $P_i$, meaning that the decision maker feels less confident about $P_i$. As a general rule of thumb, the set
of parameters $\gamma := (\gamma_i)$ ought to be chosen in such a way that the resulting uncertainty set includes those that are statistically indistinguishable from the nominal model; alternatively one can think of $\gamma$ as a set of “tuning parameters” that the decision maker can add robustness to his/her control policy by solving problem (2) under some choice of $\gamma$.

### 4.3. Alternative Representation

To express objective (2) in a more explicit way, let $\Delta A_i(t) := A_i(t) - A_i(t-)$ and define the Doléans-Dade exponential

$$
\psi_i(t) := \exp \left\{ - \int_0^t \bar{\lambda}_i \theta_i(u) du \right\} \prod_{0 < u \leq t} (1 + \theta_i(u))^{\Delta A_i(u)},
$$

(3)

where we adopt the convention that zero to the zero power is 1. It is straightforward to check that $\psi_i := \{ \psi_i(t); t \geq 0 \}$ is a $\mathbb{P}_i$-local martingale; $\psi_i$ is a martingale if $E^{\mathbb{P}_i}[\psi_i(t)] = 1$ for $t \geq 0$. Our next result, which is essentially a version of the generalized Girsanov Theorem, describes the relationship between the change of measure from $\mathbb{P}_i$ to $\mathbb{Q}_i$ and the change of intensity for the process $A_i(t)$.

**Theorem 1.** Suppose that $\psi_i$ given by (3) is a martingale. Then the probability measure $\mathbb{Q}_i$ defined via

$$
\mathbb{Q}_i(A) := E^{\mathbb{P}_i}[\psi_i(t)1_A] \quad \text{for all} \quad A \in \mathcal{F}_t
$$

(4)

is locally absolutely continuous with respect to $\mathbb{P}_i$; moreover, $A_i(t)$ is a non-homogenous Poisson process with intensity $\lambda_i(t) = \bar{\lambda}_i(1 + \theta_i(t))$ under $\mathbb{Q}_i$. Conversely, a $\mathbb{P}_i$-locally absolutely continuous measure $\mathbb{Q}_i$ can be portrayed via (4) where $\psi_i$ is in the form of (3) with $E^{\mathbb{P}_i}[\psi_i(t)] = 1$ for $t \geq 0$.

Theorem 1 implies that a perturbation $\mathbb{Q}_i$ can be identified by an $\mathcal{F}_t$-predictable process $\theta_i$. More precisely, the change of measure from $\mathbb{P}_i$ to $\mathbb{Q}_i$ is equivalent to changing the intensity of $A_i(t)$ from $\bar{\lambda}_i$ to $\lambda_i(t) := \bar{\lambda}_i(1 + \theta_i(t))$. As a result, ambiguity towards the real-world model is equivalent to ambiguity in the $I$-dimensional process $\theta := (\theta_i)$. It is easy to see that if the $i$th component of $\theta$ is a zero process, then $\psi_i(t) = 1$ and $\lambda_i(t) = \bar{\lambda}_i$. The following result relates the Rényi divergence of $\mathbb{Q}_i$ from $\mathbb{P}_i$ to the process $\theta_i$. 
Proposition 1. Recall that \( R_\alpha^i(t) \) denotes the Rényi divergence of order \( \alpha \) of the measure \( Q_i \) induced by \( \theta_i \) from \( P_i \) on \( F_t \). Then

\[
R_\alpha^i(t) = \begin{cases} 
\frac{\tilde{\lambda}_i}{\alpha-1} \int_0^t \{(1 + \theta_i(u))^{\alpha} - \alpha \theta_i(u) - 1\} du & \text{for } \alpha \neq 1, \\
\bar{\lambda}_i \int_0^t \{(1 + \theta_i(u)) \ln(1 + \theta_i(u)) - \theta_i(u)\} du & \text{for } \alpha = 1.
\end{cases}
\]  

(5)

Upon substituting (5) into (2) and letting

\[
r(\theta) := r^\alpha(\theta) := \begin{cases} 
\sum_{i=1}^I \frac{\tilde{\lambda}_i}{\alpha-1} \{(1 + \theta_i)^\alpha - \alpha \theta_i - 1\} & \text{for } \alpha \neq 1, \\
\sum_{i=1}^I \gamma_i \bar{\lambda}_i \{(1 + \theta_i) \ln(1 + \theta_i) - \theta_i\} & \text{for } \alpha = 1,
\end{cases}
\]  

(6)

we obtain the following alternative representation of the problem (2):

\[
\max_{\theta} \lim_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta \left[ \sum_{i=1}^I \int_0^t c_i(Q_i(u)) du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) - \int_0^t r(\theta(u)) du \right];
\]  

(7)

in the above we deliberately replaced \( q \) and \( \mathbb{E}^q \) with \( \theta \) and \( \mathbb{E}^\theta \), respectively, to reflect the fact that the real-world model is induced by the process \( \theta \).

5. Heavy-Traffic Analysis

The problem (7) remains complicated and seems not amenable to exact analysis. Thus, in what follows we advance and solve an approximating SDG whose solution is analytically more tractable.

5.1. SDG

The heavy-traffic regime we focus on is the one where both the demand rates are large and the capacity balances the supply and demand. Specifically, we impose the usual heavy-traffic assumption:

\[
\sum_{i=1}^I \rho_i = 1 \quad \text{for } \rho_i := \bar{\lambda}_i m_i, \quad i = 1, \ldots, I.
\]  

(8)

Since the long-run proportion of time spent by the server on producing product \( i \) is \( \rho_i \), the system can be thought of as being critically loaded under \( \mathbb{P} \). Recall that \( \theta_i \) denotes relative deviation of \( \lambda_i \) from \( \bar{\lambda}_i \). Assume optimistically the deviation \( \lambda_i \theta_i \) to be second order.\(^1\) Then, we can approximate \( A_i \) using

\[
A_i(t) = \bar{\lambda}_i t + \bar{\lambda}_i \int_0^t \theta_i(u) du + \hat{A}_i(t) + \epsilon_i(t)
\]

\(^1\) This can be formalized by positing \( \theta_i(\cdot) = \bar{\theta}_i(\cdot)/\sqrt{n} \) for \( n := \sum_i \bar{\lambda}_i \) and some \( \bar{\theta}_i \) that is independent of \( n \).
under the measure $\mathbb{P}^{\theta}$, where $\hat{A}_i$ is a Brownian motion with zero drift and variance parameter $\hat{\lambda}_i$ and $\epsilon_i$ is an approximation error term. Also, define the centered time allocation processes as

$$Y_i(t) := \rho_i t - T_i(t) \quad \text{for} \quad i = 1, \ldots, I.$$ 

Following the approach pioneered by Harrison (1988), we can formally derive the approximating SDG. In particular, we replace the random elements $Q_i, Y_i,$ and $\xi_i$ with their respective approximations, namely, $\hat{Q}_i, \hat{Y}_i$ and $\hat{\xi}_i$, that collectively satisfy

$$\hat{Q}_i(t) = \hat{Q}_i(0) + \hat{Z}_i(t) + \int_0^t \hat{\lambda}_i \hat{\theta}_i(u) du + \mu_i \hat{Y}_i(t) - \sum_{k=0}^{N_i(t)} \hat{\xi}_i(k), \quad i = 1, \ldots, I,$$

$$U(t) := \sum_i \hat{Y}_i(t) \text{ is non-decreasing with } U(0) = 0,$$

$$\hat{Q}_i(t) \geq 0 \quad \text{for} \quad t \geq 0, \quad i = 1, \ldots, I,$$

where $\hat{Z}_i$ are independent Brownian motions with drift zero and infinitesimal variance $\sigma_i^2 = \hat{\lambda}_i(1 + \nu_i^2)$. Denote by $\hat{\Psi}_i$ the approximating type $i$ outsourcing control, i.e.,

$$\hat{\Psi}_i := (\tau_i(0), \tau_i(1), \tau_i(2), \ldots, \tau_i(m), \ldots; \hat{\xi}_i(0), \hat{\xi}_i(1), \hat{\xi}_i(2), \ldots, \hat{\xi}_i(m), \ldots).$$

By writing $\hat{Y} := (\hat{Y}_i)$ and $\hat{\Psi} := (\hat{\Psi}_i)$, we can formally state the decision maker’s problem as follows: The decision maker seeks an adapted control $(\hat{Y}, \hat{\Psi})$ that minimizes

$$\max_{\theta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta [\int_0^t \left( \sum_{i=1}^I c_i(\hat{Q}_i(u)) - r(\theta(u)) \right) du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\hat{\xi}_i(k))]$$

subject to constraints (9) – (11).

**5.2. Dimensional Reduction**

Although the SDG is simpler than the original problem it approximates, its solution is not so straightforward due to the state process $\hat{Q} := (\hat{Q}_i)$ being high dimensional. For this reason, we seek further simplification of the problem, which ultimately leads to a one-dimensional differential game, referred to as the workload problem. To start, define the one-dimensional workload process $W$ as follows:

$$W(t) := \sum_{i=1}^I m_i \hat{Q}_i(t), \quad t \geq 0,$$
which tracks the (scaled) amount of work in the system at time $t$. To deduce the system equation of
the workload, we multiply (9) by $m_i$ and sum over $i = 1, \ldots, I$ to get

$$W(t) = W(0) + B(t) + \int_0^t \zeta(u)du + U(t) - O(t),$$

where we defined

$$B(t) := \sum_{i=1}^I m_i \tilde{Z}_i(t), \quad \zeta(t) := \sum_{i=1}^I \rho_i \theta_i(t) \quad \text{and} \quad O(t) := \sum_{i=1}^I m_i \sum_{k=0}^{N_i(t)} \tilde{\xi}_i(k).$$

In the above equation, $B := \{B(t); t \geq 0\}$ is a zero-drift Brownian motion with infinitesimal variance

$$\sigma^2 = \sum_i m_i^2 \sigma_i^2,$$

$\zeta := \{\zeta(t); t \geq 0\}$ is the drift rate process subject to the control by nature, and $U(t)$
can be interpreted as the cumulative (scaled) idle time up to time $t$. In a like fashion, we interpret
$O := \{O(t); t \geq 0\}$ as the cumulative (scaled) amount of work outsourced up to time $t$.

Now, define the effective holding cost rate function by

$$h(w) = \min \left\{ \sum_{i=1}^I c_i(x_i) : m^\top x = w, x \in \mathbb{R}_+^I \right\}. \quad (12)$$

The cost rate function $h(w)$ has the following interpretation: Given the total workload $w$ can be
quickly redistributed across all classes in any way the decision maker desires, the amount of work
will be distributed in such a way that the aggregate holding cost rate is minimized. Similarly, the
effective cost rate function for nature is given as

$$r^*(z) = \min \{ r(y) : \rho^\top y = z, y_i \in \Theta_i \}. \quad (13)$$

Associated with $I$ different types of outsourcing operations, there are $I$ different outsourcing cost
functions, corresponding to $I$ different ways to push down the workload to a desired level. Specifically,
for type $i$ outsourcing operations, we can define

$$\tilde{\phi}_i(w) := (L_i + \tilde{\ell}_i w) \cdot 1_{\{w > 0\}} + 0 \cdot 1_{\{w=0\}} \quad \text{for} \quad \tilde{\ell}_i := \ell_i/m_i.$$

We interpret the cost parameter $\tilde{\ell}_i$ as the proportional cost of outsourcing one unit of work through
type $i$ outsourcing operations and shall use $\tilde{\Psi}$ to denote the outsourcing control for the workload.
Finally, we let $\tilde{\xi}_i(k) := m_i \hat{\xi}_i(k)$ for $k \geq 0$ and $i = 1, \ldots, I$. With these preparations, we can spell out the workload problem, stated as follows: The decision maker seeks an adaptive control $(U, \tilde{\Psi})$ to minimize

$$\max_{\zeta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^c \left[ \int_0^t h(W(u))du - \int_0^t r^*(\zeta(u))du + \sum_{i=1}^I \sum_{k=0}^N \tilde{\phi}_i(\tilde{\xi}_i(k)) \right]$$

subject to

$$W(t) = W(0) + B(t) + \int_0^t \zeta(u)du + U(t) - \sum_{i=1}^I \sum_{k=0}^N \tilde{\xi}_i(k),$$

(14)

where $U(t)$ is non-decreasing with $U(0) = 0$, (15)

$$W(t) \geq 0 \quad \text{for} \quad t \geq 0,$$

(16)

where $\mathbb{E}^c$ is expectation with respect to the measure $\mathbb{P}^c$ induced by effective drift rate process $\zeta$.

5.3. Characterization of the Optimal Policy

Because both players face a long-run average objective, we restrict our attention to the class of stationary Markov strategies. From nature’s perspective, this means that the drift-rate control $\zeta(t)$ at time $t$ depends on history only through the workload $W(t)$ at time $t$. Hence, from now onward we shall write $\zeta(W(t))$ in replacement of $\zeta(t)$. From the decision maker’s point of view, this suggests that the desired outsourcing policy would take a control limit form (which we shall briefly describe below). In the meantime, it is evident that a deviation from the work-conserving principle can only hurt the decision maker; to wit, the idleness process $U$ ought to satisfy

$$\int_0^t 1_{\{W(u) > 0\}} dU(u) = 0, \quad t \geq 0.$$

5.3.1. Control-Band Policy Following Ormeci et al. (2008), we define a relevant class of control policies as follows.

**Definition 1.** Given some $i \in \{1, \ldots, I\}$ and two parameters $q, s$ with $0 < q < s$, we call $(i, q, s)$ a control band policy of type $i$ with parameters $(q, s)$, if the decision maker utilizes type $i$ outsourcing operations only, and upon $W$ reaching the upper barrier $s$, the decision maker enforces a downward jump to level $q$, thereby incurring a cost of $\tilde{\phi}_i(s - q)$. 
Now, for an arbitrarily given real-valued function $\zeta(\cdot)$, define the differential operator $\Gamma_\zeta$ as

$$\Gamma_\zeta f(w) = \frac{1}{2}\sigma^2 f''(w) + \zeta(w)f'(w).$$

Now, for a fixed $s > 0$ let $\mathcal{C}^2[0, s]$ denote the space of functions that are twice differentiable up to the boundaries. Suppose that there exists some $\eta \in \mathbb{R}$ and $f \in \mathcal{C}^2[0, s]$ that collectively satisfy

$$\Gamma_\zeta f(w) + h(w) - r^*(\zeta(w)) = \eta \quad \text{for} \quad w \in (0, s) \tag{17}$$

subject to the boundary conditions

$$f'(0) = 0 \quad \text{and} \quad f(s) = \tilde{\phi}_i(s - q) + f(q). \tag{18}$$

The following proposition provides a useful identity that serves to motivate the optimality equation to be described in the next subsection.

**Proposition 2.** Suppose $\eta \in \mathbb{R}$ and $f \in \mathcal{C}^2[0, s]$ jointly satisfy (17) and (18). Then $\eta$ is the long-run average cost under the control band policy $(i, q, s)$ and the drift-rate control $\zeta(\cdot)$.

**5.3.2. Optimality Equation** We proceed in three major steps. First, using the boundary and smooth pasting conditions while taking nature’s strategic behavior into account, we identify a specific control band policy, denoted as $(i, q_i, s_i)$, that mini-maximizes the long-run average cost within the class of controls utilizing type $i$ outsourcing operations only; we denote by $\eta_i$ the resulting long-run average cost of this policy. Second, among the $I$ “best” policies, namely, $\{(i, q, s); i = 1, \ldots, I\}$, we define the candidate solution to the decision maker’s decision problem as the one yielding the lowest long-run average cost under the specified minimax criterion; that is, choose $i^*$ such that $\eta_{i^*} \leq \eta_i$ for all $i \neq i^*$; the policy $(i^*, q_{i^*}, s_{i^*})$ is then viewed as the candidate solution solving the decision maker’s problem. Third, by exploiting the structural properties of the value function associated with policy $(i^*, q_{i^*}, s_{i^*})$, we demonstrate that this strategy is indeed average cost optimal for the decision maker under the minimax criterion among all adaptive controls that the decision maker can take.

Proposition 2 motivates the following optimality equation that facilitates the identification of the policy $(i, q_i, s_i)$ as mentioned earlier: Find $q_i, s_i, \eta_i \in \mathbb{R}$ and $v \in \mathcal{C}^2[0, s_i]$ such that

$$\max_{\zeta} \left\{ \frac{1}{2}\sigma^2 v''(w) + \zeta v'(w) + h(w) - r^*(\zeta) \right\} = \eta_i, \quad w \in (0, s_i), \tag{19}$$
subject to the boundary conditions

\[ v'(0) = 0 \quad \text{and} \quad v(w) = \tilde{\phi}_i(w - q_i) + v(q_i) \quad \text{for all} \quad w \geq s_i, \]  

(20)

and necessary optimality conditions stemmed from the “principle of smooth fit”: \( v'(q_i) = v'(s_i) = \tilde{\ell}_i. \)

Letting \( r^*(x) \) be the convex conjugate of the function \( r^*(\zeta), \) i.e.,

\[ r^*(x) := \max_{\zeta} \{ x\zeta - r^*(\zeta) \} \quad \text{for} \quad x \in \mathbb{R}, \]

we can rewrite equation (19) as

\[ \frac{1}{2} \sigma^2 v''(w) + r^*(v'(w)) + h(w) = \eta_i, \quad w \in (0, s_i). \]

(21)

Because (21) does not involve the unknown function \( v \) itself, it is in essence a first-order differential equation. Motivated by this we consider the class of functions \( \{ \pi(\cdot, \eta); \eta \in \mathbb{R} \}, \) where \( \pi(\cdot, \eta) \) solves

\[ \frac{1}{2} \sigma^2 \pi_w(w, \eta) + r^*(\pi(w, \eta)) + h(w) - \eta = 0 \]

(22)

subject to the boundary condition

\[ \pi(0, \eta) = 0. \]

(23)

The policy parameters \((q_i, s_i)\) along with the average cost \( \eta_i \) can be determined through conditions:

\[ \pi(q_i, \eta_i) = \pi(s_i, \eta_i) = \tilde{\ell}_i \quad \text{and} \quad \int_{q_i}^{s_i} \pi(w, \eta_i)dw = \tilde{\phi}_i(s_i - q_i). \]

(24)

It is worth emphasizing that (23) and (24) effectively make up four constraints, which alongside (22) can be utilized to solve for four unknowns, namely \( q_i, s_i, \eta_i \) and \( \pi(\cdot, \eta_i) \).

Thus far, two questions remain unanswered. First, does the system of equations given in (22)–(24) yield a solution? Second, given the answer to the first question is “yes”, does the resulting policy \((i, q_i, s_i)\) yield the lowest cost possible if the decision maker were to choose to outsource the manufacturing needs of product \( i \) only? Towards answering the two questions, we introduce additional regularity conditions on the problem data.
Assumption 1. The cost rate function $h(w)$ is continuous and strictly increasing on $(0, \infty)$ with $h(0) = 0$; moreover, $\lim\inf_{w \to \infty} h(w)/w = c_0$ for some $c_0 \in (0, \infty]$.

Our next result not only gives a positive answer to the first question but provides key ingredients in finding answers to the second one.

Proposition 3. Suppose that Assumption 1 holds. Then the following statements are true: (i) The requirements in (22)–(24) yield unique $(q_i, s_i)$ and $\eta_i$. (ii) If letting $v$ denote the unique primitive function of $\pi(\cdot, \eta_i)$, modulo an additive constant, the pair $(v, \eta_i)$ satisfies the following quasi-variational inequality:

$$\min \left\{ \frac{1}{2} \sigma^2 v''(w) + r^*(v'(w)) + h(w) - \eta_i, \inf_{z \geq 0} \{ v(w - z) + \tilde{\phi}_i(z) - v(w) \} \right\} = 0.$$

(iii) Independent of the initial condition, the control band policy $(i, q_i, s_i)$ mini-maximizes the long-run average cost among the class of adaptive controls utilizing type $i$ outsourcing operations only.

Continuing the discussion of control band policies, we now seek a policy perceived as one that mini-maximizes the long-run average cost among all adaptive controls that the decision maker can possibly take. To that end, we propose the following algorithm to generate a specific control band policy that can be viewed as the “best-of-the-best” control band policy:

1. For each $i = 1, \ldots, I$, find the triple $(q_i, s_i, \eta_i)$ using (22)–(24).
2. Seek an index $i^* \in \{1, \ldots, I\}$ such that $\eta_{i^*} \leq \eta_i$ for all $i \neq i^*$.
3. Propose the control band policy $(i^*, s_{i^*}, \eta_{i^*})$ as the best strategy for the decision maker to take.

Theorem 2. Suppose that Assumption 1 holds. (i) Let $v$ denote the unique primitive function of $\pi(\cdot, \eta_{i^*})$, modulo an additive constant. Then the pair $(v, \eta_{i^*})$ satisfies the following optimality equation:

$$\min \left\{ \frac{1}{2} \sigma^2 v''(w) + r^*(v'(w)) + h(w) - \eta_{i^*}, \min_{z \geq 0} \left\{ v(w - z) + \tilde{\phi}_i(z) - v(w) \right\} \right\} = 0.$$

(ii) Independent of the initial condition, the control band policy $(i^*, q_{i^*}, s_{i^*})$ is average cost optimal under the minimax criterion among the class of adaptive controls.
5.4. Policy Recommendation

In this subsection, we put forth an implementable set of control strategies for the decision problem introduced in §4 based on the analytical results derived earlier. Here, an implicit assumption is that both $\alpha$ and $\gamma$ are given so that the uncertainty set is fixed. As a result, the decision maker can compute $(i^*, s_{i^*}, \eta_{i^*})$ along with $\eta_{i^*}$ and $\pi(\cdot, \eta_{i^*})$ using the approach described in §5.3.2. A guideline on how to choose $\alpha$ and $\gamma$ based on historical data will be the focus of the next subsection.

Previously, we have used $W$ to denote the approximate workload process. To avoid introducing new notation, we “overload” the use of $W(t)$ to now be the workload of the actual system at time $t$, defined by $W(t) := \sum_i m_i Q_i(t)$, where $Q_i(t)$ is the actual number of class $i$ orders awaiting processing at time $t$ and $m_i$ denotes the mean processing time for product $i$ in the nominal model. The proposed policy has two components, as described below.

**Outsourcing policy.** Whenever the workload $W$ reaches the upper barrier $s_{i^*}$, outsource $o_{i^*} := (s_{i^*} - q_{i^*})/m_{i^*}$ orders of product $i^*$ immediately, if there are $o_{i^*}$ orders of product $i^*$ awaiting processing. If the number of outstanding orders of class $i^*$, say $Q_{i^*}$, is less than $o_{i^*}$, then postpone the outsourcing operation until additional $(o_{i^*} - Q_{i^*})$ orders of product $i^*$ arrive.

**Sequencing policy.** To spell out the sequencing strategy, we postulate a unique solution $(x_{i^*})$ to the optimization problem (12) for any given $w$. For this reason, the solution $(x_{i^*})$ can be regarded as a function of $w$ and serves as the target length of the queues when the workload is at position $w$. (We wish to point out that the above stipulation merely seeks to mitigate the potential technical complexity and will be easily satisfied in various settings.) Hence, a desired sequencing policy ought to be one that tries to maintain the actual queue lengths at their respective targets $(x_{i^*})$. When the waiting cost rates are linear, i.e., $c_i(x) = C_i x$ for some constant $C_i > 0$, we recover the celebrated $c\mu$ priority rule. When all $c_i(\cdot)$ are strictly convex and satisfy $c_i(0) = c'_i(0) = 0$, it leads to the generalized $c\mu$ rule; that is, give service priority to the job class for which the index $c'_i(Q_i(t))\mu_i$ is the largest at time $t$. 
5.5. Selecting the Uncertainty Set

Choosing uncertainty sets is an important question, because it affects whether one can efficiently represent a robust control problem. A too small uncertainty set may fail to include the true model, thereby making a robust formulation lose its appeal. On the other hand, choosing a too large uncertainty set can yield robust solutions that are so conservative that the resulting solutions are of low quality for the objective, thereby losing essentially any advantage over the non-robust counterpart.

Suppose the decision maker has access to historical data suggesting that the demands for various products are best described by some model $\mathcal{M}$; e.g., demand for each product may be found to follow an ARIMA model (Shumway et al. 2000). Recall that with Rényi-type penalties, the “shape” and size of the uncertainty set are collectively determined by $\alpha$ and $\gamma$. A good uncertainty set (hence a good policy) can then be found in the following fashion.

- Propose a collection of candidate shape parameters $\{\alpha^{(1)}, \cdots, \alpha^{(A_1)}\}$ and a set of candidate size parameters $\{\gamma^{(1)}, \cdots, \gamma^{(A_2)}\}$ for $\gamma$. For each combination $(\alpha^{(k)}, \gamma^{(l)})$, solve the Bellman-Isaacs equation to find out the corresponding optimal policy. Repeating the procedure $A := A_1 \times A_2$ times gives us $A$ different robust control policies, one for each combination $(\alpha^{(k)}, \gamma^{(l)})$.

- Construct a simulation model in which demand arrivals are generated using the “true” model $\mathcal{M}$; e.g., if demand data fits into a multivariate ARIMA model, then use that model to generate “demands” in the simulation program.

- Evaluate each of the $A$ different policies found in the first step by implementing it in the simulation program and estimating the long-run average cost based on independent runs. Intuitively, a robust policy yields a low objective value when the true model is on or near the “periphery” of the uncertainty set. This motivates selecting the policy that yields the lowest cost among the $A$ testing cases; the corresponding parameter combination $(\alpha^{(k)}, \gamma^{(l)})$ can then be thought of as one that characterizes a good uncertainty set (as far as Rényi-type penalties are concerned).

The data-driven method described above is believed to strike a balance between tractability and practicability in the following sense. On the one hand, an analytical model often requires the model
primitives to be “simple”; (even slightly) more sophisticated models may render the resulting decision problem intractable. However, false or oversimplified assumptions may unknowingly cause the resulting decision action to deviate significantly from optimal ones, thereby hurting profitability of the manufacturing firm. On the other hand, despite being distribution-free, simulation-based optimization is typically computationally intensive and expensive. Provided adequate time and computational power, it is believable that a simulation-based search may ultimately locate the “best” policy, but sufficient computational resources may be difficult to access for a regular manufacturing firm. In addition, without analytical theories it is not even clear whether the search should focus on control band policies; furthermore, there is a lack of a mechanism with which the appropriateness of the simplifying model (i.e., the nominal model) can be evaluated. By contrast, the proposed method combines the benefits of both approaches in that it exploits both analytical tractability (enabled by a simplifying model) and historical data through a purposeful search in the policy space via simulations.

6. A Numerical Study

In this section, we present a numerical study to demonstrate the value of the proposed policy. In §6.1 we discuss how we execute the simulation and perform the statistical estimation. In §6.2 we set up a numerical example reflecting a real-world make-to-order system. In §6.3 we report and discuss the numerical findings. In §EC.2 we clarify how we numerically solve the optimality equation in detail.

6.1. Simulation Execution & Statistical Estimation

The first step of the numerical program is to calculate $\sigma^2$ for the workload process. Recall that we can achieve this by utilizing model inputs including $\bar{\lambda}$, $\mu$, and $\nu$. Then, by solving the workload problem, we are able to uniquely pin down the optimal control band parameters $(q_i, s_i)$ given the manufacturer’s only outsourced product $i$, which yields the “optimal” long-run average cost $\eta_i$. By comparing each cost $\eta_i$, we could also determine where (at which queue) impulse control ought to take place, and furthermore, determine the optimal control parameters $q_i^\star$, $s_i^\star$.

Our simulations use $r = 100$ i.i.d. replications of our queuing system observed over a time interval of length of 2,000 after a warm-up period of length of 100 to allow the system that started empty
to approach the steady state. For a random variable $X$, its mean $\mathbb{E}[X]$ is estimated by the sample averages of the $r$ values, which should be Gaussian. Hence, the 95% confidence interval can be constructed in the usual way with $t_{0.025}(r-1)$.

6.2. Setup

Consider a make-to-order system producing two types of products. The time to produce a product of class 1 (resp. class 2) is assumed to be exponentially distributed with a rate $\mu_1 = 30$ (resp. rate $\mu_2 = 90$); hence $m_1 = 1/30$, $m_2 = 1/90$ and $\nu_1 = \nu_2 = 1$. The demand for each product is assumed to follow a non-homogeneous Poisson process whose intensity is a Cox–Ingersoll–Ross (CIR) process. Specifically, $\lambda_i(t)$ is the solution to the following stochastic differential equation

$$d\lambda_i(t) = k_i(\bar{\lambda}_i - \lambda_i(t))dt + \tilde{\sigma}_i \sqrt{\lambda_i(t)}d\tilde{B}_i(t), i = 1, 2,$$

(25)

where $\tilde{B}_i(t)$ is standard Brownian motion that is independent of everything else. In the above, parameter $k_i$ corresponds to the speed of adjustment to the mean $\bar{\lambda}_i$, and $\tilde{\sigma}_i$ measures the model volatility. In the present example, we choose $\bar{\lambda}_1 = 20$, $\bar{\lambda}_2 = 30$, $k_1 = k_2 = 1$ and $\tilde{\sigma}_1 = \tilde{\sigma}_2 = 2$, and we can generate sample paths of (25) via the Milstein scheme. It is straightforward to verify that $\bar{\lambda}_1/\mu_1 + \bar{\lambda}_2/\mu_2 = 1$; hence the desired heavy-traffic condition is satisfied. Also, a direct calculation gives $\sigma^2 = 0.0519$.

The next step is to compute the holding cost function and penalty function for workload. The cost data includes the fixed outsourcing cost parameters, $L_1 = 8$, $L_2 = 6$, the proportional outsourcing cost parameters, $l_1 = 1.5$, $l_2 = 1.8$, and two quadratic holding cost rate functions

$$c_i(x_i) = a_i x_i^2, \quad \text{for} \quad i = 1, 2,$$

(26)

where we choose $a_1 = a_2 = 0.1$. Note that with quadratic holding costs as in (26), we can solve the minimization problem embedded in (12) to get

$$h(w) = a_1 \left( \frac{a_2 m_1}{a_1 m_2^2 + a_2 m_1^2} \right)^2 w^2 + a_2 \left( \frac{a_1 m_2}{a_1 m_2^2 + a_2 m_1^2} \right)^2 w^2.$$

(27)

Moreover, the quadratic holding costs give rise to a queue-ratio type rule: If $a_1 \mu_1 Q_1 > (<) a_2 \mu_2 Q_2$, give priority to producing product 1 (product 2) with the tie broken in an arbitrary fashion.
It is assumed that the decision maker uses Rényi divergence to characterize his/her ambiguity in demands. We need to solve an optimization problem for workload which aims to find an effective cost rate function (13) with Rényi type penalties (6). Our goal is to minimize

\[
\gamma_1 \bar{\lambda}_1 (1 + y_1) \ln(1 + y_1) + \gamma_2 \bar{\lambda}_2 (1 + y_2) \ln(1 + y_2) - \gamma_1 \bar{\lambda}_1 y_1 - \gamma_2 \bar{\lambda}_2 y_2
\]

subject to \( \rho_1 y_1 + \rho_2 y_2 = \zeta \).

Further, deriving (28) implies the objective function as

\[
\min \gamma_1 \bar{\lambda}_1 (1 + y_1) \ln(1 + y_1) + \gamma_2 \bar{\lambda}_2 (1 + \frac{\zeta - \rho_1 y_1}{\rho_2}) \ln(1 + \frac{\zeta - \rho_1 y_1}{\rho_2}) - \gamma_1 \bar{\lambda}_1 y_1 - \gamma_2 \bar{\lambda}_2 \frac{\zeta - \rho_1 y_1}{\rho_2}.
\]

First order condition gives that

\[
\gamma_1 \bar{\lambda}_1 \left( \ln(1 + y_1) + 1 \right) - \gamma_2 \bar{\lambda}_2 \frac{\rho_1}{\rho_2} \left( \ln(1 + \frac{\zeta - \rho_1 y_1}{\rho_2}) + 1 \right) - \gamma_1 \bar{\lambda}_1 + \gamma_2 \bar{\lambda}_2 \frac{\rho_1}{\rho_2}
\]

\[
= \gamma_1 \bar{\lambda}_1 \ln(1 + y_1) - \gamma_2 \bar{\lambda}_2 \frac{\rho_1}{\rho_2} \ln \left( 1 + \frac{\zeta - \rho_1 y_1}{\rho_2} \right) = 0.
\]

Solving the above equation yields

\[
\zeta = \rho_2 (1 + y_1) \left( \frac{\gamma_1 \bar{\lambda}_1 \rho_2}{\gamma_2 \bar{\lambda}_2 \rho_1} \right) + \rho_1 y_1 - \rho_2.
\]

From (29) we deduce that the expression \( \rho_2 (1 + y_1) \left( \frac{\gamma_1 \bar{\lambda}_1 \rho_2}{\gamma_2 \bar{\lambda}_2 \rho_1} \right) + \rho_1 y_1 \) is strictly increasing with \( y_1 \). Therefore, for a given \( \zeta \), equation (29) produces a unique root \( y_1 \). Then \( y_2 = (\zeta - \rho_1 y_1)/\rho_2 \) further shows \( y_2 \) is unique as well. Finally, we can run a point-wise optimization program to find the value of \( r^*(\zeta) \). And further, the algorithm for solving the optimality equation (22) to (24) can be found in the online appendix.

6.3. Results

With all the preparations given in the previous sections, we can run the numerical program. We test some cases with different pairs of \((\gamma_1, \gamma_2)\). We can clearly observe from table 1 that the asymptotic bound achieved in the numerical program is strictly decreasing in \( \gamma_1 \) when \( \gamma_2 \) is fixed, and strictly decreasing in \( \gamma_2 \) when \( \gamma_1 \) is fixed. Intuitively speaking, a larger value of \( \gamma_i \) entails a higher penalty for the malevolent agent to act on the \( i \)th product, hence little incentive for the agent to distort the nominal model. The impact on each \( \gamma_i \) differs in quantity. In this example, varying the value of \( \gamma_1 \)
makes $\eta_1$ and $\eta_2$ drop faster than varying the value of $\gamma_2$. In more detail, when both $\gamma_1$ and $\gamma_2$ are large enough, the cost $\eta_i$ will converge to their “limiting value”, in which case the malevolent agent chooses to do nothing, or equivalently, the decision maker has complete confidence over the demands in the nominal model. Details of optimal control band parameters $(q_i, s_i)$ and “optimal” long-run average cost $\eta_i$ can be found in table EC.1 in online appendix. Furthermore, for this particular example, we can see that no matter which value the pair $(\gamma_1, \gamma_2)$ takes, the asymptotic bound $\eta_2$ is always greater than $\eta_1$, which implies that one should adopt policy parameters $(q_1, s_1)$ for the workload problem: Whenever the workload exceeds $s_1$ we should enforce a downward jump of size $(s_1 - q_1)$ through outsourcing a number of class 1 orders as described in §5.4.

It can be seen from Figure 2 that the estimated average cost first decreases and then increases in $\gamma_1$ (resp. $\gamma_2$) with $\gamma_2$ (resp. $\gamma_1$) fixed. It is also worth pointing out that for very large $\gamma_1$ and $\gamma_2$ (very small $1/\gamma_1$ and $1/\gamma_2$), the resulting control policy will be the one obtained by completely ignoring model uncertainty. Hence, the difference between the “limiting value” of each plot and the minimum value on that plot can be regarded as the value of robustness. In the present example, the value of robustness is $9.0966 - 8.2743 = 0.8223$, which is approximately an improvement of 9%.

7. A General Framework

Thus far, we have demonstrated how a decision maker can incorporate model uncertainty and ambiguity aversion into decision making within a game-theoretical setting. In particular, through a change
of measure, we show that a set of Rényi-type penalties can be translated into a penalty rate function $r(\cdot)$ applied to the probability distortion process. In this respect, the decision maker’s ambiguity about the model is fully captured by the function $r(\cdot)$. Although the Rényi divergence gives rise to a versatile family of uncertainty sets, it is nonetheless restrictive for reasons that will shortly become clear. The key is to observe that the penalty imposed on nature does not have to be in the form as in (6). In principle, the decision maker can specify any penalty form as s/he desires, as long as the resulting uncertainty set is believed to capture concerns about model misspecification. From now on, let $p(\cdot)$ denote the selected penalty function that maps from $\mathbb{R}^l$ to $\mathbb{R}_+$; hence, when nature picks $\theta(t)$ at time $t$, it makes a penalty payment to the decision maker at the rate of $p(\theta(t))$. This motivates a general robust control formulation to be described below.

The robust control problem corresponding to the penalty function $p(\cdot)$ can be described as the following: The decision maker seeks an adapted strategy $(T, \Psi)$ to minimize

$$\max_{\theta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta \left[ \int_0^t \left( \sum_{i=1}^I c_i(Q_i(u)) - p(\theta(u)) \right) \, du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) \right],$$

(30)
where $\mathbb{E}^\theta$ denotes the expectation taken with respect to the measure $\mathbb{P}^\theta$ induced by $\theta$. Of course, upon replacing $p$ with $r$, one recovers the problem (7).

To illustrate the usefulness of this approach, consider a common scenario where the decision maker believes/knows that the demand rate for product $i$ is confined in a range, $[\bar{\lambda}_i - \kappa_i, \bar{\lambda}_i + \kappa_i]$, where $\bar{\lambda}_i$ is the long-term average, and accordingly, $\kappa_i$ can be interpreted as the degree of model uncertainty. Clearly, this simple scenario cannot be captured through the use of Rényi-type penalties. Nonetheless, it falls under the general framework specified by (30) if one lets $p(\theta) = \sum_i p_i(\theta_i)$ for

$$
p_i(x) = \begin{cases} 0 & \text{for } |x| \leq \kappa_i, \\
\infty & \text{for } |x| > \kappa_i.
\end{cases}
$$

To briefly explain, the maximizer is able to select $\theta_i(t)$ from the interval $[\bar{\lambda}_i - \kappa_i, \bar{\lambda}_i + \kappa_i]$ freely at no expense while never considering values outside the range, which corresponds to the decision maker’s belief that the demand rate for product $i$ is confined in the range $[\bar{\lambda}_i - \kappa_i, \bar{\lambda}_i + \kappa_i]$ with no outcome being more likely than others.

Replicating the development in §5.1, we can easily arrive at the SDG in which the decision maker chooses $(\hat{Y}, \hat{\Psi})$ to minimize

$$
\max_{\theta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta \left[ \int_0^t \left( \sum_{i=1}^I c_i(\hat{Q}_i(u)) - p(\theta(u)) \right) du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\hat{\xi}_i(k)) \right]
$$

subject to constraints (9) -- (11).

Next, paralleling (13) we define

$$
p^*(z) = \min \{ p(y) : \rho^\top y = z, y_i \in \Theta_i \}.
$$

It is nature to require that $p$ is chosen in such a way that $p^*$ achieves its minimum value at $z = 0$ and $p^*(0) = 0$. Arguing along the lines that are similar to those in §5.2, we obtain the workload problem:

Seek an adaptive control $(U, \tilde{\Psi})$ to minimize

$$
\max_{\zeta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\zeta \left[ \int_0^t h(W(u)) du - \int_0^t p^*(\zeta(u)) du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right]
$$

subject to constraints (14) -- (16).
Accordingly, let $g(x)$ be the convex conjugate of the function $p^*(\zeta)$, i.e.,

$$g(x) := \max_\zeta \{ x\zeta - p^*(\zeta) \} \quad \text{for} \quad x \in \mathbb{R},$$

where $p^*(\cdot)$ is defined by (32). Note that by definition $g(0) = \min_\zeta p^*(\zeta) = 0$. As before, we restrict attention to the class of control band policies, so the task of seeking the best policy among all control band policies which utilize type $i$ outsourcing operations only, boils down to finding $(q_i, s_i)$. Paralleling the development in §5.3.2, we can identify $(q_i, s_i)$ and the corresponding average cost $\eta_i$ by using (22)–(24) wherein we redefine the class of functions $\{\pi(\cdot, \eta); \eta \in \mathbb{R}\}$ in such a way that $\pi(\cdot, \eta)$ solves the following differential equation:

$$\frac{1}{2} \sigma^2 \pi_w(w, \eta) + g(\pi(w, \eta)) + h(w) - \eta = 0 \quad (33)$$

with the boundary condition $\pi(0, \eta) = 0$. In the sequel, the following conditions will be in force.

**Assumption 2.** The function $g$ is non-negative and Lipschitz continuous on $\mathbb{R}$.

This assumption provides a standard (sufficient) condition for equations like (33) to have a unique solution (see, e.g., chapter 3 in David et al. (2018)). Moreover, the assumption implies that the function value $g(x)$ for any fixed $x$.

**Proposition 4.** Suppose both Assumptions 1 and 2 hold. (i) The requirements given by (22)–(24) wherein $\pi$ is redefined through (33) yield unique policy parameters $(q_i, s_i)$ and average cost $\eta_i$. (ii) If letting $v$ denote the unique primitive function of $\pi(\cdot, \eta)$, modulo an additive constant, then the pair $(v, \eta)$ satisfies the following quasi-variational inequality:

$$\min \left\{ \frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta, \inf_{z \geq 0} \left[ v(w - z) + \tilde{\phi}_i(z) \right] - v(w) \right\} = 0. \quad (34)$$

(iii) Independent of the initial condition, the control band policy $(i, q_i, s_i)$ mini-maximizes the long-run average cost among the class of adaptive controls utilizing type $i$ outsourcing operations only.

Pursuing the same computational procedure as described at the end of §5.3.2, we can find a triple $(q_{i\star}, s_{i\star}, \eta_{i\star})$, corresponding to the “best-of-the-best” control band policy. As before, we denote such a policy by $(i\star, q_{i\star}, s_{i\star})$. 
Theorem 3. Suppose that both Assumptions 1 and 2 hold. (i) Denote by $v$ the unique primitive function of $\pi(\cdot, \eta^{\star})$, modulo an additive constant. Then the pair $(v, \eta^{\star})$ satisfies the following optimality equation:

$$\min \left\{ \frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta^{\star}, \min \inf_{z \geq 0} \left[ v(w - z) + \tilde{\phi}_i(z) \right] - v(w) \right\} = 0.$$  \hfill (35)

(ii) Independent of the initial condition, the control band policy $(i^{\star}, q^{\star}, s^{\star})$ is average cost optimal under the minimax criterion among the class of adaptive controls.

8. Concluding Remarks

This paper studies the joint order outsourcing and sequencing of a multiclass make-to-order manufacturing system with model uncertainty and facing fixed plus proportional costs for outsourcing. Model uncertainty is captured through the notion of Rényi divergence, thereby extending the commonly used entropic approach in the literature. (This approach is further extended to incorporate a broader category of uncertainty sets.) We present a robust control formulation that involves a second player that promotes robustness to model misspecification. This formulation can be interpreted as a two-player zero-sum stochastic game.

By considering the system in a suitably heavy-traffic regime, we derive and solve an approximate SDG whose state-descriptor is driven by Brownian motions. It is demonstrated that the SDG can be further turned into a one-dimensional stochastic game whose state-descriptor is the workload process. The optimal control strategy for the decision maker is shown to be in a control-band form. The optimal strategy for nature is a state-dependent drift-rate control of the workload process that serves to resolve model uncertainty. Interpreting the solution to the SDG in the context of the original problem, we propose a class of joint outsourcing and sequencing strategies. The value of ambiguity is demonstrated through a simulation study wherein we compared different robust control strategies obtained with different degrees of ambiguity aversion in mind.

Although the paper focuses on a make-to-order manufacturing system, the solution framework may be adapted to other production strategies, such as make-to-stock manufacturing systems whereby
inventories are allowed. Moreover, our approach to characterizing ambiguity using Rényi divergence and its further extensions as demonstrated in this paper may be well-suited for other examples of stochastic dynamic programming problems. These are interesting directions worth pursuing in future research.

References


Online Appendix

This e-companion is organized into four parts. Section EC.1 collects proofs for the main results in the paper. Section EC.2 describes our numerical scheme used for solving the Bellman-Issacs equation. Section EC.4 presents proofs for some auxiliary results.

EC.1. Proofs of the Main Results

This part of the e-companion gives proofs for Proposition 1, Proposition 4 and Theorem 3. Since Proposition 3 and Theorem 2 are special cases of Proposition 4 and Theorem 3, respectively, their proofs are omitted. The proof of Proposition 2 is elementary, so we omit it as well.

Proof of Proposition 1. Because the result for $\alpha = 1$ is known, we restrict attention to cases where $\alpha \neq 1$. To start, Let $\tilde{\alpha} := \alpha - 1$. Direct calculation gives

$$
\psi_i^\tilde{\alpha}(t) = \exp \left\{ \tilde{\alpha} \int_0^t \ln(1 + \theta_i(u))dA_i(u) - \tilde{\alpha} \int_0^t \bar{\lambda}_i \theta_i(u)du \right\}
$$

(EC.1)

Now, consider a partition $\{u_i\}$ of $[0, t]$, such that $0 = u_0 < u_1 < \cdots < u_m = t$. It follows that

$$
\exp \left\{ \tilde{\alpha} \int_0^t \ln(1 + \theta_i(u))dA_i(u) \right\} = \lim \exp \left\{ \sum_k \tilde{\alpha} \ln(1 + \theta_i(u_k))(A_i(u_{k+1}) - A_i(u_k)) \right\},
$$

where the limit is in probability and is taken as $\Delta := \max_k |u_{k+1} - u_k| \to 0$. Fixing $\{u_k\}$, we have

$$
\mathbb{E}^{Q_i} \left[ \exp \left\{ \sum_k \tilde{\alpha} \ln(1 + \theta_i(u_k))(A_i(u_{k+1}) - A_i(u_k)) \right\} \right]
$$

(EC.2)

where step (a) is due to independent increments and step (b) follows from the piece-wise constant approximation of a non-homogeneous Poisson process plus using the moment generating function for a Poisson random variable. Note that the piece-wise constant approximation is valid due to the local
integrability of $\theta_i$; see, e.g., (Kim and Whitt 2014). By our hypothesis, $\theta_i$ is bounded. Therefore, we can apply the dominated convergence theorem to get

$$
\mathbb{E}^Q_i \left[ \exp \left\{ \bar{\alpha} \int_0^t \ln(1 + \theta_i(u))dA_i(u) \right\} \right] = \mathbb{E}^Q_i \left[ \lim \exp \left\{ \sum_k \bar{\alpha} \ln(1 + \theta_i(u_k))(A_i(u_{k+1}) - A_i(u_k)) \right\} \right]
$$

$$
= \lim \mathbb{E}^Q_i \left[ \exp \left\{ \sum_k \bar{\alpha} \ln(1 + \theta_i(u_k))(A_i(u_{k+1}) - A_i(u_k)) \right\} \right]
$$

where, again, the limit is taken as $\Delta := \max_k |u_{k+1} - u_k| \to 0$. In light of (EC.2),

$$
\mathbb{E}^Q_i \left[ \exp \left\{ \bar{\alpha} \int_0^t \ln(1 + \theta_i(u))dA_i(u) \right\} \right] = \exp \left\{ \int_0^t \bar{\lambda}_i [(1 + \theta_i(u))^\alpha - (1 + \theta_i(u))] du \right\}.
$$

Taking expectation of (EC.1) and substituting for the preceding expression, we deduce that

$$
\mathcal{R}_i^\alpha(t) := \frac{1}{\alpha - 1} \ln \mathbb{E}^Q_i [\psi_i(t)^{\alpha-1}] = \frac{1}{\alpha - 1} \left\{ \int_0^t \bar{\lambda}_i [(1 + \theta_i(u))^\alpha - (1 + \theta_i(u))] du - \bar{\alpha} \int_0^t \bar{\lambda}_i \theta_i(u) du \right\},
$$

which, after further simplification, leads to the desired result.

**Proof of Proposition 4.** The proof of part (i) involves a series of results, including Lemmas EC.1, EC.2 and EC.3, all of which are concerned with the properties of the function $\pi(\cdot, \cdot)$.

**Lemma EC.1.** Suppose that Assumptions 1 and 2 hold. Then (i) $\pi(w, \cdot)$ is strictly increasing for any fixed $w > 0$; also (ii) for any fixed $\eta$, $\lim_{w \to \infty} \pi(w, \eta)$ not only exists but

$$
\lim_{w \to \infty} \pi(w, \eta) \in \{-\infty, \infty\}.
$$

To proceed, let us define

$$
\eta^* := \inf \left\{ \eta \in \mathbb{R} : \sup_{w \geq 0} \pi(w, \eta) = \infty \right\}.
$$

with the usual convention that $\inf \emptyset = \infty$.

**Lemma EC.2.** Suppose that Assumptions 1 and 2 hold. Then $\eta^* \in (0, \infty]$ and

$$
\sup_{w \geq 0} \pi(w, \eta) = \begin{cases} 
-\infty & \text{if } \eta < \eta^*, \\
\infty & \text{if } \eta \in [\eta^*, \infty] \cap \mathbb{R}.
\end{cases} \quad \text{(EC.3)}
$$

With regard to (24), we need to gain a better understanding of the function $\pi(\cdot, \cdot) - \bar{\ell}_i$. For this purpose, we define

$$
\eta_{\ast,i} := \inf \left\{ \eta > 0 : \sup_{w \geq 0} \pi(w, \eta) \geq \bar{\ell}_i \right\}.
$$
**Lemma EC.3.** Suppose that Assumptions 1 and 2 hold. Then $\eta_{s,i} \in (0, \eta^*)$. Moreover, the equation $\pi(w, \eta) = \bar{\ell}_i$ uniquely defines two $C^1$ functions $w_1(\cdot), w_2(\cdot): (\eta_{s,i}, \eta^*) \mapsto (0, \infty)$ such that $w_1(\eta) < w_2(\eta)$ for all $\eta \in (\eta_{s,i}, \eta^*)$; in addition,

$$w_1(\cdot) \text{ is strictly decreasing, } w_2(\cdot) \text{ is strictly increasing,}$$  

(EC.4)

$$\lim_{\eta \downarrow \eta_{s,i}} w_1(\eta) = \lim_{\eta \uparrow \eta^*} w_2(\eta) = \infty,$$

(EC.5)

and $g(\bar{\ell}_i) + h(w) - \eta > 0$ for all $w > w_2(\eta)$.  

(EC.6)

We are now going to prove part (i) of Proposition 4. To this end, let us define

$$v(\eta) := \int_{u_1(\eta)}^{w_2(\eta)} \left( \pi(u, \eta) - \bar{\ell}_i \right) du - L_i \quad \text{for} \quad \eta \in (\eta_{s,i}, \eta^*),$$

(EC.7)

where $w_1(\cdot)$ and $w_2(\cdot)$ are given as in Lemma EC.3. Our goal is to seek some $\eta_i$ such that $v(\eta_i) = 0$. To this end, we differentiate (EC.7) with respect to $\eta$ and then use the identities

$$\pi(w_1(\eta), \eta) = \pi(w_2(\eta), \eta) = \bar{\ell}_i$$

to obtain

$$v_\eta(\eta) = \int_{w_1(\eta)}^{w_2(\eta)} \pi_\eta(u, \eta) du > 0 \quad \text{for} \quad \eta \in (\eta_{s,i}, \eta^*),$$

(EC.8)

where the inequality follows from Lemma EC.1(i). In addition, from (EC.3) in Lemma EC.2, the continuity of $\pi$ in $\eta$ and Lemma EC.3, it follows that

$$\lim_{\eta \downarrow \eta_{s,i}} v(\eta) = -L_i < 0 \quad \text{and} \quad \lim_{\eta \uparrow \eta^*} v(\eta) = \infty.$$  

(EC.9)

On combining (EC.8) and (EC.9), we conclude that there exists a unique point $\eta_i \in (\eta_{s,i}, \eta^*)$ such that $v(\eta_i) = 0$. Next, letting $q_i := w_1(\eta_i)$ and $s_i := w_2(\eta_i)$ completes the proof of part (i).

To prove part (ii) of the proposition, note that establishing (34) boils down to showing that

$$\frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_i \geq 0 \quad \text{for} \quad w > s_i$$

$$v(w - z) - v(w) + \bar{\ell}_i z + L_i \geq 0 \quad \text{for} \quad z \geq 0, \quad w \in \mathbb{R}_+.$$
for any function $v$ such that $v'(w) = \pi(w, \eta_i)$. The first inequality follows as a direct consequence of (EC.6), while the second inequality follows by a straightforward calculation involving (EC.7) and the fact that

$$
\pi(w, \eta_i) = \begin{cases} 
< \bar{\ell}_i & \text{for } w < q_i, \\
> \bar{\ell}_i & \text{for } w \in (q_i, s_i), \\
= \bar{\ell}_i & \text{for } w > s_i.
\end{cases}
$$

Thus, we complete the proof for part (ii) of the proposition.

Towards proving part (iii) of Proposition 4, let us define

$$
J_i(w) := \max_\zeta \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\zeta \left[ \int_0^t h(W(u)) du - \int_0^t p^*(\zeta(u)) du + \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right],
$$

and let

$$
\delta_i(k) := v(W(\tau_i(k))) - v(W(\tau_i(k)-)).
$$

From (34) it follows that $v(y) - v(x) \leq \tilde{\phi}_i(y - x)$, and so

$$
-\delta_i(k) \leq \tilde{\phi}_i(W(\tau_i(k)-) - W(\tau_i(k))) = \tilde{\phi}_i(\tilde{\xi}_i(k)) \quad \text{for } k = 0, 1, 2, \ldots. \tag{EC.10}
$$

On the other hand, applying the generalized Itô’s formula, we obtain, for $t \geq 0$,

$$
\mathbb{E}^\zeta[v(W(t))] = v(w) + \mathbb{E}^\zeta \left[ \int_0^t \left( \frac{1}{2} \sigma^2 v''(W(u)) + \zeta(u)v'(W(u)) \right) du \right] + \mathbb{E}^\zeta \left[ \sum_{k=0}^{N_i(t)} \delta_i(k) \right].
$$

On substituting (EC.10) into above identity and using (34), we deduce

$$
v(w) \leq \mathbb{E}^\zeta \left[ \int_0^t (h(W(u)) - p^*(\zeta(u)) - \eta_i) du \right] + \mathbb{E}^\zeta \left[ \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right] + \mathbb{E}^\zeta \left[ v(W(t)) \right] + \mathbb{E}^\zeta \left[ \int_0^t (g(v'(W(u))) + p^*(\zeta(u)) - \zeta(u)v'(W(u))) du \right]. \tag{EC.11}
$$

Now, consider a special drift-rate control $\zeta^\#(W)$, defined as

$$
\zeta^\#(W) := \inf_\zeta \max \{ v'(W)\zeta - p^*(\zeta) \}. \tag{EC.12}
$$

Clearly $\zeta^\#(\cdot)$ is an adaptive control satisfying

$$
g(v'(W)) + p^*(\zeta^\#(W)) - \zeta^\#(W)v'(W) = 0. \tag{EC.13}
$$
On combining (EC.11) and (EC.13), we see that
\[
v(w) \leq \mathbb{E}^{\zeta^\#} \left[ \int_0^t (h(W(u)) - p^* (\zeta(u)) - \eta_i) \, du \right] + \mathbb{E}^{\zeta^\#} \left[ \sum_{k=0}^{\mathcal{N}_i(t)} \tilde{\phi}_i (\tilde{\xi}_i(k)) \right] + \mathbb{E}^{\zeta^\#} [v(W(t))]. \tag{EC.14}
\]
Now, dividing both sides of (EC.14) by \( t \), taking the \( \limsup \) as \( t \to \infty \) and using the definition of \( J_i(w) \) plus the fact that \( \zeta^\# \) is an adaptive control, we get
\[
\eta_i \leq J_i(w) + \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^{\zeta^\#} [v(W(t))]. \tag{EC.15}
\]
If \( \limsup_{t \to \infty} (1/t) \mathbb{E}^{\zeta^\#} [v(W(t))] \leq 0 \), then the desired conclusion holds trivially as a result of (EC.15).

Now, suppose for the sake of contradiction
\[
\limsup_{t \to \infty} (1/t) \mathbb{E}^{\zeta^\#} [v(W(t))] > 0.
\]
We now argue that this hypothesis inevitably leads to \( J_i(w) = \infty \), which again yields the desired result. To do so, we adopt the ingenious argument used by Ormeci et al. (2008) in their optimality proof. To begin with, put \( a := \limsup_{t \to \infty} (1/t) \mathbb{E}^{\zeta^\#} [v(W(t))] > 0 \). Then there exists some constant \( \bar{t} > 0 \) such that \((1/t) \mathbb{E}^{\zeta^\#} [v(W(t))] > a/2 \) for \( t \geq \bar{t} \). Since \( v \) has bounded derivatives, it is Lipschitz continuous. Hence there exists some constant \( l > 0 \) such that
\[
v(W(t)) - v(w) \leq l|W(t) - w| \leq l(W(t) + w) \quad \text{for} \quad t \geq 0. \tag{EC.16}
\]
Taking expectation on both sides of (EC.16), we see that
\[
\mathbb{E}^{\zeta^\#} [v(W(t))] - v(w) \leq l \left( \mathbb{E}^{\zeta^\#} [W(t)] + w \right) \quad \text{for} \quad t \geq 0,
\]
which implies that
\[
\mathbb{E}^{\zeta^\#} [W(t)] \geq \frac{1}{l} \left[ v(w) + ta/2 \right] - w = l_1 t + l_2 \quad \text{for} \quad t \geq \bar{t},
\]
where \( l_1 > 0 \) and \( l_2 \in \mathbb{R} \) are two fixed constants. Thus,
\[
J_i(w) \geq \limsup_{t \to \infty} \mathbb{E}^{\zeta^\#} \left[ \frac{1}{l} \int_0^t h(W(u)) \, du \right] = \infty
\]
and we have shown that \( \eta_i \leq J_i(w) \).
In the presence of the maximizing player, we still need to verify that $\zeta^*$ is indeed the maximizer’s best response given the decision maker will commit to the control band policy $(i, q_i, s_i)$. For this purpose, we can easily write down the Bellman equation for the maximizer’s problem: Seek $v_m \in C^2(0, s_i)$ and $\eta_m \in \mathbb{R}$ such that

$$\max_{\zeta} \left\{ \frac{1}{2} \sigma^2 v_m''(w) + \zeta v_m'(w) + h(w) - p^*(\zeta) \right\} = \eta_m,$$

subject to the boundary conditions

$$v_m'(0) = 0 \quad \text{and} \quad v_m(s_i) = \tilde{\phi}_i(s_i - q_i) + v_m(q_i).$$

Comparing these with (19) and (20), we immediately conclude that $v_m = v$ and $\eta_m = \eta_i$. Therefore, the strategy $\zeta^*$ defined by (EC.12) is the maximizer’s best response given the decision maker chooses to adopt the control band policy $(i, q_i, s_i)$.

Finally, noting that $\eta_i$ is the long-run average cost when the decision maker implements $(i, q_i, s_i)$ and the maximizer employs the drift-rate control $\zeta^*$ (cf. Proposition 2) completes the proof.

**Proof of Theorem 3.** Towards proving part (i) of the theorem, note that by definition, the value function $v$ automatically satisfies

$$\frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_i \geq 0 \quad \text{and} \quad \inf_{z \geq 0} \left[ v(w - z) + \tilde{\phi}_i(z) - v(w) \right] \geq 0, \quad w \geq 0.$$ 

Hence to establish (34), it suffices to argue

$$\inf_{z \geq 0} \left[ v(w - z) + \tilde{\phi}_i(z) - v(w) \right] \geq 0 \quad \text{for} \quad w \geq 0, \quad i \neq i^*.$$  

(EC.17)

To this end, we rule out three uninteresting cases: (i) $L_i \geq L_{i^*}, \tilde{\ell}_i \geq \tilde{\ell}_{i^*}$, (ii) $L_i \leq L_{i^*}, \tilde{\ell}_i < \tilde{\ell}_{i^*}$, and (iii) $L_i < L_{i^*}, \tilde{\ell}_i \leq \tilde{\ell}_{i^*}$. Indeed, from condition (i) it follows that

$$\inf_{z \geq 0} \left[ v(w - z) + \tilde{\phi}_i(z) - v(w) \right] \geq \inf_{z \geq 0} \left[ v(w - z) + \tilde{\phi}_{i^*}(z) - v(w) \right] \geq 0,$$
hence (EC.17) trivially satisfied. On the other hand, if condition (ii) or (iii) holds, then it is immediate that $\eta_i < \eta_{i^*}$, contradicting our hypothesis that $\eta_{i^*} = \min_i \eta_i$. Therefore, with loss of generality, we may assume either (a) $L_i < L_{i^*}, \bar{\ell}_i > \bar{\ell}_{i^*}$ or (b) $L_i > L_{i^*}, \bar{\ell}_i < \bar{\ell}_{i^*}$ for $i \neq i^*$.

We first focus on case (a). Suppose by way of contradiction that

$$\inf_{z \geq 0} \left[ v(w-z) + \bar{\phi}_i(z) \right] - v(w) < 0$$

for some $w$. Thus

$$\sup_{0 \leq q \leq s} \int_q^s \left[ v'(z) - \bar{\ell}_i \right] > L_i.$$

To write out the maximum on the left-hand side more explicitly, we define the upper-level set of function $v'$ as $\mathbb{L}_{v'}(l) := \{ x \geq 0 : v'(x) \geq l \}$. Note that the proof of part (ii) in Proposition 4 implies that $v'$ is quasi-concave on $[0, s_{i^*}]$ plus $v'(s_{i^*}) = \bar{\ell}_{i^*}$. Therefore, there exists some $\bar{q}$ and $\bar{s}$ such that $\mathbb{L}_{v'}(\bar{\ell}_i) = [\bar{q}, \bar{s}] \subset [q_{i^*}, s_{i^*}]$. Moreover, by the definition of $q_{i^*}$ and $s_{i^*}$,

$$\tilde{L} := \int_{\bar{q}}^{\bar{s}} \left[ v'(z) - \bar{\ell}_i \right] = \sup_{0 \leq q \leq \bar{s}} \int_q^s \left[ v'(z) - \bar{\ell}_i \right] > L_i.$$

Now, consider a new type of outsourcing operations whose cost function is given as

$$\tilde{\phi}(w) := (\tilde{L} + \bar{\ell}_iw) \cdot 1_{\{w > 0\}} + 0 \cdot 1_{\{w = 0\}}.$$

By construction, a control band policy with parameters $(\bar{q}, \bar{s})$ and cost function $\tilde{\phi}$ shares the same value function with $(i^*, q_{i^*}, s_{i^*})$, and so the average cost must be equal to $\eta_{i^*}$. On the other hand, it is immediate from the relationship $\tilde{L} > L_i$ that $\eta_{i^*} > \eta_i$. This leads to a contradiction because $\eta_{i^*}$ is chosen to satisfy $\eta_{i^*} \leq \eta_i$ for all $i$. This completes our proof for case (a). The proof for case (b) is very similar; thus we leave it as an exercise. So we complete the proof of part (i).

The proof of part (ii) follows closely the steps in the proof of part (iii) in Proposition 4. Thus, we only highlight the key differences. To start, let

$$J(w) := \max_{\zeta} \limsup_{t \to \infty} \frac{1}{l} \mathbb{E}[\zeta] \left[ \int_0^t h(W(u))du - \int_0^t p^*(\zeta(u))du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\xi_i(k)) \left| W(0) = w \right| \right],$$
and define $\delta_i(k)$ in the same way as we did in the proof of Proposition 4. Now using (35) instead of (34) this time, we conclude that, for all $i$, $v(y) - v(x) \leq \tilde{\phi}_i(y - x)$; thus for all $k = 1, 2, \ldots$ and $i = 1, \ldots, I$, we have

$$-\delta_i(k) \leq \tilde{\phi}_i(\tilde{\xi}_i(k)).$$  \hspace{1cm} (EC.18)

Next by applying the generalized Itô's formula, we obtain, for $t \geq 0$,

$$E^c[v(W(t))] = v(w) + E^c \left[ \int_0^t \left( \frac{\sigma^2}{2} v''(W(u)) + \zeta(u)v'(W(u)) \right) du \right] + E^c \left[ \sum_{i=1}^{I} \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right].$$

On substituting (EC.18) into above identity and using (35), we deduce

$$v(w) \leq E^c \left[ \int_0^t (h(W(u)) - p^*(\zeta(u)) - \eta_i) du \right] + E^c \left[ \sum_{i=1}^{I} \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right]$$

$$+ E^c [v(W(t))] + E^c \left[ \int_0^t (h(W(u)) - \zeta(u)v'(W(u))) du \right].$$

The rest of the proof proceeds in exactly the same fashion as the proof of Proposition 4. First, by choosing $\zeta = \zeta^\#$ with $\zeta^\#$ given as in (EC.12), one can formally show that $\eta_{i^\#} \leq J(w)$. Second, one can easily argue that $\zeta^\#$ is the maximizer's best response — when the decision maker chooses $(i^\#, q_{i^\#}, s_{i^\#})$, the maximizer will follow $\zeta^\#$ and never deviate from it.

**EC.2. Numerical Algorithm for Solving the Optimality Equation**

To find the solution of the optimality equation (19), we start with an initial guess of $v$, denoted as $v_0$, that solves

$$\frac{1}{2} \sigma^2 v''_0(w) + h(w) = \eta_i, \quad w \in (0, s_i)$$  \hspace{1cm} (EC.19)

subject to the boundary conditions $v'_0(0) = 0$, $v_0(s_i) = \tilde{\phi}_i(s_i - q_i) + v_0(q_i)$ and necessary optimality conditions $v'_0(q_i) = v'_0(s_i) = \tilde{\ell}_i$. Notice that (EC.19) is a second-order linear ordinary differential equation, so we can solve it analytically. Then for each $w \in (0, s_i)$, we seek $\zeta_0 := \zeta_0(w)$ that maximizes $\{\zeta_0 v'_0(w) - r^*(\zeta_0)\}$. Recall function $r^*$ has been given in section 6.2. The next step is to find $v_1$ such that

$$\frac{1}{2} \sigma^2 v''_1(w) + \zeta_0 v'_1(w) + h(w) - r^*(\zeta_0) = \eta_i, \quad w \in (0, s_i)$$  \hspace{1cm} (EC.20)
subject to the same boundary conditions and necessary optimality conditions as mentioned previously.

We can solve EC.20 numerically via finite difference method (FDM).

In general, using the $k$th estimate of $v$, denoted as $v_k$, we can find $\zeta_k := \zeta_k(w)$ that maximizes

$$\{\zeta_k v'_k(w) - r^*(\zeta_k)\},$$

and further solve the ordinary differential equation

$$\frac{1}{2} \sigma^2 v''_{k+1}(w) + \zeta_k v'_{k+1}(w) + h(w) - r^*(\zeta_k) = \eta_i, \quad w \in (0, s_i)$$  (EC.21)

subject to the set of boundary conditions and necessary optimality conditions by using FDM to get $v_{k+1}$, the $(k+1)$th estimate of $v$. Repeating these steps we obtain an iterative procedure that generates a sequence $\{v_{k+1}, \zeta_k\}$ which is expected to converge to the optimal solution when $k \to \infty$. Although we do not attempt to rigorously prove the desired convergence result, our extensive numerical experiments suggest convergence happens after a few iterations. The algorithm terminates when the iteration error $||v_{k+1} - v_k||$ and $||\zeta_{k+1} - \zeta_k||$ become sufficiently small.

EC.3. Additional Numerical Results

<table>
<thead>
<tr>
<th>$(\gamma_1, \gamma_2)$</th>
<th>$q_1$</th>
<th>$s_1$</th>
<th>$\eta_1$</th>
<th>$\eta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6, 6)</td>
<td>0.065</td>
<td>0.575</td>
<td>74.713</td>
<td>97.499</td>
</tr>
<tr>
<td>(6, 10)</td>
<td>0.065</td>
<td>0.570</td>
<td>71.544</td>
<td>95.367</td>
</tr>
<tr>
<td>(6, 20)</td>
<td>0.070</td>
<td>0.570</td>
<td>69.326</td>
<td>92.256</td>
</tr>
<tr>
<td>(6, 100)</td>
<td>0.070</td>
<td>0.565</td>
<td>67.655</td>
<td>89.947</td>
</tr>
<tr>
<td>(10, 6)</td>
<td>0.095</td>
<td>0.530</td>
<td>48.670</td>
<td>62.037</td>
</tr>
<tr>
<td>(10, 10)</td>
<td>0.105</td>
<td>0.525</td>
<td>45.472</td>
<td>57.687</td>
</tr>
<tr>
<td>(10, 20)</td>
<td>0.105</td>
<td>0.525</td>
<td>43.253</td>
<td>54.662</td>
</tr>
<tr>
<td>(10, 100)</td>
<td>0.110</td>
<td>0.520</td>
<td>41.583</td>
<td>52.386</td>
</tr>
<tr>
<td>(20, 6)</td>
<td>0.140</td>
<td>0.520</td>
<td>32.868</td>
<td>40.071</td>
</tr>
<tr>
<td>(20, 10)</td>
<td>0.150</td>
<td>0.520</td>
<td>29.844</td>
<td>35.908</td>
</tr>
<tr>
<td>(20, 20)</td>
<td>0.160</td>
<td>0.525</td>
<td>27.755</td>
<td>33.099</td>
</tr>
<tr>
<td>(20, 100)</td>
<td>0.170</td>
<td>0.530</td>
<td>26.214</td>
<td>31.025</td>
</tr>
<tr>
<td>(100, 6)</td>
<td>0.190</td>
<td>0.540</td>
<td>22.850</td>
<td>26.473</td>
</tr>
<tr>
<td>(100, 10)</td>
<td>0.210</td>
<td>0.550</td>
<td>20.257</td>
<td>22.973</td>
</tr>
<tr>
<td>(100, 20)</td>
<td>0.225</td>
<td>0.560</td>
<td>18.565</td>
<td>20.718</td>
</tr>
<tr>
<td>(100, 100)</td>
<td>0.235</td>
<td>0.570</td>
<td>17.345</td>
<td>19.130</td>
</tr>
<tr>
<td>(10000, 10000)</td>
<td>0.260</td>
<td>0.590</td>
<td>15.428</td>
<td>16.660</td>
</tr>
</tbody>
</table>
Table EC.1 provides the optimal thresholds and asymptotic bounds for the simulation program in the main paper.

**EC.4. Proof of Auxiliary Results**

*Proof of Lemma EC.1.* Part (i) of the lemma follows directly from the definition of \( \pi(w, \eta) \) and the comparison principle for the solutions of first-order ordinary differential equations. With reference to part (ii), we start by arguing that

\[
\lim_{w \to \infty} \inf \pi(w, \eta), \quad \lim_{w \to \infty} \sup \pi(w, \eta) \in \{-\infty, \infty\}.
\]

(EC.22)

Suppose for the sake of contradiction that \( \lim_{w \to \infty} \inf \pi(w, \eta) \in (-\infty, \infty) \). Then choose a sequence \( w_n \to \infty \) such that

\[
\lim_{n \to \infty} \pi(w_n, \eta) = \lim_{w \to \infty} \inf \pi(w, \eta) \quad \text{and} \quad \lim_{n \to \infty} \pi_w(w_n, \eta) = 0.
\]

If assuming \( \limsup_{w \to \infty} \pi(w, \eta) \in (-\infty, \infty) \), then choose a sequence \( w_n \in \infty \) in a similar fashion. In either case, we obtain via direct calculation that

\[
0 = \lim_{n \to \infty} \pi_w(w_n, \eta) = \lim_{n \to \infty} \frac{-2}{\sigma^2} \left[ g(\pi(w_n, \eta)) + h(w_n) - \eta \right] = -\infty,
\]

where the last equality is owing to Assumption 1. However, the preceding calculation provides the desired contradiction, so we must have (EC.22). To complete the proof of part (ii), it suffices to argue that the number of solutions of \( \pi(w, \eta) = 0 \) is finite. To this end, we fix any \( \eta \in \mathbb{R} \), and evaluate the solvability of \( \pi(w, \eta) = 0 \) on \((0, \infty)\). Assumption 1 implies that there exist at most one point \( w(>0) \) that satisfies \( h(w) = \eta \). Moreover, from (22) it follows that

\[
\pi_w(w, \eta) = -\frac{2}{\sigma^2} [h(w) - \eta] \quad \text{for all} \quad w > 0 \quad \text{s.t.} \quad \pi(w, \eta) = 0.
\]

(EC.23)

This observation along with the boundary condition \( \pi(0, \eta) = 0 \) allow us to conclude that the number of solutions of \( \pi(w, \eta) = 0 \) over \((0, \infty)\) is no greater than the number of solutions of \( h(w) = \eta \) on \((0, \infty)\), which is at most one. This shows that the number of solutions of \( \pi(w, \eta) = 0 \) on \( \mathbb{R} \) is finite, hence completing the proof of the lemma.

\[ \Box \]
**Proof of Lemma EC.2.** In view of (22) and the non-negativity of \( h \) and \( g \), we see that

\[
\pi(w, \eta) \leq 0 \quad \text{for all } w \in \mathbb{R} \text{ and } \eta \leq 0,
\]

which in turn implies that \( \eta^* \in (0, \infty] \). Because \( \pi(w, \cdot) \) is strictly increasing for any fixed \( w > 0 \),

\[
\sup_{w \geq 0} \pi(w, \eta) < \infty \quad \text{for } \eta < \eta^* \quad \text{and} \quad = \infty \quad \text{for } \eta \in (\eta^*, \infty] \cap \mathbb{R}.
\]

In view of Lemma EC.1(ii), proving (EC.3) boils down to showing that

\[
\lim_{w \to \infty} \pi(w, \eta^*) = \infty \quad \text{if } \eta^* < \infty.
\]

Suppose for the sake of contradiction that \( \lim_{w \to \infty} \pi(w, \eta^*) = -\infty \). This limit, in conjunction with Assumption 1, implies that there exists \( \hat{w} > 0 \) such that

\[
\pi(w, \eta^*) < 0 \quad \text{and} \quad h(w) - \eta^* > \chi > 0 \quad \text{for all } w \geq \hat{w},
\]

where \( \chi \) is any given constant. Now, combining the second inequality in (EC.24) with (EC.23), we see that, for any \( \eta \in [\eta^*, \eta^* + \chi] \),

\[
\pi_w(w, \eta) < 0 \quad \text{for all } w \geq \hat{w} \quad \text{such that} \quad \pi(w, \eta) = 0.
\]

However, this calculation together with the fact that \( \lim_{w \to \infty} \pi(w, \eta) = \infty \) for all \( \eta > \eta^* \) implies that there exists no \( w \geq \hat{w} \) such that \( \pi(w, \eta) = 0 \) when \( \eta \in (\eta^*, \eta^* + \chi] \), and hence

\[
\pi(w, \eta) > 0 \quad \text{for all } w \geq \hat{w} \quad \text{and} \quad \eta \in (\eta^*, \eta^* + \chi] .
\]

But the preceding inequality contradicts the first inequality in (EC.24), due to the continuity of \( \pi \) in \( \eta \). Thus, we must have \( \sup_{w \geq 0} \pi(w, \eta^*) = \infty \), hence (EC.3). \( \square \)

**Proof of Lemma EC.3.** By the fact that \( h \geq 0 \) and \( \bar{\ell}_i > 0 \), it follows that \( \eta_{*,i} > 0 \). Also, by the definitions of \( \eta_{*,i}, \eta^* \) and the continuity of \( \pi \), we see that \( \eta_{*,i} < \eta^* \). Also, from (22) it follows that

\[
\pi_w(w, \eta) = \frac{2}{\sigma^2} \left[ g(\bar{\ell}_i) + h(w) - \eta \right] \quad \text{for } w > 0 \quad \text{such that} \quad \pi(w, \eta) = \bar{\ell}_i.
\]

In view of the definitions of \( \eta_{*,i} \) and \( \eta^* \), (EC.3) in Lemma EC.2, the fact that \( \pi(0, \eta) = 0 \) and the continuity of \( \pi \), the preceding calculation yields the following:
(a) If $\eta < \eta^*_i$, then the equation $\pi(w, \eta) = \tilde{\ell}_i$ has no solution over $(0, \infty)$;

(b) If $\eta \in (\eta^*_i, \eta^*)$, then the equation $\pi(w, \eta) = \tilde{\ell}_i$ has one solution $w_1(\eta) > 0$ such that

$$h(w_1(\eta)) + g(\tilde{\ell}_i) - \eta < 0,$$

and one solution $w_2(\eta) > w_1(\eta)$ such that

$$h(w_2(\eta)) + g(\tilde{\ell}_i) - \eta > 0.$$

(c) If $\eta > \eta^*$, then the equation $\pi(w, \eta) = \tilde{\ell}_i$ has one solution $w_1(\eta) > 0$ such that

$$h(w_1(\eta)) + g(\tilde{\ell}_i) - \eta < 0.$$

Now, (EC.6) follows from Assumption 1 and (b). Moreover, the first assertion in (EC.5) follows from (a) and (b), whereas the second assertion follows from (b) and (c) as well as (EC.4). Hence, to finish the proof, it suffices to establish (EC.4). To this end, differentiate $\pi(w_j(\eta), \eta) = \tilde{\ell}_i$ with respect to $\eta$ to get

$$\frac{\partial w_j}{\partial \eta}(\eta) = -\frac{\pi_w(w_j(\eta), \eta)}{\pi_u(w_j(\eta), \eta)} = \frac{\sigma^2 \pi_u(w_j(\eta), \eta)}{2 \left( h(w_j(\eta)) + g(\tilde{\ell}_i) - \eta \right)}$$

for all $\eta \in (\eta^*_i, \eta^*)$. In view of Lemma EC.1(i) and (b), the above calculation implies that $w_1(\cdot)$ is strictly decreasing and $w_2(\cdot)$ is strictly increasing. This completes the proof. $\square$