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, an optimal control policy is typically derived on the premise that a precise probabilistic model is given. In many situations, however, the underlying probabilistic model is merely a simplification of the real-world scenario that arises due to a lack of operational data to calibrate the model and/or tractability considerations. Thus, a policy derived from such a simplification may perform sub-optimally if the assumed model does not capture reality well. This paper advances a modeling paradigm that accounts for model uncertainty in controlling a multi-product make-to-order manufacturing system with an outsourcing mechanism. It focuses on model misspecification in demand for different products and introduces different robust control formulations using a notion of Rényi divergence to describe ambiguity. Focusing on one formulation, we develop an approximating problem, which is effectively a stochastic differential game. We illustrate how the stochastic differential game can be solved via dynamic programming and, based on its solution, propose an implementable control rule in the context of the original make-to-order system. We also present a data-driven method for selecting an appropriate uncertainty set. Numerical experiments are conducted to demonstrate the value of building “robustness” into decision-making.

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

1. Introduction

This paper concerns 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 possesses the option of subcontracting or outsourcing its manufacturing needs for those products at a fixed plus proportional outsourcing cost in response to unexpected upticks in demand for the products.

One example that fits within this framework is additive manufacturing, also known as 3D printing, which utilizes 3D printers to produce a variety of products on-demand. Printable products are ones that are printed directly from a computer model. As a result, 3D printing operations cannot stockpile inventory in order to decouple demand variability from production. In the presence of demand surges, a subcontractor, such as a professional 3D printing bureau, can be called upon to
supplement in-house manufacturing efforts (Kantaros et al. 2022). Orders for 3D-printed products can be prioritized based on their relative urgency; recent academic research has well documented the benefits of using 3D printing in a make-to-order environment with heterogeneous demands; see, for example, Li et al. (2019). 3D printing technology is on the rise, as evidenced by its wide adoption by big companies like Nike, New Balance, and Adidas, which use 3D printers to produce athletes’ shoes, custom-made shoes, and sneakers. As a potentially transformational technology, 3D printing is expected to have a massive impact on the future supply chain relationships (Arbabian and Wagner 2020).

A make-to-order system can be modeled as a queuing system, 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 techniques (Plambeck et al. 2001, Çelik and Maglaras 2008, Rubino and Ata 2009). While the problem of controlling a make-to-order system has received considerable attention, the majority of the research assumes the probability law describing realized demand is precisely known; as a result, the real-world performance of a policy would depend on how faithful the probabilistic model is to reality.

In practice, the demand model may be misspecified for a variety of reasons. For one thing, many demand models are parametric in nature, which means that parameter estimates may be uncertain. As a result, a model may be misspecified due to a lack of historical data to calibrate the model. Model misspecification can also arise due to simplifying assumptions such as “Markovian” and “stationarity” made for tractability considerations, despite the fact that the real-world situation may be much more complex. In principle, one can always come up with a more complex, “high-fidelity” model to describe realized demand for a given product. For example, a doubly stochastic Poisson process with an auto-regressive intensity process may be well suited to capture auto-correlation among demand arrivals. Such a model, however, is bound to introduce new modeling and computational challenges. For one thing, to optimize decisions, one would need to know the value of intensity over time, which is rarely observable. Another thing is that, even if its value can be observed, the problem dimension is raised, making the problem computationally excruciating. As a result, a simpler, low-fidelity model may be adopted instead, introducing model errors.

A Bayesian framework may prove useful when exact knowledge of the model parameters is lacking. This approach starts by assuming a parametric family of distributions to which the true distribution belongs 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, since the effectiveness of Bayesian models depends on the choice of the prior, they can be effective if one is certain about how to choose the prior. In reality, one is unlikely to be sure about the prior, so incorrect parameter values
may prevail. An alternative modeling paradigm is robust optimization (RO), introduced in part to alleviate the tension between specifying a high-fidelity model (that better replicates reality) and achieving tractability (possibly by adopting a low-fidelity model). RO addresses model misspecification by introducing the so-called uncertainty set, which is thought to contain the true model parameters; see Bertsimas and Thiele (2006) for an excellent survey of early works in this field, as well as Mamani et al. (2017), Bandi et al. (2019), Sun and Van Mieghem (2019) for some recent contributions. Although classical RO has proven useful in a variety of applications, it is inherently a static approach. As a result, it may produce overly conservative solutions for problems that involve not only here-and-now but also wait-and-see decisions.

This paper aims at presenting a framework that can yield good decision rules for a make-to-order manufacturing system when (a) the correct demand model is known but difficult to describe or calibrate using the available data, or (b) the decision maker is not perfectly sure about the probabilistic model governing realized demands. As such, the decision maker faces model uncertainty in contemplating control strategies. We adopt an approach that is closer in spirit to the line of research pioneered by Petersen et al. (2000), Hansen and Sargent (2001). The idea is to extract a nominal model from the available data and add a malevolent second player (“nature”) that perturbs the nominal model within some prescribed limits. Nature’s malevolence serves as the decision maker’s tool for analyzing the fragility of alternative decision rules. Perturbations allow random shocks to feed back into the state process. In this regard, the uncertainty set is only vaguely specified, rendering the approach different from the classical RO setting.

In greater detail, we assume demands for each product arrive according to a non-homogeneous Poisson process. However, the decision maker has no idea what the true value of arrival intensity is, beyond knowing that it fluctuates around some long-term average (which hints at the nominal model). Thus, the decision maker acts as if the intensity is chosen by nature. The decision maker calculates the expected long-run average cost while assuming that nature always creates the worst-case scenario and seeks a joint sequencing and outsourcing rule to minimize it (minimax criterion). We consider two alternative formulations of robust control problems: the constraint problem and the penalty problem. The former shares high-level similarities with classical RO settings in that it imposes direct constraints restricting the magnitude of the perturbation to the nominal model. The latter constrains nature’s actions by penalizing deviations from the nominal model—to what extent nature is punished depends on how much the decision maker distrusts the nominal model. Both formulations use a notion of Rényi divergence to measure the gap between the nominal model and a perturbation to it. Both formulations can be interpreted as stochastic games in which the decision maker seeks to take the best action in response to nature’s alternation of the belief reflected in the nominal model. Under suitable conditions, we establish the equivalence of these two formulations.
Our consideration of an infinite time horizon is motivated by the fact that there is typically no fixed end date for a manufacturing firm to close down. And we adopt an average cost criterion as opposed to the more traditional discounted cost criterion for a couple of reasons. First, the concept of a discount factor is generally alien to manufacturing settings, despite being natural in many applications, such as finance. Second, researchers have traditionally focused on the discounted problem because that version of the dynamic programming operator has favorable contraction properties that make analysis easier. Such consideration, however, has no bearing on the methodologies adopted in this paper. Third, an average cost problem is more amenable to analysis since it avoids dealing with a second-order differential equation.

Our choice of Rényi divergence to construct uncertainty sets is motivated by three key observations. First, Rényi divergence, as a means of quantifying the discrepancy between two probability measures, has a multitude of applications. For instance, it is used in variational inference (Li and Turner 2016), uncertainty quantification for rare events (Dupuis et al. 2020), and coding theory and hypothesis testing (Van Erven and Harremos 2014), among others. It is also believed to have several advantages over the commonly used Kullback-Leibler (KL) divergence, including the ability to compare heavy-tailed distributions and certain nonabsolutely continuous distributions; see, for example, Song and Ermon (2019). Despite being a special family of divergence measures in its own right, Rényi divergence encompasses or is closely related to important divergence measures such as the Bhattacharyya distance, KL divergence, \(\chi^2\)-divergence, and total variation distance, among others. Second, since Rényi divergence usefully generalizes (and thus incorporates as a special case) KL divergence, it offers decision makers greater flexibility in terms of constructing the uncertainty set. Indeed, our extensive numerical experiments show that the use of Rényi divergence leads to better control rules compared to using KL divergence only. Third, Rényi divergence enables a tractable representation of the distance between models, which simplifies analysis and computation. While \(f\)-divergence is another powerful generalization of KL divergence, because it does not translate to a tractable representation of the distance between models in the specific context considered by this paper, we do not consider this generalization in this work.

Focusing on one formulation, we develop an approximating problem that is analytically more tractable. We refer to this approximation as the stochastic differential game (SDG). 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. We demonstrate that the SDG can further be converted to an equivalent one-dimensional differential game, whose state descriptor is one-dimensional and tracks the amount of work in the system over time. The conversion thus 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 a single band control 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 the decision maker’s ambiguity aversion. We summarize the contributions in our paper as follows.

• First, it investigates joint sequencing and outsourcing control for a make-to-order system with fixed plus proportional costs for outsourcing operations under model uncertainty. To the best of our knowledge, this is the first paper that accounts for model misspecification and fixed outsourcing costs in the literature that jointly considers these controls. Moreover, the use of Rényi divergence to describe model uncertainty extends the celebrated entropic approach in the literature.

• Second, assuming the capacity is adequately utilized, we derive and solve an SDG that approximates the original penalty problem. In the SDG, the maximizing player (i.e., nature) can control the drift of the underlying state process, making the problem more challenging. To be more specific, one cannot identify a pathwise optimal solution for the minimizing player (the decision maker). Therefore, one would need to deal with a complex nonlinear differential equation with free boundaries. The fixed outsourcing costs are also a crucial part of the solution analysis of the SDG. When these fixed costs are zero, outsourcing controls boil down to singular controls in the SDG, and one can read off the cheapest product to outsource when the backlog of work is judged to be excessive. In the presence of non-zero fixed outsourcing costs, however, it is not immediately clear which product is the cheapest in terms of outsourcing. To resolve it, we demonstrate that solution of the game can be found by repetitively solving the differential equation a finite number of times, each time using a different set of boundary conditions. The optimality proof uses a novel level-set argument.

• Third, it proposes a data-driven method that allows one to identify a proper uncertainty set (i.e., an appropriate level of ambiguity) in a computationally efficient manner. Specifically, the method allows for a meaningful combination of an analytical framework with computer simulations to generate and evaluate a family of closely related control rules. The output of this method is a control rule thought to be the best for real-world implementation.

2. Literature Review
This work 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 include Rubino and Ata (2009), Ata and Tongarlik (2013), Kim and Ward (2013), Ghamami and Ward (2013). In the context of managing make-to-order systems, a few papers (Celik 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 work incorporating outsourcing decisions typically assumes only proportional costs, whereas we consider fixed costs in addition to many practical settings, which are common in many practical settings.

Closer in spirit to our work is 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 who promotes robustness. The nominal model is believed to be the best representation of the real-world scenario based on one’s limited knowledge, while a perturbation to it accounts for the possibility of model misspecification. This concept can be traced back to the early works of Petersen et al. (2000), Hansen and Sargent (2001) 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), and probability of lifetime ruin (Bayraktar and Zhang 2015), among others. Usually, the problem can be translated into a two-player stochastic game, which can then be solved through a dynamic programming equation. Addressing certain fundamental questions associated with this equation often makes up the major technical hurdles in the analysis. Additionally, these papers use the notion of KL divergence to capture ambiguous beliefs. This paper is one of the first to apply divergence measures beyond KL divergence to characterize model uncertainty in stochastic dynamic programming problems. In this respect, we feel that the present paper signifies an important step forward in expanding this powerful modeling paradigm.

Our work shares some similarities with a recent paper by Cohen (2019) that studies a Brownian control problem arising from the heavy-traffic approximation of a multiclass M/M/1 queue under model uncertainty. However, our work distinguishes Cohen (2019) in several aspects. (i) Cohen (2019) uses an infinite-horizon discounted cost formulation, whereas we use a long-run average cost criterion. (ii) In Cohen (2019), workload control incurs only proportional cost; thus, the workload control is of the barrier type, and the cheapest class to outsource or reject is determined as a result of the workload reformulation. However, we incur both fixed and proportional costs in our workload control, leading to a single band control, and determining the cheapest queue to push down requires
more involved analysis. (iii) Cohen (2019) employs the entropic method, which results in a quadratic nonlinear term in the dynamic programming equation. With Rényi-type ambiguities, the nonlinear term of our dynamic programming equation is not necessarily quadratic. The main challenge in Cohen (2019) is analyzing a second-order differential equation with a quadratic nonlinear term, whereas the major hurdle in this paper is dealing with the general nonlinear term in the dynamic programming equation plus free boundary conditions. As a result, the methodologies in Cohen (2019) and this paper do not encompass each other.

Methodologically speaking, the paper is related to impulse control of Brownian systems. Two-sided impulse control of a Brownian motion having a constant drift rate has been widely studied in the literature; see, for example, 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 prove that a control band policy \((a_1, a_2, b_1, b_2)\) with \(a_1 < a_2 < b_1 < b_2\) is optimal. 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; and (ii) their problems are of a single-class nature, whereas ours is inherently a multiclass model.

3. Nominal Model

All random quantities of interest in this section are on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) endowed with a filtration \((\mathcal{F}_t)\) contained in \(\mathcal{F}\). Consider a make-to-order manufacturing system that offers \(I\) different products, indexed 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, arrive according to a Poisson process with rate \(\bar{\lambda}_i\). For each \(i\), \(A_i(t)\) denotes the number of class \(i\) orders placed up to time \(t\). Denote by \(\bar{\lambda} := (\bar{\lambda}_i)\) the arrival rate vector that collects all the arrival rates. The time it takes to process a class \(i\) order follows an exponential distribution with rate \(\mu_i\) and mean \(m_i = 1/\mu_i\). We use \(S_i(t)\) to denote the number of completed class \(i\) orders up to \(t\) if the server is constantly working on class \(i\) orders. The decision maker has discretion over the 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 $i$. Clearly, $(t - \sum_i T_i(t))$ tracks the total idle time up to time $t$. The decision maker is able to 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 this, the decision maker is able to reduce its backlog of class $i$ orders by $x$ units instantly; for convenience, we will refer to this type of outsourcing activity as type $i$ outsourcing operation.

To proceed, let $\Psi := (\Psi_i)$, where its $i$th component is specified by

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

where $0 = \tau_i(0) < \tau_i(1) < \tau_i(2) < \cdots < \tau_i(k) < \cdots$ is a sequence of time epochs at which a type $i$ outsourcing operation is performed, and $\{\xi_i(k); k \geq 0\}$ represents the sequence of batch sizes of the consecutive type $i$ outsourcing operations. Let $N_i(t) := \sup\{k \geq 0 : \tau_i(k) \leq t\}$, so that $N_i(t)$ tracks the number of type $i$ outsourcing operations performed up to $t$.

Denote by $Q_i(t)$ 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

$$Q_i(t) = Q_i(0) + A_i(t) - S_i(T_i(t)) - \sum_{k=0}^{N_i(t)} \xi_i(k) \quad \text{for} \quad i = 1, \ldots, I. \quad (1)$$

The system incurs costs 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. Therefore, the total backlog penalty accrues at the rate of $\sum_i c_i(Q_i(t))$.

A control policy can then be represented by the pair of controls $(T, \Psi)$. Adopting the long-run average cost criterion, the decision maker seeks an adaptive control $(T, \Psi)$ to minimize

$$\limsup_{t \to \infty} \frac{1}{t} \mathbb{E} \left[ \sum_{i=1}^{I} \int_0^t c_i(Q_i(u)) du + \sum_{i=1}^{I} \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) \right]. \quad (2)$$

The primary goal of this paper is to formulate and solve robust versions of (2).

4. Robust Control Problem

In reality, the demand rate for product $i$ may be some function of time (possibly random), $\lambda_i := \{\lambda_i(t); t \geq 0\}$, rather than a constant value $\bar{\lambda}_i$. With this in mind, the decision maker has a keen interest in pursuing control rules that are robust against such departures. To capture the idea that the decision maker distrusts the nominal model (which assumes constant arrival rates), assume that the decision maker acts as if there is a second player (i.e., nature) who chooses $\lambda := (\lambda_i)$
strategically. Associate with each \( \lambda_i \), one can define the perturbation process \( \theta_i := \{ \theta_i(t); t \geq 0 \} \), where \( \theta_i(t) := (\lambda_i(t) - \bar{\lambda}_i) / \bar{\lambda}_i \) describes the relative deviation of \( \lambda_i \) from its nominal value \( \bar{\lambda}_i \) at time \( t \). We require that each \( \theta_i \) be such that \( \theta_i(t) \in \Theta_i := [a_i, b_i] \) for all \( t \geq 0 \) with \(-1 < a_i < 0 < b_i < \infty \) and remains constant between consecutive state changes (with possible randomization at times when state changes). Let \( \theta := (\theta_i) \). It is straightforward to check that \( \lambda \) and \( \theta \) are determined by each other. For this reason, we will henceforth treat \( \theta \) rather than \( \lambda \) as nature’s decision process.

As is customary in the stream of research pioneered by Petersen et al. (2000), Hansen and Sargent (2001), to reflect the decision maker’s aversion towards model uncertainty, we assume nature chooses \( \theta \) adaptively with the goal of inflating the decision maker’s cost to the greatest extent possible. The consideration of the worst-case not only conforms to the general idea underpinning RO, but also, as argued by Epstein and Schneider (2003), shields the decision maker from what they call “dynamic inconsistency” under appropriate conditions.

Degrees of model uncertainty are managed by imposing appropriate constraints on nature, which aim to limit the size of the perturbations to the nominal model. To that end, we define, for each \( i = 1, \ldots, I \), 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 \( \Delta A_i(t) := A_i(t) - A_i(t-) \). It is easy to see that if \( \theta_i \equiv 0 \), then \( \psi_i \equiv 1 \) and hence \( \lambda_i \equiv \bar{\lambda}_i \). It is also straightforward to check that \( \psi_i := \{ \psi_i(t); t \geq 0 \} \) is a martingale. Thus, if using \( \mathbb{P}_i \) to denote the (marginal) law governing \( A_i \) in the nominal model, we can define a new measure \( \mathbb{Q}_i \) via

\[
\frac{d\mathbb{Q}_i}{d\mathbb{P}_i} \bigg|_{F_t} = \psi_i(t) \quad \text{for} \quad t \geq 0.
\]  

(4)

The Girsanov theorem for filtered Poisson processes (see, for example, Proposition 8.4.5.1 in Jeanblanc et al. (2009)) then tells us that \( A_i \) is a filtered Poisson process with intensity \( \lambda_i(t) = \bar{\lambda}_i(1 + \theta_i(t)) \) under the measure \( \mathbb{Q}_i \) induced by \( \psi_i \).

Identity (4) implies that shifting the intensity of \( A_i \) from \( \bar{\lambda}_i \) to \( \lambda_i \) corresponds to changing the measure from \( \mathbb{P}_i \) to the induced measure \( \mathbb{Q}_i \). Notably, this relationship brings the induced measure \( \mathbb{Q}_i \) to the forefront, providing justification for using some notion of distance between the measures, \( \mathbb{P}_i \) and \( \mathbb{Q}_i \), to quantify the size of the perturbations \( \theta_i \). Correspondingly, limiting the size of \( \theta_i \) would mean controlling the distance between the two measures. In this paper, the distance between each \((\mathbb{P}_i, \mathbb{Q}_i)\) pair is evaluated via Rényi divergence.

In general, the Rényi divergence of a measure \( \tilde{\mathbb{P}} \) with respect to a reference measure \( \mathbb{P} \) of order \( \alpha \neq 1 \) can be defined as

\[
\mathcal{R}^\alpha(\tilde{\mathbb{P}} \| \mathbb{P}) := \frac{1}{\alpha - 1} \ln \int \left( \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \right)^\alpha d\mathbb{P} = \frac{1}{\alpha - 1} \ln \int \left( \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \right)^{\alpha - 1} d\tilde{\mathbb{P}}.
\]
As shown by Van Erven and Harremos (2014), given $\hat{\mathbb{P}}$ and $\mathbb{P}$ fixed, $\mathcal{R}^{\alpha}$ is continuous in $\alpha$ on $(0,1) \cup (1,\infty)$, and it tends to KL divergence as $\alpha$ approaches 1. This leads to defining the Rényi divergence of order $\alpha = 1$ as KL divergence, so one may assume $\alpha \in (0,\infty)$. It is also possible to demonstrate that $\mathcal{R}^{\alpha}(\hat{\mathbb{P}}||\mathbb{P})$ is convex in $\hat{\mathbb{P}}$ when $\alpha \in (0,1]$ and quasi-convex in $\hat{\mathbb{P}}$ when $\alpha \in (1,\infty)$. Let $\mathcal{R}^\alpha_i(t)$ denote the Rényi divergence of order $\alpha$ with respect to $\mathbb{P}_i$ on $\mathcal{F}_t$. The following result links $\mathcal{R}^\alpha_i(t)$ to $\theta_i$.

**Proposition 1.** For each fixed $t \geq 0$,

$$
\mathcal{R}^\alpha_i(t) = \frac{\bar{\lambda}_i}{\alpha - 1} \int_0^t \{(1 + \theta_i(u))^\alpha - \alpha \theta_i(u) - 1\} \, du,
$$

if $\alpha \neq 1$ and $\mathcal{R}^\alpha_i(t) = \bar{\lambda}_i \int_0^t \{(1 + \theta_i(u)) \ln(1 + \theta_i(u)) - \theta_i(u)\} \, du$ if $\alpha = 1$.

Based on this result, one can now put each induced measure $\mathbb{Q}_i$ back into the background and treat the explicit expression for $\mathcal{R}^\alpha_i(t)$ as its definition, allowing us to focus solely on the perturbation processes $\theta_i$ rather than the measures they induce.

We assume the decision maker restricts attention to stationary deterministic policies under which no queue can grow without a bound. Requiring queue lengths to be in a bounded region is innocuous and intuitive. To briefly explain, whether to outsource the production of a product depends on whether the cost of keeping the corresponding orders in the queue outweighs the cost of outsourcing some of these orders. Without a bound, the holding cost rate associated with a queue can approach infinity as the queue grows large. Thus, there will be a tipping point where holding on is no longer more cost-effective than outsourcing. This insight has been confirmed by our extensive numerical studies, including those shown in Figures 2–4; see also Figures EC.1–EC.3. The insight can also be gleaned from the solution to the approximating SDG (to be presented later on), which holds that allowing the amount of work in the system to grow to infinity is not ideal, implying that all queues should be in a bounded region under optimal min-max control.

We now have all the vocabulary we need to state the **constraint problem**. It is one in which the decision maker seeks some $(T, \Psi)$ to minimize

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

subject to $\liminf_{t \to \infty} \frac{1}{t} \mathbb{E}^{\theta} [\mathcal{R}^\alpha_i(t)] \leq \beta_i$ for $i = 1, \ldots, I$;

in the above, we have added a superscript $\theta$ to the expectation operator to emphasize the fact that the arrival intensities now follow $\bar{\lambda} + \theta$ rather than being given by $\bar{\lambda}$ in the nominal model.

The constraint problem involves $I$ model-error constraints, with $\beta_i$ measuring the extent to which the decision maker distrusts the law that governs class $i$ arrivals in the nominal model. To briefly
explain, a large value of $\beta_i$ allows $\lambda_i$ to deviate further from $\bar{\lambda}_i$, whereas a small value of $\beta_i$ forces $\lambda_i$ to stay close to its nominal value $\bar{\lambda}_i$. It is clear that when $\beta_i = 0$, $\lambda_i$ collapses to the constant function of $\bar{\lambda}_i$. In this respect, one may think of $\beta := (\beta_i)$ as a set of tuning parameters that the decision maker can use to add robustness to his control strategy by solving the constraint problem under some choice of $\beta$.

Alternatively, one may consider the following penalty problem wherein the decision maker’s objective is to seek some $(T, \Psi)$ to minimize

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

(5)

Each $\gamma_i$ in the penalty problem quantifies the decision maker’s level of confidence in the law (specified in the nominal model) that governs class $i$ arrivals: A large value of $\gamma_i$ penalizes nature severely for perturbations $\theta_i$, whereas a small value of $\gamma_i$ levies light penalties, incentivizing nature to take more drastic actions to inflate the objective value.

The set of parameters $\gamma := (\gamma_i)$ can also be viewed as the Lagrange multipliers associated with $I$ model-error constraints in the constraint problem. The simultaneous minimization and maximization suggest that the link between the constraint problem and the penalty problem will not be a direct consequence of the Lagrange multiplier theorem. Nonetheless, we can draw a connection between the two problems. The analysis is based on the theory underpinning the “convex analytic method”, which was first published in Bhatnagar and Borkar (1995) and further developed in the book by Altman (1999). Although the theory was originally intended for discrete-time models, the uniformization technique allows us to convert our continuous-time model to a discrete-time equivalent so that the theory can be applied. Let the optimal values of the constraint problem and penalty problem be denoted by $C^*_{\text{constraint}}(\beta)$ and $C^*_{\text{penalty}}(\gamma)$, respectively.

**Proposition 2.** For each fixed $\beta$, if $\gamma^* \succeq 0$ minimizes $C^*_{\text{penalty}}(\gamma) + \langle \beta, \gamma \rangle$, then $C^*_{\text{constraint}}(\beta) = C^*_{\text{penalty}}(\gamma^*) + \langle \beta, \gamma^* \rangle$. Furthermore, for the penalty problem, $C^*_{\text{penalty}}(\gamma^*)$ can be achieved when both players adopt stationary deterministic policies.

Proposition 2 shows that it is meaningful to consider the “Lagrangian relaxation” of the constraint problem. Hence, we will from now limit attention to the penalty problem. Utilizing the results from Proposition 1, we obtain the following alternative representation of (5):

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

(6)

where we have defined

$$
r(\theta) := r^\alpha(\theta) := \sum_{i=1}^I \frac{\gamma_i \bar{\lambda}_i}{\alpha - 1} \{(1 + \theta_i)^\alpha - \alpha \theta_i\}
$$

for $\alpha \neq 1$ and $\sum_{i=1}^I \gamma_i \bar{\lambda}_i \{(1 + \theta_i) \ln (1 + \theta_i) - \theta_i\}$ for $\alpha = 1$. 

5. Heavy-Traffic Analysis

Problem (6) remains complicated and suffers from the curse of dimensionality as the number of classes increases. Moreover, an exact analysis yields limited structural insights into decision-making. For these reasons, we advance and solve an SDG that is deemed more tractable than the original penalty problem it approximates.

5.1. SDG

The limiting regime we focus on is the one where both the demand volume and production capacity are large and the capacity balances the supply and demand. To be more specific, we impose the following heavy-traffic assumption:

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

(7)

Because the server’s long-run proportion of time spent on producing product \( i \) is \( \rho_i \), the system can be thought of as critically loaded if the nominal model is correct. Assume that nature employs \( \theta_i \) to generate the demand rate for product \( i \). Assuming optimistically that \( \theta_i \) is an order of magnitude smaller than \( \bar{\lambda}_i \), we can approximate \( A_i \) using

\[ A_i(t) = \bar{\lambda}_i t + \bar{\lambda}_i \int_{0}^{t} \theta_i(u) du + \hat{A}_i(t) + \epsilon_i^o(t), \]  

(8)

where \( \hat{A}_i \) is a Brownian motion with zero drift and variance parameter \( \bar{\lambda}_i \) and \( \epsilon_i^o \) is an approximation error term. Define, for each \( i \), the centered time allocation process as

\[ Y_i(t) := \rho_i t - T_i(t). \]

(10)

\[ U(t) := \sum_{i} Y_i(t) \text{ is non-decreasing with } U(0) = 0, \quad \text{and} \]

(11)

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

(12)

where \( \hat{Z}_i \) are independent Brownian motions with drift zero and infinitesimal variance \( \sigma_i^2 = 2\bar{\lambda}_i \).

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 $\hat{Ψ} := (\hat{Ψ}_i)$, we can formally state the decision maker’s problem as one that seeks an adapted control $(\hat{Y}, \hat{Ψ})$ that minimizes
\[
\max_\theta \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta \left[ \int_0^t \left( \sum_{i=1}^I c_i(\hat{Q}_i(u)) - r(\theta(u)) \right) \, du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\hat{ξ}_i(k)) \right]
\]
subject to constraints (10)−(12).

5.2. Dimensional Reduction
Although the SDG is simpler than the original problem it approximates, its solution is not as simple due to the high dimensionality of the state process $\hat{Q} := (\hat{Q}_i)$. For this reason, we seek further simplification of the problem, which eventually leads to a one-dimensional differential game, termed 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 serves as an approximation for the amount of work in the system at time $t$. To deduce the system equation of the workload process, we multiply (10) by $m_i$ and sum over $i$ to get
\[
W(t) = W(0) + B(t) + \int_0^t \zeta(u) \, du + U(t) - O(t),
\]
where we defined
\[
B(t) := \sum_{i=1}^I m_i \hat{Z}_i(t), \quad \zeta(t) := \sum_{i=1}^I \rho_i \theta_i(t) \quad \text{and} \quad O(t) := \sum_{i=1}^I m_i \sum_{k=0}^{N_i(t)} \hat{ξ}_i(k).
\]
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 $U(t)$ approximates the cumulative idle time up to $t$. Similarly, $O(t)$ approximates the cumulative amount of work outsourced up to $t$.

For the workload problem, we can define the effective holding cost rate function as
\[
h(w) = \min \left\{ \sum_{i=1}^I c_i(x_i) : m^\top x = w, x \in \mathbb{R}_+^I \right\}.
\]
This cost rate function has an intuitive interpretation: Given that the total workload $w$ can be instantly 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, we can define nature’s “cost rate function” as
\[
r^*(z) = \min \left\{ r(y) : \rho^\top y = z, y_i \in \Theta_i \right\}.
\]
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. For type $i$ outsourcing operations, we 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 can thus interpret $\tilde{\ell}_i$ as the proportional cost of outsourcing one unit of work through type $i$ outsourcing operations, and we denote by $\tilde{\Psi}$ the outsourcing rule for the workload process. Finally, by letting $\tilde{\xi}_i(k) := m_i \hat{\xi}_i(k)$ for $k \geq 0$ and $i = 1, \ldots, I$, we can spell out the workload problem, which states that the decision maker seeks some adaptive control $(U, \tilde{\Psi})$ to minimize

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

s.t. $W(t) = W(0) + B(t) + \int_0^t \zeta(u)du + U(t) - \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \tilde{\xi}_i(k)$, \hspace{1cm} (15)

$$U(t) \text{ is non-decreasing with } U(0) = 0, \quad \text{and}$$

$$W(t) \geq 0 \quad \text{for} \quad t \geq 0.$$ \hspace{1cm} (16)

where the superscript $\zeta$ means that there is a state-dependent perturbation $\zeta$ to the drift of the underlying process. To minimize technicalities, we assume that both the minimizing player and the maximizing player limit their attention to stationary deterministic policies.

5.3. Characterization of the Optimal Solution

Because nature’s drift-rate control only depends on the current workload, in this case, we will now write $\zeta(W(t))$ instead of $\zeta(t)$; for the decision maker, this means that an outsourcing rule would be in the form of a control limit (which we will briefly describe below). It is evident that a deviation from the work-conserving principle can only hurt the decision maker, so the idleness process $U$ ought to satisfy

$$\int_0^t 1_{\{W(u)>0\}}du = 0, \quad t \geq 0.$$ \hspace{1cm} (17)

5.3.1. Control Band Policy

Following Ormeci et al. (2008), we define a relevant class of control rules 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 $C^2[0, s]$ denote the space of functions that are twice differentiable up to the boundaries. Suppose that there exists some $\eta \in \mathbb{R}$ and $f \in C^2[0, s]$ that collectively satisfy

$$\Gamma_\zeta f(w) + h(w) - r^*(\zeta(w)) = \eta \quad \text{for} \quad w \in (0, s) \quad (18)$$

subject to the boundary conditions

$$f'(0) = 0 \quad \text{and} \quad f(s) = \tilde{\phi}_i(s - q) + f(q). \quad (19)$$

The following proposition provides a useful identity that motivates the optimality equation described in the following subsection.

**Proposition 3.** Suppose $\eta \in \mathbb{R}$ and $f \in C^2[0, s]$ jointly satisfy $(18)$ and $(19)$. 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

The analysis proceeds in three 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 strategy. Second, 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. In more formal terms, we select $i^*$ so that $\eta_{i^*} \leq \eta_i$ for all $i \neq i^*$, with the control band policy $(i^*, q_{i^*}, s_{i^*})$ viewed as a potential solution to the decision maker’s problem. Third, by exploiting the structural properties of the value function associated with the control band 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 3 motivates the following optimality equation that facilitates the identification of the control band 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

$$\max_\zeta \left\{ \frac{1}{2}\sigma^2 v''(w) + \zeta v'(w) + h(w) - r^*(\zeta) \right\} = \eta_i, \quad w \in (0, s_i), \quad (20)$$

subject to the boundary conditions

$$v'(0) = 0, \quad v(s_i) = \tilde{\phi}_i(s_i - q_i) + v(q_i), \quad \text{and} \quad v(w) = \tilde{\ell}_i(w - s_i) + v(s_i) \quad \text{for} \quad w > s_i, \quad (21)$$

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

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

$$g(x) := \max_\zeta \left\{ x \zeta - r^*(\zeta) \right\} \quad \text{for} \quad x \in \mathbb{R},$$
we can rewrite (20) as
\[ \frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) = \eta_i, \quad w \in (0, s_i). \] (22)

Because (22) does not involve the unknown function \( v \), it is in essence a first-order differential equation. This motivates us to consider the class of functions \( \{ \pi(\cdot, \eta); \eta \in \mathbb{R} \} \), where \( \pi(\cdot, \eta) \) solves
\[ \frac{1}{2} \sigma^2 \pi_w(w, \eta) + g(\pi(w, \eta)) + h(w) - \eta = 0 \] (23)
subject to the boundary condition
\[ \pi(0, \eta) = 0. \] (24)
The parameter pair \( (q_i, s_i) \) and the average cost \( \eta_i \) are determined through conditions:
\[ \pi(q_i, \eta_i) = \pi(s_i, \eta_i) = \tilde{t}_i \quad \text{and} \quad \int_{q_i}^{s_i} \pi(w, \eta_i) \, dw = \tilde{\phi}_i(s_i - q_i). \] (25)

It is noteworthy that (24) and (25) effectively make up four constraints, which alongside (23) should be sufficient to pin down four unknowns, i.e., \( q_i, s_i, \eta_i \) and \( \pi(\cdot, \eta_i) \).

Thus far, two questions remain unanswered. First, does the system of equations given in (23)–(25) yield a solution? Second, given that the answer to the first question is “yes”, does the control band 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? To answer these two questions, we need additional regularity conditions on the problem data, which we formally record below.

**Assumption 1.** \( h(w) \) is continuous and strictly increasing on \( (0, \infty) \) with \( h(0) = 0 \); moreover, \( \lim_{w \to \infty} h(w) = \infty \).

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

**Proposition 4.** Suppose Assumption 1 holds. Then the following statements are true. (i) The requirements in (23)–(25) uniquely determine \( q_i, s_i, \) and \( \eta_i \). (ii) If letting \( v \) denote the unique primitive function of \( \pi(\cdot, \eta_i) \), modulo an additive constant, then \( (v, \eta_i) \) satisfies the following quasi-variational inequality:
\[ \min \left\{ \frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_i, \inf_{z \geq 0} \left[ v(w - z) + \tilde{\phi}_i(z) \right] - v(w) \right\} = 0. \] (26)
(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.
In studying a joint two-sided impulse control and drift rate control problem, Cao and Yao (2018) arrive at a dynamic programming equation similar to (26). Thus, our detailed construction of the solution to (26) bears some similarities to that in Cao and Yao (2018). Because the problem in Cao and Yao (2018) involves impulse control on both sides, their analysis involves one additional degree of freedom, namely, the value of $v'$ at zero. Here, since we consider a “one-sided” impulse control, the left boundary of $v'$ is fixed at zero.

According to Proposition 4, for each fixed $i$, one can appeal to (23)–(25) to find a triple $(q_i, s_i, \eta_i)$, which leads to the best control rule given only type $i$ outsourcing operations can be used. Clearly, there exists a winning index $i^* \in \{1, \ldots, I\}$ such that $\eta_{i^*} \leq \eta_i$ for all $i \neq i^*$. Intuitively, the resulting control band policy $(i^*, s_{i^*}, \eta_{i^*})$ can be thought of as one that mini-maximizes the long-run average cost among all adaptive controls that can utilize only one type of outsourcing operation. Hence, at this stage, it is natural to wonder whether this rule mini-maximizes the long-run average cost among all adaptive controls that the decision maker can possibly take. The answer to it does not seem to be obvious, given that combining different outsourcing operations is likely to produce better results than relying solely on one type of outsourcing operation. Nonetheless, we establish that the control band policy $(i^*, s_{i^*}, \eta_{i^*})$ is overall optimal. The key is to ensure that the solution to (26) with $i = i^*$ also satisfies the variational inequality given in Theorem 1 below. The proof uses a level-set argument, which we believe carries some methodological novelty.

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

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

and (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**

We propose an implementable control rule based on the solution to the SDG. Here we take both $\alpha$ and $\gamma$ as given and fixed. Using the method described in §5.3.2, one can calculate $(i^*, s_{i^*}, \eta_{i^*})$ and $\pi(\cdot, \eta_{i^*})$. A procedure regarding how to select $\gamma$ for a given $\alpha$ based on historical data will be the focus of the Section 7.

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 $m_i$ the mean class $i$ service time. The proposed control has two components, as described below.

---

1 Previously, we have used $W$ to denote the approximate workload process.
Outsourcing. Whenever the workload 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. We mention that postponement can be accomplished by creating a virtual queue to hold current orders of class $i^*$ and routing new orders of this class to this virtual queue until the queue length reaches $(o_i^*)$, at which point an outsourcing operation is performed to deplete the virtual queue. As a result, from the time those orders are moved to the virtual queue until the outsourcing operation occurs, the $i^*$th queue is considered empty.

Sequencing. To fix ideas, we stipulate that there is a unique solution $(x_i^*)$ to the optimization problem (13) for each fixed $w$. (Note that this stipulation merely seeks to mitigate the potential technical complexity and will be easily satisfied in various settings.) The solution $(x_i^*)$ can then be regarded as a function of $w$ and serves as the target length of the queues when the workload is at position $w$. Hence, a desired sequencing rule 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 can recover the well-known $c \mu$ priority rule. When all $c_i(\cdot)$ are strictly convex and satisfy $c_i(0) = c_i'(0) = 0$, the generalized $c \mu$ rule emerges, which states that service priority is given to the job class whose $c_i'(Q_i(t)) \mu_i$ index is the largest at time $t$.

6. Discussion
We comment on some key aspects of the modeling framework proposed in this paper.

6.1. The Choice of Discrepancy Measure and General Service Times
We have demonstrated how the use of Rényi divergence can give rise to a versatile family of uncertainty sets. In those cases, model uncertainty is conveniently represented by the single function $r$, which determines how nature will be penalized based on her actions. An important point to note is that this way of representing model uncertainty does not require the penalty rate function to be in a specific form. In principle, the decision maker can specify any penalty form he wants, as long as the resulting uncertainty set is thought to capture model misspecification concerns.

Let $p(\cdot)$ denote a penalty function that maps from $\mathbb{R}^I$ to $\mathbb{R}_{+}$, so that when nature picks $\theta(t)$ at time $t$, it makes a penalty 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

$$\max_{T, \Psi} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta \left[ \int_0^t \left( \sum_{i=1}^I c_i(Q_i(u)) - p(\theta(u)) \right) du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) \right].$$
Of course, upon replacing \( p \) with \( r \), one recovers problem (6). Paralleling the development in §5.1, we arrive at an approximating SDG in which the decision maker chooses \((\hat{Y}, \hat{\Psi})\) to minimize

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

subject to constraints (10) – (12).

Arguing along the lines that are similar to those in §5.2, we can obtain the corresponding workload problem, in which the decision maker seeks an adaptive control \((U, \tilde{\Psi})\) to minimize

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

subject to constraints (15) – (17),

where \( p^*(z) = \min \{ p(y) : \rho^\top y = z, y_i \in \Theta_i \} \). At this point, we would like to emphasize that only two properties of \( r^* \) are crucial to the proofs of Proposition 4 and Theorem 1: (i) \( r^* \) attains its minimum value at \( z = 0 \) with \( r^*(0) = 0 \); and (ii) its convex conjugate is non-negative and Lipschitz continuous. As a result, all the analytical results established in the previous section apply to all \( p^* \) that possess these two properties.

Thus far, we have assumed service times to be exponentially distributed. As far as heavy-traffic analysis is concerned, this assumption can be easily relaxed to allow for general service time distributions without affecting the main results established in Section 5. Indeed, with general service times, the process \( S_i(t) \), which represents the number of class \( i \) products manufactured over time if the server was constantly working on class \( i \) orders, can be viewed as a renewal process with cycles having mean \( m_i \) and coefficient of variation \( \nu_i \). As a result, \( \hat{S}_i \) in (9) becomes a Brownian motion with zero drift and variance parameter \( \bar{\lambda}_i \nu_i^2 \), whereas \( \hat{Z}_i \) in (10) becomes a Brownian motion with zero drift and infinitesimal variance \( \sigma_i^2 = \bar{\lambda}_i(1 + \nu_i^2) \).

6.2. The Proposed Control Rule

Our proposed outsourcing rule allows the decision maker to postpone outsourcing if the workload reaches \( s_{i^*} \) while there are insufficient outsourceable orders in the queue—outsourcing will occur only when enough class \( i^* \) orders accumulate. This is the simplest and perhaps the most faithful interpretation of the solution derived in §5.3. When \( s_{i^*} \) is small, one expects such postponements to occur only infrequently. When \( s_{i^*} \) is large, one would expect postponements to happen regularly. Long postponements can defeat the purpose of reducing an excessive amount of work in the system, because orders in the virtual queue still accrue holding costs. Fortunately, when the system operates in the assumed heavy-traffic regime, long postponements can be very rare. This is because in heavy traffic, queue lengths are an order of magnitude smaller than the demand volume. In other
words, stochastic fluctuations have a second-order effect when compared with demand volume and processing capacity, implying that relevant control parameters such as $q_i$ and $s_i$ should be an order of magnitude smaller than $\bar{\lambda}_i$s. (Since control actions are designed to mitigate the negative effects of stochastic fluctuations, they ought to be of the same order of magnitude.) Therefore, as the demand volume grows, the time taken to accumulate sufficient outsourceable orders becomes negligible.

For the SDG, work can be redistributed instantly among various classes. Obviously, in practical situations, an instant redistribution of work is infeasible. Nonetheless, we can argue that any desired corrections to the queues with respect to their respective targets can be made relatively quickly. The intuition can be explained as follows: Classes whose actual queue lengths fall short of the targets will temporarily lose access to the server, thereby facing only an inflow to the queues (due to demand arrivals). Because demand rates are an order of magnitude greater than the queue lengths, the actual queue lengths of these classes will quickly return to the desired targets. By the same token, classes whose actual queue lengths exceed the targets exclusively enjoy total service capacity, hence seeing “an underloaded system” due to the heavy-traffic condition (7) and the fact that nature’s distortions are moderate. Because the outflow from these queues (due to service completions) greatly surpasses the inflow (due to demand arrivals), the queue lengths of these classes will quickly return to the desired targets as well.

### 6.3. Selecting the Uncertainty Set

One of the primary reasons for considering robust control is to find a good balance between tractability and practicability. Considerations of tractability often necessitate the adoption of a simple model. However, such a model is likely to contain false or oversimplified assumptions, causing the resulting decision action to deviate significantly from the optimal course. On the other hand, considerations of practicability tend to favor sophisticated models with solutions that are often analytically intractable and difficult to compute. A robust control formulation mitigates the tension by taking a simple model as the nominal model and building an uncertainty set around it. The choice of an appropriate uncertainty set is thus central. A too-small uncertainty set may exclude the true model, making a robust formulation lose its appeal. By contrast, an exceedingly large uncertainty set can lead to a solution that is too conservative, resulting in a loss of advantages over its non-robust counterpart.

It is important to note that our modeling framework encompasses a class of SDGs, which are parameterized by the vector $\gamma$ (assuming the parameter $\alpha$ is held fixed). Thus, in our context, choosing an uncertainty set boils down to selecting an appropriate $\gamma$. Each of these SDGs is unique, and therefore, may lead to a distinct control rule for the decision maker. Notably, all of these
control rules are robust control rules, and their effectiveness can be evaluated using computer simulation, with all relevant distributions based on historical data. Therefore, the best uncertainty set corresponds to the $\gamma$ vector for which the robust control rule produces the lowest cost (as estimated by simulation output). In the following section, we propose a data-driven method for locating the “best” $\gamma$ vector.

7. A Data-Driven Method for Uncertainty Set Selection

In this section, we demonstrate how to utilize historical data to select $\gamma$. Here, we reduce the importance of the “shape” parameter $\alpha$ by treating it as fixed. (Of course, changing the value of $\alpha$ allows for greater flexibility in designing the uncertainty set.)

Historical data can be used to calibrate a true model that is thought to govern realized demand; the true model, though potentially complicated, oftentimes permits a simplification that serves as the nominal model to be used in the formulation of the robust control problem. A well-specified nominal model combined with a vector $\gamma$ yields a Bellman-Isaacs equation that can be solved to obtain a robust control rule, denoted as $P(\gamma)$. When such a control rule is applied to the true model, it gives rise to a long-run average cost, denoted by $C(\gamma)$. Choosing the best uncertainty set is thus a matter of identifying some vector $\gamma$ that minimizes $C(\gamma)$. When the expression for $C(\gamma)$ is known, it is natural to consider using a gradient-decent algorithm. On the negative side, finding such an expression can be notoriously difficult. On the plus side, the function value of $C$ can be efficiently evaluated through computer simulations. This motivates our proposed algorithm, which is similar in spirit to a standard gradient-decent algorithm (Boyd et al. 2004, Chapter 9) but substitutes finite difference for gradient (Spall 2005, Chapter 6). For this reason, we refer to it as the “quasi-gradient descent” algorithm.

Algorithm 1 shows how to find the location of the best $\gamma$ value for a fixed $\alpha$. In the algorithm, $\kappa$ is so-called the learning rate. $\gamma^{(j)} := \{\gamma_1^{(j)}, \ldots, \gamma_I^{(j)}\}$ is understood to be the $\gamma$ value at the end of the $j$th outer-iteration, and $\hat{C}(\gamma)$ denotes the simulation estimate of $C(\gamma)$. By choosing a small $\delta \in \mathbb{R}^+$, we use

$$\nabla C_i(\gamma^{(j)}) := \frac{\hat{C}(\gamma^{(j)} + \delta e_i) - \hat{C}(\gamma^{(j)} - \delta e_i)}{2\delta}$$

to approximate the $i$th component of the gradient of $C$ at the point $\gamma^{(j)}$, where $e_i$ denotes the unit vector whose $i$th component is one and rest components zero.

In the following, we provide an example to demonstrate the efficiency of this method. Consider a make-to-order system that produces two different types of products. Service rates are given by $\mu_1 = 30$ and $\mu_2 = 90$; thus, $m_1 = 1/30$ and $m_2 = 1/90$. The actual demand rate of product 1 (resp. product 2) follows an auto-regressive integrated moving average (ARIMA) model with mean $\bar{\lambda}_1 = 20$ (resp. mean $\bar{\lambda}_2 = 30$). For both products, we set the two ‘ARLags’ to 0.3 and −0.3
Algorithm 1 Quasi-Gradient Descent

1: initialize: $\kappa$, $\gamma(0)$, $\epsilon$
2: while $\gamma^{(j+1)} \neq \gamma^{(j)}$ do
3: $\tilde{\gamma}^{(j)} \leftarrow \gamma^{(j)}$
4: for $i = 1 : I$ do
5: while $|\nabla C_i(\tilde{\gamma}^{(j)})| > \epsilon$ do
6: $V^{(j)} \leftarrow \tilde{\gamma}^{(j)}$
7: $\tilde{\gamma}_i^{(j)} \leftarrow \max \left\{ \tilde{\gamma}_i^{(j)} - \kappa \nabla C_i(\tilde{\gamma}^{(j)}), 1 \right\}$ \hspace{1cm} $\triangleright$ Maximum ensures positiveness.
8: $\kappa \leftarrow \frac{\left| (v_i^{(j)} - \tilde{\gamma}_i^{(j)}) (\nabla C_i(v^{(j)}) - \nabla C_i(\tilde{\gamma}^{(j)})) \right|}{(\nabla C_i(v^{(j)}) - \nabla C_i(\tilde{\gamma}^{(j)}))^2}$
9: end while
10: end for
11: $\gamma^{(j+1)} \leftarrow \{\tilde{\gamma}_1^{(j)}, \ldots, \tilde{\gamma}_I^{(j)}\}$
12: end while

respectively, the MA error coefficient to 0.4 and the variance to 1. It is straightforward to verify that $\tilde{\lambda}_1/\mu_1 + \tilde{\lambda}_2/\mu_2 = 1$, so that the heavy-traffic condition holds. Also, a direct calculation gives $\sigma^2 = 0.0519$. We also set $\alpha = 2$. The cost data includes the fixed outsourcing cost parameters, $L_1 = 8, L_2 = 6$, the proportional outsourcing cost parameters, $l_1 = 1.5, l_2 = 2$, and two quadratic holding cost rate functions

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

where we choose $a_1 = 0.1$ and $a_2 = 0.2$.

The algorithm, as shown in Figure 1, approaches the minimum of $\hat{C}(\gamma)$ much faster than the exhaustive method. A desired $\gamma$ vector can be reached in just a few steps. In addition, the algorithm is computationally light-weight, making it suitable for practical use.

To see that our method of uncertainty set selection conforms to the general philosophy underlying many techniques in data-driven modeling and machine learning, it is instructive to draw useful parallels between our decision framework and regularized (or penalized) regression. To use regularized regression, one needs to specify the norm for the penalty term (L1 norm, L2 norm, etc.) and the regularization parameter (a key parameter that controls the strength of shrinkage and variable selection), two decisions that are analogous to selecting $\alpha$ and $\gamma$. To find the best regularization parameter (for a given norm), one often resorts to cross-validation, a popular data-driven technique that parallels our simulation-based evaluation.
8. Numerical Studies

This section consists of two main parts. In §8.1, we contrast the decision maker’s strategy derived from the original penalty problem with that of the SDG. In §8.2, we provide examples to demonstrate the value of incorporating robustness. Specifically, §8.2.1 assumes the demand rate of each product to follow an ARIMA model, whereas §8.2.2 assumes the demand rate of each product to follow a Cox–Ingersoll–Ross (CIR) process.

8.1. Solution Comparison Between the Penalty Problem and the SDG

In this subsection, we numerically solve the penalty problem using value iteration and the SDG via the solution to the associated Bellman-Isaacs equation. We compare their solutions to demonstrate that the heavy-traffic approximation is accurate and reliable. Section EC.1 contains a description of the state-descriptor for the penalty problem. We consider a two-class system with parameters set as follows. Let $\lambda_1 = 30$, $\lambda_2 = 40$, $\mu_1 = 60$, and $\mu_2 = 80$. The cost data includes the fixed outsourcing cost parameters $L_1 = 5$ and $L_2 = 15$, the proportional outsourcing cost parameters $\ell_1 = \ell_2 = 1.0$, and two quadratic holding cost rate $a_1 = 0.4$ and $a_2 = 0.5$.

For ease of presentation, we make a slight modification to the original model by allowing preemptive sequencing, which eliminates the need to track which class is in service. Hence, which class is in service becomes part of the decision rather than a component of the state-descriptor. Appendix EC.4 contains results for the case in which preemptive sequencing is prohibited. Note that with non-preemptive sequencing, a control rule (particularly the outsourcing rule) should be dependent on which class is in service.
Figure 2  The decision maker’s strategies derived from the penalty problem and the SDG with $\gamma_1 = \gamma_2 = 30$ when $\alpha = 2$

Figure 3  The decision maker’s strategies derived from the penalty problem and the SDG with $\gamma_1 = \gamma_2 = 30$ when $\alpha = 1$

The left panels of Figures 2, 3 and 4 depict the optimal sequencing rules obtained by solving the penalty problem and the SDG under KL ($\alpha = 1$) and Rényi ($\alpha = 1/2$ and 2) divergence. For the
original penalty problem, if the state is designated blue, the decision maker would prioritize class 1, and if the state is marked green, the decision maker would favor producing product 2. For the SDG, the decision maker should prioritize class 2 if the state is above the grey line, and class 1 if the state is below the grey line. We can see that the sequencing rule derived from the SDG is very close to that obtained by solving the penalty problem. This suggests that the generalized $c\mu$ rule is near-optimal for the original penalty problem.

The right panels of Figures 2, 3 and 4 depict the optimal outsourcing rules. For the original penalty problem, the blue squares represent the states where product 1 should be outsourced, whereas the orange squares represent the states where product 2 should be outsourced. From these plots, we can see that the system is inclined to outsource production of product 1, unless queue 2 becomes exceedingly large. This observation is in alignment with the control strategy prescribed by the SDG. In addition, the green lines represent the outsourcing rule parameters, $q_1$ and $s_1$, derived from the SDG. Hence, we can conclude that the outsourcing rule derived from the SDG, which always outsources product 1, is reasonably close to that obtained by solving the original penalty problem. We can see that the value of the parameter $\alpha$ also influences the optimal outsourcing rule. For example, some states that will trigger outsourcing under KL divergence ($\alpha = 1$) may not trigger outsourcing under the Rényi divergence of order 2.

Finally, Figure 5 depicts the long-run average cost $\eta$ generated by both the original penalty problem ($\eta_{exact}$) and the SDG ($\eta_{SDG}$) at different levels of model uncertainty. We notice that the
long-run average costs derived from the SDG are quite close to the original penalty problem values, implying that the heavy traffic approximation and dimensional reduction method are accurate. Details of such costs are displayed in Tables EC.1, EC.2 and EC.3 in Appendix EC.4.

8.2. Exposing the Value of Robustness

In this subsection, we demonstrate the value of building robustness into decision-making through two numerical examples. All simulation estimates are based on 100 i.i.d. replications of the underlying stochastic system 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.

8.2.1. ARIMA Intensity In the first example, we assume that the real-world demand model is one in which the demand rate of each product follows a non-homogeneous Poisson process with an ARIMA model for intensity. The system parameters are identical to those described in Section 7.

The quadratic holding costs give rise to a queue-ratio type rule: If $a_1\mu_1Q_1 > (<) a_2\mu_2Q_2$, give priority to producing product 1 (product 2) with the tie broken in an arbitrary fashion. For the outsourcing rule, we find that in this example, we should always outsource product 1 because $\eta_1 < \eta_2$ for all pairs of $(\gamma_1, \gamma_2)$. Details of optimal control band parameters $(q_i, s_i)$ and “optimal” long-run average cost $\eta_i$ can be found in Table EC.4 in Section EC.4.

The values of the simulated average costs $\hat{C}(\gamma)$ in this example are shown in Figure 6. We can clearly observe 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 worth emphasizing that for very large $\gamma_1$ and $\gamma_2$ (hence very small $1/\gamma_1$ and $1/\gamma_2$), the corresponding control rule can be thought of as one obtained by completely ignoring model uncertainty. This means that we can regard the difference between the “limiting value” of each plot and the minimum value on that plot as the value of robustness. For example, when $\alpha = 2$, the value of robustness is roughly $10.404 - 9.124 = 1.280$, which is approximately 10.49% improvement relative to completely ignoring model uncertainty. Here we conduct “exhaustive research” on $\gamma$ rather than using our quasi-gradient descent algorithm, because we would like to fully expose the influence of uncertainty set selection on the actual performance of the proposed control rules. Of course, using our quasi-gradient descent algorithm can also locate a “minimum value”. We find that this value is extremely close to the lowest cost reported here, with a gap of less than 1%.

8.2.2. CIR Intensity  In this example, we assume that the real-world demand model is one in which the demand rate of each product is assumed to follow a non-homogeneous Poisson process whose intensity is a CIR process. Specifically, $\lambda_i(t)$ is the solution to the following stochastic differential equation

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

(28)

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 $\sigma_i$ measures the model
Figure 7: Average costs estimated from simulation (CIR) for different choices of \((\gamma_1, \gamma_2)\) with \(\alpha = 1/2, 1\) and 2 volatility. In the present example, we choose \(\bar{\lambda}_1 = 20, \bar{\lambda}_2 = 30, k_1 = k_2 = 1\) and \(\bar{\sigma}_1 = \bar{\sigma}_2 = 2\), and we can generate sample paths of (28) via the Milstein scheme.

The optimal outsourcing rule in this example is also to outsource product 1, and the optimal sequencing rule is the generalized \(c \mu\) rule. Figure 7 depicts that the estimated average cost again first decreases and then increases in \(\gamma_1\) (resp. \(\gamma_2\)) with \(\gamma_2\) (resp. \(\gamma_1\)) fixed for both \(\alpha\). This time, the value of robustness for \(\alpha = 2\) is approximately \(9.743 - 8.721 = 1.022\), which is 11.72% improvement over fully ignoring model uncertainty. Tables EC.5 and EC.6 in Appendix EC.4 provide the confidence intervals obtained from computer simulations of all the point estimates shown in Figures 6 and 7.

The main takeaway from Figures 6 and 7 is that employing a robust control framework that accounts for ambiguity can result in significant cost savings. The decision maker who follows the policy obtained from the nominal model (hence ignoring model uncertainty entirely) will get penalized for failing to adequately prepare for model misspecification. Overemphasizing the worst-case scenario, as is the case where the uncertainty set is chosen too large, would result in an overly conservative solution. A proper level of ambiguity (i.e., an appropriate choice of the uncertainty set) can guard against model errors without producing overly conservative solutions.

9. Concluding Remarks

This paper studies the joint order outsourcing and sequencing of a multiclass make-to-order manufacturing system with model uncertainty and fixed plus proportional costs for outsourcing. Model uncertainty is captured through the notion of Rényi divergence, extending the commonly used
entropic approach in the literature. We present two robust control formulations, both involving a second player that promotes robustness to model misspecification. Both formulations can be interpreted as two-player zero-sum stochastic games. By considering the system in a suitable heavy-traffic regime, we derive and solve an approximate SDG whose state-descriptor is driven by Brownian motions. We demonstrate that the SDG can be further turned into a one-dimensional stochastic game whose state-descriptor is the workload process. The optimal control for the decision maker is shown to be in a control-band form, whereas the optimal control for nature is a state-dependent drift-rate control of the workload process that serves to resolve ambiguity aversion. Combining the solution to the SDG and a simulation program fitted to the data, we propose a data-driven procedure that can generate a family of joint outsourcing and sequencing rules. The effectiveness of the procedure is demonstrated through numerical studies.

Although the paper focuses on a make-to-order manufacturing system, the methodologies developed in this paper may be adapted to other production strategies, such as make-to-stock manufacturing systems that allow inventories. Moreover, our approach to describing the distance between models may be well-suited for other examples of stochastic dynamic programming problems. These are interesting directions worth pursuing in future research.

References


E-Companion

This e-companion is organized into five parts. Section EC.1 gives the proof of Proposition 2. Section EC.2 collects proofs of other main results. Section EC.3 describes our numerical scheme used for solving the Bellman-Isaacs equation. Section EC.4 supplies additional data on the numerical experiments. Section EC.5 presents proofs for some auxiliary results.

EC.1. Proof of Proposition 2

Throughout the proof, we will treat $\lambda$ rather than $\theta$ as nature’s decision process. (This is done without loss of generality since the two quantities determine each other.) Note that nature has a bounded action space given as $\Lambda := \{\lambda \in \mathbb{R}^I : \lambda_i \in [\bar{\lambda}_i(1 + a_i), \bar{\lambda}_i(1 + b_i)]\}$. The proof involves three major steps. Step 1 takes the decision maker’s strategy as given and reformulates the nature’s decision problem as a discrete-time constrained Markov decision process (CMDP). Step 2 establishes a Lagrange multiplier theorem for the CMDP. Step 3 establishes the desired connection between the constraint problem and the penalty problem.

**Step 1:** Let the decision maker’s strategy be given and fixed. If $\lambda$ is fixed, the system evolution can be described by the stochastic process $X(t) := (Q(t), J(t))$, where $Q(t)$ is the vector of queue length processes and $J(t)$ is a process taking values from $\{0, 1, \ldots, I\}$; in particular, $J(t) = j$ indicates that a class $j$ job is in service at time $t$ if $j \neq 0$, whereas $J(t) = 0$ indicates that the server is idle at time $t$. Note that the decision maker’s strategy not only determines the state space of $X$, denoted as $S$, but also allows for the division of $S$ into two sets: $\tilde{S}$, which includes all the states that do not trigger outsourcing, and $\bar{S}$, which includes all the states that will trigger outsourcing. Clearly,

$$\tilde{S} \cap \bar{S} = \emptyset \quad \text{and} \quad \tilde{S} \cup \bar{S} = S.$$ 

Moreover, for each $x \in \bar{S}$, there is a pair $(\tilde{k}(x), \tilde{\delta}(x))$ stating that whenever $X$ reaches the state $x$, outsource $\tilde{\delta}(x)$ units of product $\tilde{k}(x)$. For $i = 1, \ldots, I$, denote by $e_i \in \mathbb{R}^I$ the unit vector whose $i$th component is one and remaining components zero. Then, for each $x := (q, j) \in \bar{S}$, we can define two sets:

$$\chi_0(x) := \{i : (q + e_i, j) \notin \tilde{S}\} \quad \text{and} \quad \chi_1(x) := \{i : (q + e_i, j) \in \bar{S}\}.$$

Intuitively, $\chi_0(x)$ collects the indices of job classes for which a new arrival will not trigger outsourcing, whereas $\chi_1(x)$ gathers the indices of job classes for which a new arrival will trigger outsourcing, given that the current system state is $x$. Henceforth, we will simply write $\chi_0$ and $\chi_1$ in place of $\chi_0(x)$ and $\chi_1(x)$, respectively, whenever the dependence on $x$ is clear from the context.

Using the standard uniformization technique, we can construct a discrete-time equivalent of the continuous-time process $X$, denoted as $X(n) = (Q(n), J(n))$. It is worth noting that since
outsourcing happens instantly, the state space of \( \{X(n)\} \) is effectively \( \tilde{S} \). The transition law of this
discrete-time process can be specified as follows: If \( X(n) = (q, j) \) for \( j \neq 0 \), then (i) with probability \( \frac{\lambda_i(n)}{\nu} \),

\[
X(n+1) = (q + e_i, j) \quad \text{if} \quad i \in \chi_0
\]

and

\[
X(n+1) = (q + e_i - \tilde{\delta}(q + e_i, j)e_{k(q+e_i,j)}, j) \quad \text{if} \quad i \in \chi_1;
\]

(ii) with probability \( \frac{\mu_j}{\nu} \),

\[
X(n+1) = (q - e_j, \tilde{j}(q - e_j)),
\]

where \( \tilde{j} \) is determined by the specific sequencing rule chosen by the decision maker; and (iii) with
probability \( 1 - \frac{\mu_j}{\nu} \sum_{i=1}^{I} \lambda_i(n) \),

\[
X(n+1) = (q, j).
\]

In the above, the constant \( \nu \) can be chosen arbitrarily as long as it is large enough to make
the aforementioned probabilities well-defined, and the existence of such a \( \nu \) is ensured by the
boundedness of nature’s action space. Also keep in mind that \( \lambda_i(n) \) are nature’s decision variables
at stage \( n \).

It is straightforward to check that the discrete-time system satisfies the so-called Weak Accessibility condition (Definition 4.2.2. in Bertsekas (1995)). Thus, if ignoring the model-error constraints for now, the Bellman equation characterizing nature’s best actions admits the following
form (Proposition 4.2.3. in Bertsekas (1995)):

\[
\frac{\eta^*}{\nu} + \phi^*(q, j) = \max_{\lambda \in \Lambda} \left[ c(q) + \sum_{i \in \chi_0} \frac{\lambda_i}{\nu} \phi^*(q + e_i, j) + \sum_{i \in \chi_1} \frac{\lambda_i}{\nu} \phi^* \left( q + e_i - \tilde{\delta}(q + e_i, j)e_{k(q+e_i,j)}, j \right) + \frac{\mu_j}{\nu} \phi^* \left( q - e_j, \tilde{j}(q - e_j) \right) 1_{\{j\neq 0\}} + \left( 1 - \frac{\mu_j}{\nu} \sum_{i=1}^{I} \lambda_i \right) \phi^*(q, j) \right] + \sum_{i \in \chi_1} \frac{\lambda_i}{\nu} \left[ L_{k(q+e_i,j)} + \ell_{k(q+e_i,j)} \tilde{\delta}(q + e_i, j) \right] \quad \text{for all} \quad (q, j) \in \tilde{S},
\]

where \( c(q) := \frac{1}{\nu} \sum_i c_i(q_i) \). The Bellman equation implies that the outsourcing cost can be absorbed
into the unit cost, yielding an effective unit cost function:

\[
\tilde{c}(x, \lambda) := c(q) + \sum_{i \in \chi_1} \frac{\lambda_i}{\nu} \left[ L_{k(q+e_i,j)} + \ell_{k(q+e_i,j)} \tilde{\delta}(q + e_i, j) \right].
\]

Therefore, nature’s problem can be cast into a CMDP that seeks \( \lambda \) to maximize

\[
C_{ea}(X, \lambda) := \lim_{m \to \infty} \frac{1}{m} \mathbb{E} \left[ \sum_{n=1}^{m} \tilde{c}(X(n), \lambda(n)) \right]
\]
subject to
\[
D_i(\lambda) := \liminf_{m \to \infty} \frac{1}{m} \mathbb{E} \left[ \sum_{n=1}^{m} d^\alpha_i(\lambda(n)) \right] \leq \beta_i, \quad i = 1, \ldots, I, \tag{EC.1}
\]
where
\[
d^\alpha_i(\lambda) := \begin{cases} \bar{\lambda}_i \nu (\alpha - 1) & \text{for } \alpha \neq 1, \\ \frac{1}{\nu} \left\{ \lambda_i \ln (\lambda_i - \lambda_i) - \lambda_i + \bar{\lambda}_i \right\} & \text{for } \alpha = 1. \end{cases}
\]

In the following, we will refer to \(D\) as the set containing all admissible \(\lambda\) satisfying (EC.1). We will also use \(\Phi_S\) and \(\Phi_D\) to represent the set of stationary \(\lambda\)s and the set of stationary deterministic \(\lambda\)s.

**Step 2:** Recall that \(\beta := (\beta_i)\) and \(\gamma := (\gamma_i)\) are \(I\)-dimensional vectors of real numbers. Now, write \(D(\lambda) := (D_i(\lambda))\). Our main task is to establish
\[
\sup_{\lambda \in \mathcal{D}} C_{ea}(X, \lambda) = \sup_{\lambda \in \Phi_S} \inf_{\gamma \succeq 0} J^\gamma_{ea}(X, \lambda) = \inf_{\gamma \succeq 0} \sup_{\lambda \in \Phi_D} J^\gamma_{ea}(X, \lambda), \tag{EC.2}
\]
where \(J^\gamma_{ea}(X, \lambda) := C_{ea}(X, \lambda) - \langle \gamma, D(\lambda) - \beta \rangle\). Identity (EC.2) will follow from Theorem 12.7 in (Altman 1999, Chapter 12), if one can verify two conditions, referred to by Altman (1999) as the moment condition and the boundedness condition. In particular, the moment condition, which corresponds to “the near-monotonic case” in (Feinberg and Shwartz 2012, Chapter 11) ensures that \(\Phi_S\) is a dominating class among all admissible \(\lambda\)s, justifying the first equality in (EC.2). The second equality in (EC.2) follows from that \(\Phi_S\) is a convex set; see, e.g, Lemma 11.2 in (Feinberg and Shwartz 2012, Chapter 11). The last equality in (EC.2), which also appears in part (iii) of Theorem 12.7 in (Altman 1999, Chapter 12), holds because the relaxed (unconstrained) problem can be handled by solving a system of dynamic programming equations, from which one recovers a deterministic policy. We next state and verify the two conditions in turn. (Note that the statements given below are slightly different from but essentially the same as those in Altman (1999), because Altman (1999) considers a minimization problem whereas nature faces a maximization problem.)

**Condition 1** *(Moment Condition; Condition 11.21 in Altman (1999)):* For all \(\bar{z} \in \mathbb{R}\),
\[
\forall \bar{z} \in \mathbb{R}, \text{ the set } \left\{ x \in \tilde{S} : \sup_{\lambda \in \Lambda} \tilde{c}(x, \lambda) > \bar{z} \right\} \text{ is finite.} \tag{EC.3}
\]

Since \(\tilde{S}\) is a finite set, (EC.3) is trivially satisfied.

**Condition 2** *(Boundedness Condition; Condition 11.1 in Altman (1999)):* \(\tilde{c}\) is bounded from above and for each \(i = 1, 2, \cdots I\), \(d^\alpha_i\) is bounded from below.

Since \(\tilde{S}\) is a finite set and the demand rate of each product is restricted to a bounded region, we know that \(\tilde{c}\) is bounded from above. In addition, we see that for all \(\alpha\),
\[
\inf_{\lambda \in \mathcal{D}} d^\alpha_i(\lambda) = d^\alpha_i(\bar{\lambda}) = 0, \tag{EC.4}
\]
from which we can conclude each $d_i^\alpha$ is bounded from below.

Having verified the two conditions, by Theorem 12.7 in Altman (1999) we conclude that (EC.2) holds (with the decision maker’s strategy being fixed).

**Step 3:** It is easy to see that the discrete-time equivalent of the constraint problem can be described as one where the decision maker seeks $(T, \Psi)$ to minimize $\max_{\lambda \in \mathcal{D}} C_{ca}(X, \lambda)$, where, to avoid introducing new notation, we continue to use $(T, \Psi)$ to represent the decision maker’s strategy. Similarly, the discrete-time equivalent of the penalty problem can be described as one where the decision maker seeks $(T, \Psi)$ to minimize

$$\max_{\lambda \in \Phi} C_{ca}(X, \lambda) - \langle D(\lambda), \gamma \rangle.$$  

With a slight abuse of notation, let the optimal values of the discrete-time constraint problem and penalty problem be denoted by $C^*_{\text{constraint}}(\beta)$ and $C^*_{\text{penalty}}(\gamma)$, respectively. We have

$$C^*_{\text{constraint}}(\beta) := \min_{(T, \Psi)} \max_{\lambda \in \mathcal{D}} C_{ca}(X, \lambda)$$  

(EC.5)

$$\overset{(a)}{=} \min_{(T, \Psi)} \min_{\gamma \geq 0} \max_{\lambda \in \Phi} J^*_{ca}(X, \lambda) = \min_{\gamma \geq 0} \max_{(T, \Psi)} J^*_{ca}(X, \lambda)$$  

(EC.6)

$$\overset{(b)}{=} \min_{\gamma \geq 0} \left[ C^*_{\text{penalty}}(\gamma) + \langle \beta, \gamma \rangle \right],$$  

(EC.7)

where step (a) is due to (EC.2) and step (b) follows by the definition of $C^*_{\text{penalty}}(\gamma)$.

**EC.2. Proofs of Other Main Results**

This part of the e-companion gives proofs for Proposition 1, Proposition 4 and Theorem 1. The proof of Proposition 3 is elementary, so it is omitted.

**Proof of Proposition 1.** We prove the result for all $\theta_i$ that are locally integrable. Because the result for $\alpha = 1$ is known, we restrict attention to cases where $\alpha \neq 1$. To start, Let $\bar{\alpha} := \alpha - 1$. Direct calculation gives

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

(EC.8)

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

$$\exp\left\{ \bar{\alpha} \int_0^t \ln(1 + \theta_i(u))dA_i(u) \right\} = \lim_{m \to \infty} \exp\left\{ \sum_{k} \bar{\alpha} \ln(1 + \theta_i(u_k))(A_i(u_{k+1}) - A_i(u_k)) \right\},$$

EC4
where the limit is in probability and taken as \( \Delta := \max_{k} |u_{k+1} - u_k| \to 0 \). Fixing \( \{u_k\} \),

\[
\mathbb{E}^{Q_i} \left[ \exp \left\{ \sum_k \tilde{\alpha} \ln(1 + \theta_i(u_k))(A_i(u_{k+1}) - A_i(u_k)) \right\} \right] \\
\overset{(a)}{=} \prod_k \mathbb{E}^{Q_i} \left[ \exp \left\{ \tilde{\alpha} \ln(1 + \theta_i(u_k))(A_i(u_{k+1}) - A_i(u_k)) \right\} \right] \\
\overset{(b)}{=} \prod_k \exp \left\{ \lambda_i(u_k)(u_{k+1} - u_k)(e^{\tilde{\alpha} \ln(1 + \theta_i(u_k))} - 1) \right\} + o(\Delta) \\
= \exp \left\{ \sum_k \tilde{\lambda}_i(1 + \theta_i(u_k))(1 + \theta_i(u_k)^{\tilde{\alpha}} - 1)(u_{k+1} - u_k) \right\} + o(\Delta),
\]

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, for example, (Kim and Whitt 2014). By our hypothesis, \( \theta_i \) is bounded, so we can apply the dominated convergence theorem to get

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

where, again, the limit is taken as \( \Delta := \max_{k} |u_{k+1} - u_k| \to 0 \). In light of (EC.9),

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

Taking expectation of (EC.8) and substituting for the preceding expression, we deduce that

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

which, after further simplification, leads to the desired result. \( \square \)

**Proof of Proposition 4.** To begin, we introduce the following lemma that asserts non-negativity and Lipschitz continuity of the function \( g \).

**Lemma EC.1.** \( g \) is non-negative and Lipschitz continuous on \( \mathbb{R} \).

Note that the properties claimed by the lemma provide the standard (sufficient) condition for equations like (23) to have a unique solution (see, for example, chapter 3 in David et al. (2018)).

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

**Lemma EC.2.** Suppose that Assumption 1 holds. 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.3.** Suppose that Assumption 1 holds. Then \( \eta^* \in (0, \infty] \) and
\[ \sup_{w \geq 0} \pi(w, \eta) = \begin{cases} -\infty & \text{if } \eta < \eta^*, \\ \infty & \text{if } \eta \in [\eta^*, \infty) \cap \mathbb{R}. \end{cases} \] (EC.10)

To gain a better understanding of the function \( \pi(\cdot, \cdot) - \tilde{\ell}_i \), define
\[ \eta_{*,i} := \inf \left\{ \eta > 0 : \sup_{w \geq 0} \pi(w, \eta) \geq \tilde{\ell}_i \right\}. \]

**Lemma EC.4.** Suppose that Assumption 1 holds. Then \( \eta_{*,i} \in (0, \eta^*) \). Moreover, the equation \( \pi(w, \eta) = \tilde{\ell}_i \) uniquely defines two \( C^1 \) functions \( w_1(\cdot), w_2(\cdot) : (\eta_{*,i}, \eta^*) \mapsto (0, \infty) \) such that \( w_1(\eta) < w_2(\eta) \) for all \( \eta \in (\eta_{*,i}, \eta^*) \); in addition,
\[ \begin{align*}
& w_1(\cdot) \text{ is strictly decreasing, } w_2(\cdot) \text{ is strictly increasing}, \\
& \lim_{\eta \downarrow \eta_{*,i}} w_1(\eta) = \lim_{\eta \uparrow \eta^*} w_2(\eta) = \infty, \\
& \text{and } g(\tilde{\ell}_i) + h(w) - \eta > 0 \text{ for all } w > w_2(\eta). \tag{EC.11}
\end{align*} \]

We are now going to prove part (i) of Proposition 4. To this end, let us define
\[ v(\eta) := \int_{w_1(\eta)}^{w_2(\eta)} \left( \pi(u, \eta) - \tilde{\ell}_i \right) du - L_i \quad \text{for } \eta \in (\eta_{*,i}, \eta^*), \] (EC.14)
where \( w_1(\cdot) \) and \( w_2(\cdot) \) are given as in Lemma EC.4. Our goal is to seek some \( \eta_i \) such that \( v(\eta_i) = 0 \). To this end, we differentiate (EC.14) with respect to \( \eta \) and then use the identities
\[ \pi(w_1(\eta), \eta) = \pi(w_2(\eta), \eta) = \tilde{\ell}_i \]
to obtain
\[ v_{\eta}(\eta) = \int_{w_1(\eta)}^{w_2(\eta)} \pi_{\eta}(u, \eta) du > 0 \quad \text{for } \eta \in (\eta_{*,i}, \eta^*), \] (EC.15)
where the inequality follows from Lemma EC.2(i). In addition, from (EC.10) in Lemma EC.3, the continuity of \( \pi \) in \( \eta \) and Lemma EC.4, it follows that
\[ \lim_{\eta \downarrow \eta_{*,i}} v(\eta) = -L_i < 0 \quad \text{and} \quad \lim_{\eta \uparrow \eta^*} v(\eta) = \infty. \] (EC.16)
On combining (EC.15) and (EC.16), we conclude that there exists a unique point \( \eta_i \in (\eta_{*,i}, \eta^*) \) such that \( v(\eta_i) = 0 \). 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 (26) boils down to showing that
\[
\frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_i \geq 0 \quad \text{for} \quad w > s_i
\]
\[
v(w - z) - v(w) + \ell_i z + L_i \geq 0 \quad \text{for} \quad z \geq 0, \quad w \in \mathbb{R}_+,
\]
for any function \( v \) such that \( v'(w) = \pi(w, \eta_i) \). The first inequality follows as a direct consequence of (EC.13), while the second inequality follows by a straightforward calculation involving (EC.14) and the fact that
\[
\pi(w, \eta_i) = \begin{cases} 
< \ell_i & \text{for} \quad w < q_i, \\
> \ell_i & \text{for} \quad w \in (q_i, s_i), \\
= \ell_i & \text{for} \quad w > s_i.
\end{cases}
\]
Thus, we complete the proof for part (ii) of the proposition.

Towards proving part (iii) of Proposition 4, let us define
\[
J_i(w) := \max_{\zeta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\zeta \left[ \int_0^t h(W(u))du - \int_0^t r^*(\zeta(u))du + \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) \right] W(0) = w,
\]
and let
\[
\delta_i(k) := v(W(\tau_i(k))) - v(W(\tau_i(k)-)).
\]
From (26) it follows that \( v(y) - v(x) \leq \tilde{\phi}_i(y - x) \), and so
\[
-\delta_i(k) \leq \tilde{\phi}_i(W(\tau_i(k)-) - W(\tau_i(k))) = \tilde{\phi}_i(\xi_i(k)) \quad \text{for} \quad k = 0, 1, 2, \ldots \tag{EC.17}
\]
On the other hand, applying the generalized Itô’s formula, we obtain, for \( t \geq 0 \),
\[
\mathbb{E}^\zeta [v(W(t))] = v(w) + \mathbb{E}^\zeta \left[ \int_0^t \left( \frac{\sigma^2}{2} v''(W(u)) + \zeta(u)v'(W(u)) \right) du \right] + \mathbb{E}^\zeta \left[ \sum_{k=0}^{N_i(t)} \delta_i(k) \right].
\]
On substituting (EC.17) into above identity and using (26), we deduce
\[
v(w) \leq \mathbb{E}^\zeta \left[ \int_0^t (h(W(u)) - r^*(\zeta(u)) - \eta_i) du \right] + \mathbb{E}^\zeta \left[ \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\xi_i(k)) \right] + \mathbb{E}^\zeta [v(W(t))] + \mathbb{E}^\zeta \left[ \int_0^t (g(v'(W(u))) + r^*(\zeta(u)) - \zeta(u)v'(W(u))) du \right]. \tag{EC.18}
\]
Now, consider a special drift-rate control \( \zeta^\#(W) \), defined as
\[
\zeta^\#(W) := \inf \arg \max_{\zeta} \left\{ v'(W) \zeta - r^*(\zeta) \right\}. \tag{EC.19}
\]
Clearly \( \zeta^\#(\cdot) \) is an adaptive control satisfying
\[
g(v'(W)) + r^*(\zeta^\#(W)) - \zeta^\#(W)v'(W) = 0. \tag{EC.20}
\]
On combining (EC.18) and (EC.20), we see that

\[ v(w) \leq \mathbb{E}^\zeta \left[ \int_0^t (h(W(u)) - r^*(\zeta(u)) - \eta_t) \, du \right] + \mathbb{E}^\zeta \left[ \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\hat{\zeta}_i(k)) \right] + \mathbb{E}^\zeta [v(W(t))]. \] (EC.21)

Now, dividing both sides of (EC.21) 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

\[ \eta_i \leq J_i(w) + \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\zeta [v(W(t))]. \] (EC.22)

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.22). 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

\[ v(W(t)) - v(w) \leq l|W(t) - w| \leq l(W(t) + w) \quad \text{for} \quad t \geq 0. \] (EC.23)

Taking expectation on both sides of (EC.23), 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,

\[ J_i(w) \geq \limsup_{t \to \infty} \mathbb{E}^\zeta \left[ \frac{1}{l} \int_0^t (h(W(u)) - r^*(\zeta(u))) \, du \right] = \infty, \]

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 \( v_m \in \mathcal{C}^2(0, s_i) \) and \( \eta_m \in \mathbb{R} \) such that

\[ \max_\zeta \left\{ \frac{1}{2} \sigma^2 v_m''(w) + \zeta v_m'(w) + h(w) - r^*(\zeta) \right\} = \eta_m, \quad w \in (0, s_i), \]
subject to the boundary conditions

\[ v'_m(0) = 0 \text{ and } v_m(s_i) = \bar{\phi}_i(s_i - q_i) + v_m(q_i). \]

Comparing these with (20) and (21), we immediately conclude that \( v_m = v \) and \( \eta_m = \eta_i \). Therefore, the control rule \( \zeta^\# \) defined by (EC.19) 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 3) completes the proof. \( \square \)

**Proof of Theorem 1.** Towards proving part (i) of the theorem, note that by definition, the value function \( v \) automatically satisfies

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

Hence to establish (27), it suffices to argue

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

(EC.24)

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

\[
\inf_{z \geq 0} [v(w - z) + \bar{\phi}_i(z)] - v(w) \geq \inf_{z \geq 0} [v(w - z) + \bar{\phi}_i,^*(z)] - v(w) \geq 0,
\]

hence (EC.24) 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, without loss of generality, we may assume either (a) \( L_i < L_i,^*, \tilde{\ell}_i > \tilde{\ell}_i,^* \) or (b) \( L_i > L_i,^*, \tilde{\ell}_i < \tilde{\ell}_i,^* \) for \( i \neq i^* \).

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

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

for some \( w \). Thus

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

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 \( v'(s_i,^*) = \tilde{\ell}_i,^* \). Therefore, there exists some \( \tilde{q} \) and \( \tilde{s} \) such that \( \mathbb{L}_{v'}(\tilde{\ell}_i) = [\tilde{q}, \tilde{s}] \subset [q_i,^*, s_i,^*] \). Moreover, by the definition of \( q_i,^* \) and \( s_i,^* \),

\[
\tilde{L} := \int_{\tilde{q}}^{\tilde{s}} [v'(z) - \tilde{\ell}_i] = \sup_{0 \leq q \leq s} \int_q^s [v'(z) - \tilde{\ell}_i] > L_i.
\]
Now, consider a new type of outsourcing operations whose cost function is given as

\[ \tilde{\phi}(w) := (L + \bar{L}, w) \cdot 1_{\{w > 0\}} + 0 \cdot 1_{\{w = 0\}}. \]

By construction, a control band policy with parameters \((\bar{q}, \bar{s})\) and cost function \(\tilde{\phi}\) shares the same value function with \((i^*, q^*, s^*)\), and so the average cost must be equal to \(\eta^*\). On the other hand, it is immediate from the relationship \(L > \bar{L}\) that \(\eta^* > \eta_i\). 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

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

and define \(\delta_i(k)\) in the same way as we did in the proof of Proposition 4. Now using (27) instead of (26) 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

\[ -\delta_i(k) \leq \tilde{\phi}_i(\tilde{\xi}_i(k)). \]  

(EC.25)

Next by applying the generalized Itô’s formula, we obtain, for \(t \geq 0\),

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

On substituting (EC.25) into above identity and using (27), we deduce

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

\[ + \mathbb{E}^\zeta[v(W(t))] + \mathbb{E}^\zeta \left[ \int_0^t (g(v'(W(u))) + r^*(\zeta(u)) - \zeta(u)v'(W(u))) du \right]. \]

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.19), one can formally show that \(\eta^* \leq J(w)\). Second, one can easily argue that \(\zeta^#\) is the maximizer’s best response — when the decision maker chooses \((i^*, q^*, s^*)\), the maximizer will follow \(\zeta^#\) and never deviate from it. \(\square\)

**EC.3. Numerical Algorithm for Solving the Optimality Equation**

To find the solution of the optimality equation (20), we start with an initial guess of \(v\), denoted as \(v_0\), that solves

\[ \frac{1}{2}\sigma^2 v''_0(w) + h(w) = \eta_i, \quad w \in (0, s_i) \]

(EC.26)
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.26) 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 \( \{ \zeta_0 v'_0(w) - r^*(\zeta_0) \} \). The next step is to find \( v_1 \) such that

\[
\frac{1}{2} \sigma^2 v''_1(w) + \zeta_0 v'_1(w) + h(w) - r^*(\zeta_0) = \eta_i, \quad w \in (0, s_i)
\]

(E.27)

subject to the same boundary conditions and necessary optimality conditions as mentioned previously. We can solve (EC.27) 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

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

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.4. Additional Numerical Results**

The outsourcing rules with non-preemptive sequencing are shown in Figures EC.1, EC.2 and EC.3. Tables EC.1, EC.2 and EC.3 compare the optimal costs from the exact solution to those obtained by the SDG for various \((\gamma_1, \gamma_2)\) pairs. Tables EC.4 reports the optimal thresholds used in simulation as well as the asymptotic bounds. Tables EC.5 and EC.6 present the estimated long-run average costs generated from computer simulations.

<table>
<thead>
<tr>
<th>((\gamma_1, \gamma_2))</th>
<th>(\eta_{\text{exact}} ) (preemptive)</th>
<th>(\eta_{\text{exact}} ) (non-preemptive)</th>
<th>(\eta_{\text{SDG}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>((7, 7))</td>
<td>21.8938</td>
<td>24.0212</td>
<td>21.3799</td>
</tr>
<tr>
<td>((12, 12))</td>
<td>20.0744</td>
<td>21.6421</td>
<td>20.0167</td>
</tr>
<tr>
<td>((30, 30))</td>
<td>18.8208</td>
<td>19.8444</td>
<td>18.9634</td>
</tr>
<tr>
<td>((50, 50))</td>
<td>18.3907</td>
<td>19.3362</td>
<td>18.6763</td>
</tr>
<tr>
<td>((1000, 1000))</td>
<td>17.4911</td>
<td>18.6198</td>
<td>18.2935</td>
</tr>
</tbody>
</table>
Figure EC.1  Outsourcing strategies derive from the penalty problem and the SDG with $\gamma_1 = \gamma_2 = 30$ when $\alpha = 2$ with non-preemptive sequencing

Table EC.2  Long-run average cost $\eta$ obtained from numerical programs when $\alpha = 1$

<table>
<thead>
<tr>
<th>$(\gamma_1, \gamma_2)$</th>
<th>$\eta_{exact}$ (preemptive)</th>
<th>$\eta_{exact}$ (non-preemptive)</th>
<th>$\eta_{SDG}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7, 7)</td>
<td>26.4508</td>
<td>27.2448</td>
<td>26.7500</td>
</tr>
<tr>
<td>(12, 12)</td>
<td>22.6574</td>
<td>23.1730</td>
<td>22.0625</td>
</tr>
<tr>
<td>(30, 30)</td>
<td>19.7867</td>
<td>20.3445</td>
<td>19.6836</td>
</tr>
<tr>
<td>(50, 50)</td>
<td>18.9035</td>
<td>19.6659</td>
<td>19.1094</td>
</tr>
<tr>
<td>(1000, 1000)</td>
<td>17.7080</td>
<td>18.6403</td>
<td>18.2891</td>
</tr>
</tbody>
</table>

EC.5.  Proof of Auxiliary Results

*Proof of Lemma EC.1.* The non-negativity of $g$ is immediate. To establish the Lipschitz continuity of $g$, we demonstrate that $g$ is everywhere differentiable and the derivatives are uniformly
(a) Class 1 is in service

(b) Class 2 is in service

Figure EC.2 Outsourcing strategies derived from the penalty problem and the SDG with $\gamma_1 = \gamma_2 = 30$ when $\alpha = 1$ with non-preemptive sequencing

Table EC.3 Long-run average cost $\eta$ obtained from numerical programs when $\alpha = 1/2$

<table>
<thead>
<tr>
<th>$(\gamma_1, \gamma_2)$</th>
<th>$\eta_{\text{exact}}$ (preemptive)</th>
<th>$\eta_{\text{exact}}$ (non-preemptive)</th>
<th>$\eta_{\text{SDG}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7, 7)</td>
<td>34.4145</td>
<td>35.1853</td>
<td>35.9829</td>
</tr>
<tr>
<td>(12, 12)</td>
<td>28.3218</td>
<td>28.8772</td>
<td>26.9434</td>
</tr>
<tr>
<td>(30, 30)</td>
<td>21.7799</td>
<td>22.1417</td>
<td>21.0459</td>
</tr>
<tr>
<td>(50, 50)</td>
<td>20.1876</td>
<td>20.7016</td>
<td>19.7286</td>
</tr>
<tr>
<td>(1000, 1000)</td>
<td>18.0593</td>
<td>18.6757</td>
<td>17.9884</td>
</tr>
</tbody>
</table>

bounded. Using properties of conjugate functions (see, e.g., Ex. 3.40 in Boyd et al. (2004)), we
have that
\[
g'(x) = \arg\max_{\zeta \in \mathcal{Z}} x\zeta - r^*(\zeta), \quad \text{(EC.28)}
\]
where \( \mathcal{Z} := [\rho^\top a, \rho^\top b] \). The left-hand side of (EC.28) ought to be understood as subderivative (or subgradient) if the right-hand side contains multiple elements. Hence, the desired Lipschitz continuity will follow if we can show that the right-hand side of (EC.28) is a singleton for all \( x \in \mathbb{R} \).

We prove this by showing that \( r^* \) is strictly convex. For that purpose, pick arbitrarily \( \zeta_1, \zeta_2 \in \mathcal{Z} \) and a convex combination \( \zeta_3 := \alpha \zeta_1 + (1 - \alpha) \zeta_2 \) for \( \alpha \in (0, 1) \) (where we have overloaded the notation \( \alpha \)

\textbf{Figure EC.3} Outsourcing strategies derived from the penalty problem and the SDG with \( \gamma_1 = \gamma_2 = 30 \) when \( \alpha = 1/2 \) with non-preemptive sequencing
Table EC.4  Asymptotic bounds and outsourcing parameters for $\alpha = 1/2, 1$ and 2

$$
\begin{array}{cccccccc}
\gamma_1, \gamma_2 & \alpha = 1/2 & & & \alpha = 1 & & & \alpha = 2 \\
(8, 8) & 0.016 & 0.882 & 25.26 & (6, 6) & 0.058 & 0.573 & 17.93 \\
(8, 12) & 0.016 & 0.877 & 24.58 & (6, 12) & 0.061 & 0.566 & 17.30 \\
(8, 20) & 0.016 & 0.873 & 24.06 & (6, 20) & 0.062 & 0.564 & 17.05 \\
(8, 200) & 0.017 & 0.870 & 23.40 & (6, 200) & 0.063 & 0.561 & 16.73 \\
(12, 8) & 0.039 & 0.668 & 19.34 & (12, 6) & 0.105 & 0.502 & 13.66 \\
(12, 12) & 0.040 & 0.659 & 18.63 & (12, 12) & 0.116 & 0.498 & 13.06 \\
(12, 20) & 0.042 & 0.654 & 18.09 & (12, 20) & 0.120 & 0.497 & 12.83 \\
(12, 200) & 0.043 & 0.649 & 17.41 & (12, 200) & 0.126 & 0.496 & 12.53 \\
(20, 8) & 0.073 & 0.542 & 15.37 & (20, 6) & 0.140 & 0.495 & 12.19 \\
(20, 12) & 0.081 & 0.531 & 14.69 & (20, 12) & 0.159 & 0.499 & 11.61 \\
(20, 20) & 0.086 & 0.524 & 14.17 & (20, 20) & 0.166 & 0.502 & 11.40 \\
(20, 200) & 0.093 & 0.518 & 13.51 & (20, 200) & 0.177 & 0.506 & 11.11 \\
(200, 8) & 0.159 & 0.501 & 11.23 & (200, 6) & 0.214 & 0.526 & 10.34 \\
(200, 12) & 0.197 & 0.517 & 10.63 & (200, 12) & 0.251 & 0.552 & 9.87 \\
(200, 20) & 0.229 & 0.536 & 10.19 & (200, 20) & 0.267 & 0.565 & 9.69 \\
(200, 200) & 0.272 & 0.568 & 9.58 & (200, 200) & 0.289 & 0.581 & 9.46 \\
\end{array}
$$

Table EC.5  Estimated average costs from simulations for ARIMA intensity with $\alpha = 1/2, 1$ and 2

$$
\begin{array}{cccccccc}
\gamma_1, \gamma_2 & \alpha = 1/2 & & & \alpha = 1 & & & \alpha = 2 \\
(8, 8) & 15.822 & ±3E-1 & & & & & \\
(8, 12) & 15.655 & ±3E-1 & & & & & \\
(8, 20) & 15.189 & ±3E-1 & & & & & \\
(8, 30) & 14.776 & ±3E-1 & & & & & \\
(8, 50) & 15.309 & ±3E-1 & & & & & \\
(8, 200) & 15.526 & ±2E-1 & & & & & \\
(12, 8) & 11.700 & ±3E-1 & & & & & \\
(12, 12) & 11.010 & ±3E-1 & & & (4, 4) & & \\
(12, 20) & 10.746 & ±3E-1 & & & & & \\
(12, 30) & 10.994 & ±3E-1 & & & & & \\
(12, 50) & 11.395 & ±3E-1 & & & & & \\
(12, 200) & 11.370 & ±3E-1 & & & & & \\
(20, 8) & 9.687 & ±2E-1 & & & & & \\
(20, 12) & 9.680 & ±2E-1 & & & & & \\
(20, 20) & 9.391 & ±2E-1 & & & & & \\
(20, 30) & 9.360 & ±2E-1 & & & & & \\
(20, 50) & 9.724 & ±3E-1 & & & & & \\
(20, 200) & 9.568 & ±2E-1 & & & & & \\
(30, 8) & 9.551 & ±2E-1 & & & & & \\
(30, 12) & 9.413 & ±2E-1 & & & & & \\
(30, 20) & 9.419 & ±3E-1 & & & & & \\
(30, 30) & 9.339 & ±2E-1 & & & & & \\
(30, 50) & 9.424 & ±2E-1 & & & & & \\
(30, 200) & 9.471 & ±2E-1 & & & & & \\
(200, 8) & 9.423 & ±2E-1 & & & & & \\
(200, 12) & 9.414 & ±2E-1 & & & & & \\
(200, 20) & 9.790 & ±3E-1 & & & & & \\
(200, 30) & 9.737 & ±2E-1 & & & & & \\
(200, 50) & 9.875 & ±3E-1 & & & & & \\
(200, 200) & 10.381 & ±3E-1 & & & & & \\
\end{array}
$$

for convenience). We intend to show that

$$ r^*(\zeta_3) < \alpha r^*(\zeta_1) + (1 - \alpha)r^*(\zeta_2). $$
Table EC.6  Estimated average costs from simulations for CIR intensity with $\alpha = 1/2, 1$ and $2$

<table>
<thead>
<tr>
<th>$\alpha = 1/2$</th>
<th>$\alpha = 1$</th>
<th>$\alpha = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\gamma_1, \gamma_2)$</td>
<td>Simulated Cost</td>
<td>$(\gamma_1, \gamma_2)$</td>
</tr>
<tr>
<td>(8, 8)</td>
<td>14.364±E-1</td>
<td>(6, 4)</td>
</tr>
<tr>
<td>(8, 12)</td>
<td>14.137±E-1</td>
<td>(6, 6)</td>
</tr>
<tr>
<td>(8, 20)</td>
<td>14.121±E-1</td>
<td>(6, 8)</td>
</tr>
<tr>
<td>(8, 30)</td>
<td>13.983±E-1</td>
<td>(6, 12)</td>
</tr>
<tr>
<td>(8, 50)</td>
<td>14.047±E-1</td>
<td>(6, 20)</td>
</tr>
<tr>
<td>(8, 200)</td>
<td>14.078±E-1</td>
<td>(6, 200)</td>
</tr>
<tr>
<td>(12, 8)</td>
<td>10.693±E-1</td>
<td>(8, 4)</td>
</tr>
<tr>
<td>(12, 12)</td>
<td>10.357±E-1</td>
<td>(8, 6)</td>
</tr>
<tr>
<td>(12, 20)</td>
<td>10.234±E-1</td>
<td>(8, 8)</td>
</tr>
<tr>
<td>(12, 30)</td>
<td>10.235±E-1</td>
<td>(8, 12)</td>
</tr>
<tr>
<td>(12, 50)</td>
<td>10.279±E-1</td>
<td>(8, 20)</td>
</tr>
<tr>
<td>(12, 200)</td>
<td>10.305±E-1</td>
<td>(8, 200)</td>
</tr>
<tr>
<td>(20, 8)</td>
<td>9.158±E-1</td>
<td>(12, 4)</td>
</tr>
<tr>
<td>(20, 12)</td>
<td>9.022±E-1</td>
<td>(12, 6)</td>
</tr>
<tr>
<td>(20, 20)</td>
<td>8.977±E-1</td>
<td>(12, 8)</td>
</tr>
<tr>
<td>(20, 30)</td>
<td>8.929±E-1</td>
<td>(12, 12)</td>
</tr>
<tr>
<td>(20, 50)</td>
<td>8.938±E-1</td>
<td>(12, 20)</td>
</tr>
<tr>
<td>(20, 200)</td>
<td>8.985±E-1</td>
<td>(12, 200)</td>
</tr>
<tr>
<td>(30, 8)</td>
<td>8.916±E-1</td>
<td>(20, 4)</td>
</tr>
<tr>
<td>(30, 12)</td>
<td>8.864±E-1</td>
<td>(20, 6)</td>
</tr>
<tr>
<td>(30, 20)</td>
<td>8.857±E-1</td>
<td>(20, 8)</td>
</tr>
<tr>
<td>(30, 30)</td>
<td>8.822±E-1</td>
<td>(20, 12)</td>
</tr>
<tr>
<td>(30, 50)</td>
<td>8.874±E-1</td>
<td>(20, 20)</td>
</tr>
<tr>
<td>(30, 200)</td>
<td>8.890±E-1</td>
<td>(20, 200)</td>
</tr>
<tr>
<td>(20, 8)</td>
<td>8.891±E-1</td>
<td>(200, 4)</td>
</tr>
<tr>
<td>(20, 12)</td>
<td>8.935±E-1</td>
<td>(200, 6)</td>
</tr>
<tr>
<td>(20, 20)</td>
<td>9.106±E-1</td>
<td>(200, 8)</td>
</tr>
<tr>
<td>(20, 30)</td>
<td>9.238±E-1</td>
<td>(200, 12)</td>
</tr>
<tr>
<td>(20, 50)</td>
<td>9.344±E-1</td>
<td>(200, 20)</td>
</tr>
<tr>
<td>(20, 200)</td>
<td>9.616±E-1</td>
<td>(200, 200)</td>
</tr>
</tbody>
</table>

To that end, we note that for each $\zeta$, $r^\star(\zeta)$ is the optimal objective value of a convex optimization problem. Denote by $\theta(\zeta)$ the optimal solution to the convex optimization problem. It is easy to check that $\alpha \theta(\zeta_1) + (1 - \alpha) \theta(\zeta_2)$ is a feasible solution to the convex optimization problem that defines $r^\star(\zeta_3)$. Thus

$$r^\star(\zeta_3) \leq r(\alpha \theta(\zeta_1) + (1 - \alpha) \theta(\zeta_2)) < \alpha r(\theta(\zeta_1)) + (1 - \alpha) r(\theta(\zeta_2)) = \alpha r^\star(\zeta_1) + (1 - \alpha) r^\star(\zeta_2),$$

where the second inequality is due to the strict convexity of $r$. This demonstrates that $r^\star$ is strictly convex, implying that the right-hand side of (EC.28) is a singleton for all $x \in \mathbb{R}$. The desired Lipschitz continuity thus follows. □

Proof of Lemma EC.2. 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

$$\liminf_{w \to \infty} \pi(w, \eta), \quad \limsup_{w \to \infty} \pi(w, \eta) \in \{-\infty, \infty\}. \quad \text{(EC.29)}$$
Suppose for the sake of contradiction that \( \liminf_{w \to \infty} \pi(w, \eta) \in (-\infty, \infty) \). Then choose a sequence \( w_n \to \infty \) such that
\[
\lim_{n \to \infty} \pi(w_n, \eta) = \liminf_{w \to \infty} \pi(w, \eta) \quad \text{and} \quad \lim_{n \to \infty} \pi_w(w_n, \eta) = 0.
\]

If assuming \( \limsup_{w \to \infty} \pi(w, \eta) \in (-\infty, \infty) \), then choose a sequence \( w_n \to \infty \) in a similar fashion. In either case, we obtain via direct calculation that
\[
0 = \lim_{n \to \infty} \pi_w(w_n, \eta) = \lim_{n \to \infty} -\frac{2}{\sigma^2} [g(\pi(w_n, \eta)) + h(w_n) - \eta] = -\infty,
\]
where the last equality is owing to Assumption 1. However, the preceding calculation provides the desired contradiction, so we must have (EC.29). 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 (23) it follows that
\[
\pi_w(w, \eta) = -\frac{2}{\sigma^2} [h(w) - \eta] \quad \text{for all} \quad w > 0 \quad \text{s.t.} \quad \pi(w, \eta) = 0. \tag{EC.30}
\]

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. \(\square\)

**Proof of Lemma EC.3.** In view of (23) and the non-negativity of \( h \) and \( g \), we see that
\[
\pi(w, \eta) \leq 0 \quad \text{for all} \quad w \in \mathbb{R} \quad \text{and} \quad \eta \leq 0,
\]
which in turn implies that \( \eta^* \in (0, \infty] \). Because \( \pi(w, \cdot) \) is strictly increasing for any fixed \( w > 0 \),
\[
\sup_{w \geq 0} \pi(w, \eta) < \infty \quad \text{for} \quad \eta < \eta^* \quad \text{and} \quad = \infty \quad \text{for} \quad \eta \in (\eta^*, \infty] \cap \mathbb{R}.
\]

In view of Lemma EC.2(ii), proving (EC.10) boils down to showing that
\[
\lim_{w \to \infty} \pi(w, \eta^*) = \infty \quad \text{if} \quad \eta^* < \infty.
\]

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
\[
\pi(w, \eta^*) < 0 \quad \text{and} \quad h(w) - \eta^* > \chi > 0 \quad \text{for all} \quad w \geq \hat{w}, \tag{EC.31}
\]
where \( \chi \) is any given constant. Now, combining the second inequality in (EC.31) with (EC.30), we see that, for any \( \eta \in [\eta^*, \eta^* + \chi] \),
\[
\pi_w(w, \eta) < 0 \quad \text{for all} \quad w \geq \hat{w} \quad \text{such that} \quad \pi(w, \eta) = 0.
\]
However, this calculation together with the fact that \( \lim_{x \to \infty} \pi(w, \eta) = \infty \) 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

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

But the preceding inequality contradicts the first inequality in (EC.31), due to the continuity of \( \pi \) in \( \eta \). Thus, we must have \( \sup_{w \geq 0} \pi(w, \eta^*) = \infty \), hence (EC.10).

**Proof of Lemma EC.4.** By the fact that \( h \geq 0 \) and \( \tilde{\ell}_i > 0 \), 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 (23) it follows that

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

In view of the definitions of \( \eta_{*,i} \) and \( \eta^* \), (EC.10) in Lemma EC.3, 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.13) follows from Assumption 1 and (b). Moreover, the first assertion in (EC.12) follows from (a) and (b), whereas the second assertion follows from (b) and (c) as well as (EC.11). Hence, to finish the proof, it suffices to establish (EC.11). To this end, differentiate \( \pi(w_j(\eta), \eta) = \tilde{\ell}_i \) with respect to \( \eta \) to get

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

for all \( \eta \in (\eta_{*,i}, \eta^*) \). In view of Lemma EC.2(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. \( \square \)