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

An optimal control policy for make-to-order manufacturing systems is typically derived assuming the availability of a precise probabilistic model. However, in practice, the underlying probabilistic model may be a simplification of the actual scenario due to tractability considerations or the lack of operational data for calibration. Consequently, policies derived from such simplifications may perform poorly if the assumed model does not accurately capture reality. To address this issue, this paper proposes a modeling paradigm that employs a simplified (low-fidelity) model while accounting for potential sources of error or uncertainty that may arise from its use. The focus is on controlling a multi-product make-to-order manufacturing system with an outsourcing mechanism, and the paper specifically addresses model misspecification in the demand for different products. Specifically, the paper introduces a robust control formulation in the form of a two-player zero-sum game where model uncertainty is described using the notion of Rényi divergence. The paper further presents an approximating problem under the heavy-traffic assumption that effectively constitutes a stochastic differential game. The solution to the stochastic differential game is derived using dynamic programming, and based on this solution, an implementable control rule is proposed for the original make-to-order system. Additionally, the paper presents a simulation-based 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 approximation; diffusion approximation; stochastic differential games

1. Introduction

This paper addresses the distributionally robust control of a make-to-order manufacturing system that produces multiple types of products with a shared capacity. While most items are produced in-house, the system also has the option of subcontracting or outsourcing its manufacturing needs for certain products at a fixed plus proportional outsourcing cost in response to unexpected upticks in demand. An example that is applicable to this context is additive manufacturing, also known as 3D printing. This technology utilizes 3D printers to produce a variety of printable products on demand. Since products are printed directly from a computer model, 3D printing operations cannot stockpile inventory in order to decouple demand variability from production. In the event of demand surges, subcontractors, such as professional 3D printing bureaus, can be engaged to supplement in-house manufacturing efforts. The prioritization of orders for 3D-printed products can be based on their
relative urgency, and 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 transformative technology, 3D printing is expected to have a massive impact on 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 using either stochastic dynamic programming (Carr and Duenyas 2000, Öner-Közen and Minner 2017) or optimal control techniques (Çelik and Maglaras 2008, Rubino and Ata 2009). However, existing research on controlling a make-to-order system typically assumes that the probability law describing realized demand is precisely known and faithfully represents reality. In practice, however, the demand model may be misspecified due to tractability considerations, as real-world situations are often much more complex than what simple models can capture. For instance, a doubly stochastic Poisson process with an auto-regressive intensity process may be suitable for capturing auto-correlation among demand arrivals. However, to optimize decisions, one would need to know the value of intensity over time, which is rarely observable. Moreover, even if its value can be observed, accounting for it in modeling increases the problem dimension, making the problem computationally challenging. In general, even though it is often possible to come up with a more complex “high-fidelity” model to describe demand realization, such a model is bound to introduce new analytical and computational challenges. As a result, a decision-maker may resort to a “low-fidelity” model by making simplifying assumptions, such as “Markovian” and “stationarity,” on the demand realization process, giving rise to model misspecification.

Given that a simplified model may carry model errors, the decision-maker needs to account for the biases created by the model and develop strategies to mitigate their negative impacts. Robust optimization (RO) is a modeling paradigm that is well-suited for addressing such an issue and often employed to resolve the trade-off between using a high-fidelity model that better reflects reality and adopting a low-fidelity model for tractability. It addresses model misspecification by introducing an uncertainty set that is believed to contain the parameters of the true model. Recent contributions to this field include Bandi et al. (2019), Sun and Van Mieghem (2019). Although classical RO has demonstrated its usefulness in various applications, it is a static approach and may produce excessively conservative solutions for problems that involve both immediate and wait-and-see decisions.

This paper presents a framework that provides effective decision rules for a make-to-order manufacturing system when the correct demand model is known (or can be reliably estimated) but highly complex and the decision-maker is able to find a simplified model believed to capture the essential features of the real-world system. Thus, the decision-maker faces model uncertainty when
contemplating control strategies. The framework is close in spirit to a stream of research pioneered by Petersen et al. (2000), Hansen and Sargent (2001). The main idea is to extract a nominal model and add a malevolent second player (“nature”) that perturbs the nominal model within some prescribed limits. The malevolence of nature serves as a tool for the decision-maker to explore the fragility of candidate decision rules. Nature’s perturbations enable random shocks to feed back into the state process, allowing the uncertainty set to be only loosely specified, which distinguishes this approach from the classical RO setting.

In greater detail, we assume demands for each product arrive according to a non-homogeneous Poisson process, but the decision-maker does not know the true value of arrival intensity beyond knowing that it may fluctuate around some long-term average (which hints at the nominal model). Hence, the decision-maker treats the intensity as if it were chosen by nature. The decision-maker then calculates the expected long-run average cost while assuming that nature will create the worst-case scenario and seeks a joint sequencing and outsourcing rule to minimize it (using the minimax criterion). This gives rise to a robust control problem in the form of a stochastic two-player zero-sum game where nature’s actions are constrained by penalizing deviations from the nominal model. The extent to which nature is punished depends on how much the decision-maker distrusts the nominal model. In this formulation, the decision-maker seeks to take the best action in response to nature’s alternation of the beliefs reflected in the nominal model.

The present study adopts an infinite time horizon as manufacturing firms usually do not have a predetermined end date, and we choose an average cost criterion rather than a discounted cost criterion because time discounting is less common in manufacturing settings (Ormeci et al. 2008) and an average cost problem is generally easier to handle as it does not require dealing with a second-order differential equation. Furthermore, we employ Rényi divergence, a measure of the difference between two probability distributions, to construct uncertainty sets for two main reasons. First, Rényi divergence provides greater flexibility in modeling compared to Kullback-Leibler (KL) divergence, as demonstrated in our numerical experiments, where it leads to better control rules compared to using KL divergence only. Second, Rényi divergence leads to a tractable representation of the distance between models, simplifying analysis and computation. While other generalizations of KL divergence, such as f-divergence, could potentially be considered, they do not seem to be able to translate into a tractable representation of the distance between models in the present context.

As the original robust control problem appears to be analytically challenging, we employ the “heavy-traffic approximation” to make further headway. This approach assumes that both demand and service capacity are high, and the server utilization is close to 100%, meaning that the production server must be busy most of the time to meet the demand. This approximation scheme aligns with the framework pioneered by Harrison (1988), where a Brownian control problem emerges as the
heavy-traffic limit of a sequence of queuing control problems. Since we formulate our robust control problem as a stochastic game, the heavy-traffic approximation can lead us to a stochastic differential game (SDG). We show that the SDG can be further converted to an equivalent one-dimensional differential game, referred to as the workload problem, where the state-descriptor is one-dimensional and tracks the amount of work in the system over time. The conversion reduces the dimension of the problem while dictating how sequencing decisions ought to be made to achieve the lowest holding cost possible. The solution to the one-dimensional problem comprises a control-band policy for the decision-maker and a drift-rate control for nature. Whenever the workload exceeds an upper barrier, the decision-maker promptly reduces it by outsourcing the manufacturing of one particular product, bringing it back to a lower threshold level. Between two consecutive outsourcing operations, nature uses a state-dependent drift-rate control to resolve the decision-maker’s ambiguity aversion.

Our presentation and analysis of the heavy-traffic approximation are informal to some extent. For instance, we do not provide a rigorous proof showing that the SDG arises as the heavy-traffic limit of a sequence of robust control problems. (We do not even utilize a sequence of problems in developing the SDG.) Furthermore, there is no attempt to prove that the control rule, derived from the solution to the workload problem (and therefore the SDG), is asymptotically optimal in some proper sense. In our own assessment, presenting a rigorous proof for asymptotic optimality would be an arduous task and would deviate from our two primary objectives of using an approximate analysis: (i) to gain fundamental insights into the optimal solutions in heavy traffic and (ii) to develop effective and easy-to-implement control strategies. However, much of our analysis draws upon results that are known for simpler systems, and therefore, we expect that the approximating solution, derived from the SDG, is close to the “exact” solution to the original robust control problem. In summary, we regard the insights gained from our analysis as providing a deeper understanding of the optimal solution to the original robust control problem in heavy traffic.

We view the contributions in our paper as trifolds. 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. Despite its clear focus on the control of make-to-order systems, to the best of our knowledge, this paper is the first to address a long-run average control problem while accounting for model misspecification, and our use of Rényi divergence to describe model uncertainty extends the widely used entropic approach found in the existing literature. Second, assuming the capacity is adequately utilized, we derive and solve an SDG that approximates the original robust control problem. In this SDG, since nature controls the drift of the state process, a pathwise optimal solution cannot be identified for the decision-maker. We demonstrate that the solution to the game can be found by solving a nonlinear differential equation with a set of free boundaries. Third, this paper proposes a simulation-based approach that enables the identification of an appropriate level of
ambiguity in a computationally efficient manner. This approach combines an analytical model with computer simulations to generate and evaluate a family of closely related control rules, producing a control rule that is deemed the most suitable for real-world implementation.

2. Literature Review

This work draws on the literature on controlling queues. When dealing with Poisson arrivals and linear delay cost rates, the \( c \mu \) rule, which assigns static priority levels to jobs based on their \( c_i \mu_i \) values, is known to minimize delay costs by assigning (Cox and Smith 1991). Some extensions of this rule take into account more sophisticated cost structures (Van Mieghem 1995, Akan et al. 2012) and/or job abandonment from the queue, including Rubino and Ata (2009), Ata and Tongarlak (2013), Kim and Ward (2013), Ghamami and Ward (2013). Several papers have explored the combination of economic levers (e.g., pricing) with operational decisions in managing make-to-order systems, such as Çelik and Maglaras (2008), Ata and Olsen (2013). Our paper differs from these studies in one crucial aspect: these papers assume an accurate probabilistic model for optimization, whereas we consider potential model misspecification, leading to a min-max optimal control problem.

This paper contributes to the literature on sequential decision-making under ambiguous beliefs, which typically involves a nominal model, considered to be a parsimonious (presumably oversimplified) representation of the real-world scenario and a malevolent agent, who can create alternative models by perturbing the nominal model. This concept can be traced back to Petersen et al. (2000) and Hansen and Sargent (2001), and has been applied in various domains, including dynamic pricing (Lim and Shanthikumar 2007), corporate investment (Nishimura and Ozaki 2007), and the probability of lifetime ruin (Bayraktar and Zhang 2015), among others. The typical problem formulation is in the form of a two-player stochastic game that can be solved using a dynamic programming equation. In addition to tackling different problems, these papers employ KL divergence to capture ambiguous beliefs. In contrast, the present paper goes beyond KL divergence to characterize model uncertainty in stochastic dynamic programming problems, representing an important step forward in expanding this modeling paradigm.

Our work shares similarities with a recent paper by Cohen (2019) that explores a Brownian control problem arising from the Brownian approximation of a multiclass M/M/1 queue under model uncertainty. However, there are key distinctions between our work and his. First, Cohen (2019) assumes a finite buffer for all queues, with pre-defined buffer sizes that are not subject to optimization. Consequently, for the workload problem developed therein, a barrier-type (singular) control is implemented at a predetermined level, meaning that the sole control lever is job sequencing. Cohen (2019) shows the \( c \mu \) rule is optimal in the pathwise sense, meaning it is independent of the solution to the dynamic programming equation. As a result, the choice of uncertainty set has no
bearing on the decision-maker’s control strategy. In contrast, our model involves outsourcing and the derived outsourcing strategy is not pathwise and, therefore, is influenced by the choice of the uncertainty set. Second, Cohen (2019) considers an infinite-horizon discounted cost criterion, resulting in solving a second-order nonlinear differential equation with fixed boundary conditions, whereas we consider a long-run average cost criterion, leading to solving a first-order differential equation with free boundary boundaries. Our use of Rényi-type divergence further distinguishes our work from Cohen (2019). In summary, the approach of Cohen (2019) is not applicable to our problem, even if we also employ the entropic method to describe model uncertainty and eliminate fixed outsourcing costs.

Methodologically, our paper is related to the impulse control of Brownian systems. Harrison et al. (1983) and Dai and Yao (2013b) consider discounted cost formulations, and Ormeci et al. (2008) and Dai and Yao (2013a) study average cost problems. Via explicit solutions, these studies reveal that a control band policy is optimal. Our problem cannot be solved using their methods since the drift rate is not constant and subject to the control by nature. Thus, obtaining explicit solutions to the optimality equations is impossible. Instead, we adopt an approach that does not require explicit solutions to the value function. A few recent papers (Yao 2017, Cao and Yao 2018) have considered joint drift-rate control and impulse control for Brownian models in joint pricing and inventory control problems. Our work differs from theirs in two aspects: (i) their problems belong to the class of cost minimization problems, while ours adopts a min-max criterion; (ii) their problems deal with a single-class model, while ours inherently deals with multiclass models.

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 $\mathbb{F} := (\mathcal{F}_t)$ contained in $\mathcal{F}$. We focus on 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 $i$ (also known as class $i$ orders) arrive according to a Poisson process with rate $\bar{\lambda}_i$. The number of class $i$ orders placed up to time $t$ is denoted by $A_i(t)$, and $\bar{\lambda} := (\bar{\lambda}_i)$ denotes the arrival rate vector that collects all the arrival rates. The time required 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 time $t$, assuming that 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 (first-in-first-out) sequencing principle within each queue. Therefore, sequencing decisions can be described by an $I$-dimensional time allocation process $T := (T_i)$, where $T_i(t)$ represents the amount of time spent by the server on producing product $i$. The total idle time up to time $t$ can be calculated by $(t - \sum T_i(t))$. 
The decision-maker has the option to outsource manufacturing needs at a fixed plus proportional cost. Specifically, outsourcing a batch size of \( x \) class \( i \) orders incurs a cost

\[
\phi_i(x) := (L_i + \ell_i x) \cdot 1_{\{x>0\}},
\]

where \( L_i \) and \( \ell_i \) are fixed and proportional costs, respectively. By doing this, the decision-maker is able to instantly reduce the backlog of class \( i \) orders by \( x \) units. We will refer to this type of outsourcing activity as a type \( i \) outsourcing operation. Let \( \Psi := (\Psi_i) \), where the \( i \)th component is given by

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

where \( 0 = \tau_i(0) < \tau_i(1) < \tau_i(2) < \cdots \) 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 time \( t \). Denote by \( Q_i(t) \) the number of outstanding orders of class \( i \) in the system at time \( t \), and let \( \Psi_i := (\Psi_i) \). Assuming that the system initially has \( Q_i(0) \) class \( i \) orders, we can describe the dynamics of \( Q_i(t) \) using the equation

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

The cost of holding a backlog of class \( i \) orders is incurred at a rate of \( c_i(Q_i(t)) \). Hence, the total backlog penalty is accumulated at the rate of \( \sum_i c_i(Q_i(t)) \). Throughout this paper, we assume that the cost rate functions satisfy the following assumption.

**Assumption 1.** For every class \( i \), \( c_i(\cdot) : \mathbb{R}_+ \mapsto \mathbb{R}_+ \) is a known function that is continuous and monotonically increasing towards infinity.

A control policy can be represented by the pair of controls \((T, \Psi)\). Using the long-run average cost criterion, the decision-maker aims to find an adaptive control \((T, \Psi)\) that minimizes

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

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 a product \( i \) may fluctuate over time, possibly in a random manner. As a result, a time-dependent arrival rate function \( \lambda_i := \lambda_i(t); t \geq 0 \) emerges instead of a constant rate \( \bar{\lambda}_i \). To account for this, the decision-maker is interested in devising control policies that are robust against such variations. To model this scenario, we assume that the decision-maker acts as if there is a second
player (nature) who strategically chooses the vector of demand rate processes $\lambda := (\lambda_i)$. We further define a perturbation process $\theta_i := \{\theta_i(t); t \geq 0\}$ for each $\lambda_i$, where $\theta_i(t) := (\lambda_i(t) - \bar{\lambda}_i)/\bar{\lambda}_i$ represents the relative deviation of $\lambda_i$ from its nominal value $\bar{\lambda}_i$ at time $t$. We require that each $\theta_i$ satisfies the condition that $\theta_i(t) \in \Theta_i := [a_i, b_i]$ for all $t \geq 0$, where $-1 < a_i < 0 < b_i < \infty$. The reason for imposing a lower bound of $-1$ on $a_i$ is to ensure that the values of $\theta_i$ are also bounded below by $-1$. If $\theta_i(t) < -1$ at time $t$, $\lambda_i(t)$ will become negative at that time, which does not make practical sense. Henceforth, let $\theta := (\theta_i)$ denote the perturbation process. Note that $\lambda$ and $\theta$ are uniquely determined by each other. For this reason, we will treat $\theta$ as nature’s decision process instead of $\lambda$.

To capture the decision-maker’s aversion towards model uncertainty, we follow the convention established by Petersen et al. (2000) and Hansen and Sargent (2001), by assuming that nature selects a $\theta$ to maximize the decision-maker’s cost, leading up to a worst-case analysis. One major benefit of a worst-case analysis is that it provides a conservative estimate of the solution’s performance. By assuming that the model input takes on its worst possible values within the specified uncertainty set, the decision-maker can design a solution that performs well even under the most unfavorable scenario. Put differently, if the decision-maker is willing to accept the solution’s performance even in the most unfavorable scenario, he can be confident in accepting any other outcomes that may arise due to model uncertainty. The worst-case approach also aligns with the core concept of robust optimization, as it frees the decision-maker from the need to fully specify a probability distribution governing uncertain model input. It ensures that the solution is resilient to uncertainties about the model input and provides a safety net against unexpected changes in the model input.

The degree of model uncertainty can be accounted for by imposing constraints on nature’s perturbations to the nominal model. This calls for a measure that can be used to quantify the extent to which perturbations are made. For this purpose, we define, for each $\mathbb{F}$-predictable $\theta_i$, the Doléans-Dade exponential:

$$\psi_i(t) := \exp\left\{-\int_0^t \bar{\lambda}_i \theta_i(u) du\right\} \prod_{0 < u \leq t} (1 + \theta_i(u))^{\Delta A_i(u)}, \quad (3)$$

where $\Delta A_i(t) := A_i(t) - A_i(t^-)$; see Definition 9.4.3.1 in (Jeanblanc et al. 2009, §9.4.3). Note that if $\theta_i \equiv 0$, then $\psi_i \equiv 1$, implying $\lambda_i \equiv \bar{\lambda}_i$. Moreover, the process $\psi_i := \{\psi_i(t); t \geq 0\}$ is a martingale. We denote the marginal distribution of $A_i$ in the nominal model by $\mathbb{P}_i$, and define a new measure $Q_i$ using the Radon-Nikodym derivative:

$$\frac{dQ_i}{d\mathbb{P}_i} \bigg|_{\mathcal{F}_t} = \psi_i(t) \quad \text{for} \quad t \geq 0. \quad (4)$$

By the Girsanov theorem for filtered Poisson processes (Jeanblanc et al. 2009, Proposition 8.4.5.1), $A_i$ is a filtered Poisson process with intensity $\lambda_i(t) = \bar{\lambda}_i(1 + \theta_i(t))$ under the induced measure $Q_i$. 

Identity (4) suggests that changing the intensity of $A_i$ from $\bar{\lambda}_i$ to $\lambda_i$ is equivalent to changing the measure from $\mathbb{P}_i$ to the induced measure $Q_i$. This relationship brings the induced measure $Q_i$ to the forefront, allowing the use of a notion of distance between the measures $\mathbb{P}_i$ and $Q_i$, to measure the size of the perturbations $\theta_i$. Therefore, limiting the size of $\theta_i$ would correspond to controlling the distance between the two measures. In this paper, the distance between each $(\mathbb{P}_i, Q_i)$ pair is evaluated using the 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

$$ \mathcal{R}_\alpha(\tilde{\mathbb{P}} \parallel \mathbb{P}) := \frac{1}{\alpha - 1} \ln \int \left( \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \right)^\alpha d\mathbb{P} = \frac{1}{\alpha - 1} \ln \int \left( \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \right)^{\alpha - 1} d\tilde{\mathbb{P}}. $$

When $\alpha = 1$, the Rényi divergence reduces to the relative entropy or KL divergence (Van Erven and Harremos 2014), so one may assume $\alpha \in (0, \infty)$. Let $\mathcal{R}_\alpha^i(t)$ denote the Rényi divergence of $Q_i$ of order $\alpha$ with respect to $\mathbb{P}_i$ on $F_t$. The following result links $\mathcal{R}_\alpha^i(t)$ to $\theta_i$.

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

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

The proof for the case where $\alpha = 1$ appeared in Lim and Shanthikumar (2007). However, we are unaware of other work prior to the submission of this work that utilizes Rényi divergence to construct an uncertainty set for problems of this type. Therefore, we believe our proof for the case where $\alpha \neq 1$ is novel. The detailed proof is presented in §EC.2.

After obtaining the explicit expression for $\mathcal{R}_\alpha^i(t)$, we can put each induced measure $Q_i$ back into the background and treat the expression as the definition of $\mathcal{R}_\alpha^i(t)$. This approach enables us to focus solely on the perturbation processes $\theta_i$, rather than the measures they induce. With this, we have all the vocabulary we need to present our robust control problem. In this problem, the decision-maker seeks to find $(T, \Psi)$ to minimize

$$ \max_{\theta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta \left[ \sum_{i=1}^I \int_0^t c_i(Q_i(u)) du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) - \sum_{i=1}^I \gamma_i \mathcal{R}_\alpha^i(t) \right], \quad (5) $$

where the superscript $\theta$ is added to the expectation operator to emphasize that the arrival intensities now follow $\bar{\lambda} + \theta$ instead of $\bar{\lambda}$ in the nominal model. This formulation constrains nature’s actions by penalizing deviations from the nominal model. The extent to which nature is punished depends on how much the decision-maker distrusts the nominal model: a large value of $\gamma_i$ imposes severe penalties on nature for perturbations $\theta_i$, whereas a small value of $\gamma_i$ applies lighter penalties, encouraging nature to take more drastic actions to increase the objective value. In this respect, each parameter $\gamma_i$ can
be interpreted as the decision-maker’s confidence level in the law governing class \(i\) arrivals specified in the nominal model. We assume that the decision-maker only considers stationary deterministic policies. Then, it is logical to limit nature to only considering stationary deterministic as well. This is because, after the decision-maker chooses a strategy (which is stationary and deterministic), nature faces a Markov decision process (MDP) in which she aims to optimize a long-run average objective.

To gain deeper insights into how model uncertainty and ambiguity-aversion are captured by Problem (5), we present an alternative formulation in the next section. This alternative formulation shares high-level similarities with classical RO settings, as it imposes direct constraints to restrict the magnitude of the perturbation caused by nature on the nominal model.

5. An Alternative Formulation

Consider the following problem, where the decision-maker seeks to find \((T, \Psi)\) 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 \(\limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta [R^\star_i(t)] \leq \beta_i \) for \(i = 1, \ldots, I\).

This formulation involves \(I\) model-error constraints, where each \(\beta_i\) measures the degree of mistrust that the decision-maker has in the law governing class \(i\) arrivals in the nominal model. A larger value of \(\beta_i\) allows \(\lambda_i\) to deviate more from \(\bar{\lambda}_i\), whereas a smaller value of \(\beta_i\) forces \(\lambda_i\) to remain close to its nominal value. Thus, we can consider \(\beta = (\beta_i)\) as a set of tuning parameters that the decision-maker can use to enhance the robustness of his control strategy.

It is also useful to review Problem (5) as the “Lagrange relaxation” of Problem (6), with the parameters \(\gamma = (\gamma_i)\) being interpreted as the Lagrange multipliers associated with the \(I\) model-error constraints in Problem (6). As a consequence, we will henceforth refer to Problems (5) and (6) as the penalty problem and the constraint problem, respectively. It is natural to wonder if one can establish a Lagrange multiplier theorem that connects these two problems. However, due to the simultaneous minimization and maximization, it is not immediately clear if this is feasible. Nevertheless, we are able to prove such a result with additional regularity conditions. Let \(C^\star_{\text{constraint}}(\beta)\) and \(C^\star_{\text{penalty}}(\gamma)\) denote the optimal values of the constraint problem and penalty problem, respectively. The following result provides a formal connection between \(C^\star_{\text{constraint}}(\beta)\) and \(C^\star_{\text{penalty}}(\gamma)\).

**Proposition 2.** Suppose the decision-maker restricts attention to stationary deterministic policies that prevent queues from growing unbounded, and nature is allowed to choose from the class of stationary policies (deterministic and randomized). Then, for each fixed \(\beta\), if \(\gamma^* \succeq 0\) minimizes \(C^\star_{\text{penalty}}(\gamma) + \langle \beta, \gamma \rangle\), then \(C^\star_{\text{constraint}}(\beta) = C^\star_{\text{penalty}}(\gamma^*) + \langle \beta, \gamma^* \rangle\). Furthermore, for the penalty problem, \(C^\star_{\text{penalty}}(\gamma^*)\) can be achieved when both players adopt stationary deterministic policies.
Remark 1. In Proposition 2, we allow nature to choose from the class of stationary policies because, once the decision-maker selects a strategy (which is stationary and deterministic), nature confronts a constrained Markov decision process (CMDP), where she aims to optimize a long-run average objective subject to some constraints. Furthermore, the requirement for queues not to grow unbounded under the decision-maker’s control, despite being driven by technical considerations, is intuitive. When deciding whether to outsource the production of a product, one must weigh the cost of holding orders in the queue against the cost of outsourcing some of those orders. If there is no limit on queue length, then the holding cost can increase substantially as the queue grows. Hence, there will be a point at which holding orders is no longer cost-effective, and outsourcing becomes more appealing. Our extensive numerical studies confirm this insight; see, e.g., Figures 2 and EC.2.

Proposition 2 demonstrates that it is meaningful to consider and solve the penalty problem. Its proof employs the “convex analytic method,” which was initially introduced in Bhatnagar and Borkar (1995) and further developed in Altman (1999). Although the Lagrange multiplier theorems for the infinite-horizon discounted problem and the finite-time problem were respectively established by Hansen et al. (2006) and Lim and Shanthikumar (2007), we cannot adopt the arguments in their proofs since we tackle a long-run average problem, which is fundamentally different from the infinite-horizon discounted and finite-time problems. The detailed proof is provided in §EC.1.

6. Diffusion Analysis

In light of Proposition 1, we define

\[
 r(\theta) := r^\alpha(\theta) := \sum_{i=1}^{I} \frac{\gamma_i \lambda_i}{\alpha - 1} \{ (1 + \theta_i)^\alpha - \alpha \theta_i - 1 \}
\]

for \( \alpha \neq 1 \) and \( r(\theta) := r^1(\theta) := \sum_{i=1}^{I} \gamma_i \lambda_i \{ (1 + \theta_i) \ln(1 + \theta_i) - \theta_i \} \) for \( \alpha = 1 \). This allows us to express the penalty problem more explicitly, yielding one where the decision-maker aims to minimize

\[
 \max_{\theta} \limsup_{t \to \infty} \frac{1}{t} E^\theta \left[ \sum_{i=1}^{I} \int_0^t c_i(Q_i(u)) \, du + \sum_{i=1}^{I} \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) - \int_0^t r(\theta(u)) \, du \right]. \tag{7}
\]

Problem (7) remains complicated and suffers from the curse of dimensionality as the number of classes increases. Moreover, an exact analysis yields limited structural insights. For these reasons, we advance and solve an SDG that is deemed more tractable than the original penalty problem.

6.1. SDG

The operating 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 “critical-loading” assumption:

\[
 \sum_{i=1}^{I} \rho_i = 1 \quad \text{for} \quad \rho_i := \bar{\lambda}_i m_i, \quad i = 1, \ldots, I. \tag{8}
\]
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 $\theta_i \bar{\lambda}_i$ is an order of magnitude smaller than $\bar{\lambda}_i$, we approximate $A_i$ using

$$A_i(t) = \bar{\lambda}_i t + \bar{\lambda}_i \int_0^t \theta_i(u) du + \hat{A}_i(t) + \epsilon^a_i(t),$$

where $\hat{A}_i$ is a Brownian motion with zero drift and variance parameter $\bar{\lambda}_i$ and $\epsilon^a_i$ is an approximation error term. Define, for each $i$, the centered time allocation process as $Y_i(t) := \rho_i t - T_i(t)$. Note that $\sum_i Y_i(t)$ tracks the cumulative idleness up to time $t$. Similarly, we approximate $S_i \circ T_i$ using

$$S_i(T_i(t)) = \mu_i T_i(t) + \hat{S}_i(t) + \epsilon^d_i(t) = \bar{\lambda}_i t - \mu_i Y_i(t) + \hat{S}_i(t) + \epsilon^d_i(t),$$

where $\hat{S}_i$ is a Brownian motion with zero drift and variance parameter $\bar{\lambda}_i$ and $\epsilon^d_i$ is an approximation error term. Substituting (9) and (10) into (1), ignoring the approximation error terms, and replacing the random elements $Q_i, Y_i,$ and $\xi_i$ with their respective approximations, namely, $\hat{Q}_i, \hat{Y}_i$ and $\hat{\xi}_i$, we obtain

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

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

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

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

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

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

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

subject to constraints (11) – (13).

### 6.2. 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, leading 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 (11) 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 define

\[
B(t) := \sum_{i=1}^{I} m_i \hat{Z}_i(t), \quad \zeta(t) := \sum_{i=1}^{I} \rho_i \hat{\theta}_i(t) \quad \text{and} \quad O(t) := \sum_{i=1}^{I} m_i \sum_{k=0}^{N_i(t)} \hat{\xi}_i(k). \quad (15)
\]

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

\[
h(w) := \min \left\{ \sum_{i=1}^{I} c_i(x_i) : m^\top x = w, x \in \mathbb{R}_I^I \right\}.
\]

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

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

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 \( \tilde{\xi}_i(k) := m_i \tilde{\xi}_i(k) \) 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

\[
\max_{\zeta} \lim_{t \to \infty} \frac{1}{t} \mathbb{E}^\zeta \left[ \int_0^t h(W(u))du - \int_0^t r^*(\zeta(u))du + \sum_{i=1}^{I} \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right], \quad (18)
\]
\[ \text{s.t. } W(t) = W(0) + B(t) + \int_0^t \zeta(u) \, du + U(t) - \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \xi_i(k), \]  
\[ U(t) \text{ is non-decreasing with } U(0) = 0, \]  
\[ W(t) \geq 0 \quad \text{for } t \geq 0, \]  

where the superscript \( \zeta \) in the expectation operator means that there is a state-dependent perturbation \( \zeta \) to the drift of the underlying process.

**Proposition 3.** The SDG (11)–(14) is equivalent to the workload problem (18)–(21) in the following sense: every admissible control \((\hat{Y}, \hat{\Psi})\) for the SDG yields an admissible policy \((U, \tilde{\Psi})\) for the workload problem, and these two policies share the same cost. Furthermore, for any admissible policy \((U, \tilde{\Psi})\) of the workload problem, there exists an admissible policy \((\hat{Y}, \hat{\Psi})\) for the SDG, and its cost is no less than that of the policy \((U, \tilde{\Psi})\) for the workload problem.

### 6.3. Characterization of the Optimal Solution

Because nature’s drift-rate control only depends on the current workload, 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

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

#### 6.3.1. 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

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

subject to the boundary conditions

\[ f'(0) = 0 \quad \text{and} \quad f(s) = \tilde{\phi}_i(s - q) + f(q). \]
The proposition presented below offers a useful identity that motivates the optimality equation described in the subsequent subsection. The proof of this identity is a routine application of Itô’s formula, which we omit here.

**Proposition 4.** Suppose $\eta \in \mathbb{R}$ and $f \in C^2[0,s]$ jointly satisfy (22) and (23). Then $\eta$ is the long-run average cost under the control band policy $(i,q,s)$ and the drift-rate control $\zeta(\cdot)$.

### 6.3.2. 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 $(i,q,s)$, 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 $(i^*,q^*,s^*)$ 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 $(i^*,q^*,s^*)$, 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 4 motivates the following optimality equation that facilitates the identification of the control band policy $(i,q,s)$ as mentioned earlier: find $q_i, s_i, \eta_i \in \mathbb{R}$ and $v \in C^2[0,s_i]$ such that

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

subject to the boundary conditions

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

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

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

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

we can rewrite (24) as

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

Because (26) 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

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

(27)
subject to the boundary condition

$$\pi(0, \eta) = 0.$$  \hfill (28)

The parameter pair \((q_i, s_i)\) and the average cost \(\eta_i\) are determined through conditions:

$$\pi(q_i, \eta_i) = \pi(s_i, \eta_i) = \tilde{\ell}_i \quad \text{and} \quad \int_{q_i}^{s_i} \pi(w, \eta_i) \, dw = \tilde{\phi}_i(s_i - q_i).$$  \hfill (29)

The two equations, (28) and (29) effectively make up four constraints. However, two questions remain unanswered. First, does the system of equations given in (27)–(29) yield a solution (i.e., suffice to pin down the four unknowns, \(q_i, s_i, \eta_i\) and \(\pi(\cdot, \eta_i)\))? 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. Based on the definition of \(h\) and Assumption 1, it follows that \(h(w)\) is a continuous and strictly increasing function over the interval on \(\mathbb{R}_+\), with \(h(0) = 0\). Additionally, \(h(w)\) goes to infinity as \(w\) approaches infinity. Our next result not only gives a positive answer to the first question but also provides key ingredients for finding answers to the second one.

**Proposition 5.** Suppose Assumption 1 holds. Then the following statements are true. (i) The requirements in (27)–(29) uniquely determine \(q_i, s_i,\) and \(\eta_i\). (ii) If we define \(v(\cdot)\) such that its first-order derivative is equal to \(\pi(\cdot, \eta_i)\) on the interval \([0, s_i)\), and \(v(w) = v(s_i) + (w - s_i)\tilde{\ell}_i\) for \(w \geq s_i\), then the pair \((v, \eta_i)\) satisfies the following quasi-variational inequality:

$$\min \left\{ \frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_i, \inf_{0 \leq z \leq w} \left[ v(w - z) + \tilde{\phi}_i(z) \right] - v(w) \right\} = 0.$$  \hfill (30)

(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 (30). Thus, our detailed construction of the solution to (30) 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 \(v'(0)\). Here, since we consider a “one-sided” impulse control, the left boundary of \(v'(0)\) is fixed at zero. The detailed proof of Proposition 5 is given in §EC.2.

According to Proposition 5, for each fixed \(i\), one can appeal to (27)–(29) 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 type \(i^* \in \{1, \ldots, I\}\) such that \(\eta_{i^*} \leq \eta_i\) for all \(i \neq i^*\). Intuitively, the resulting control band policy \((i^*, q_{i^*}, s_{i^*})\) 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. It is natural to wonder if this rule mini-maximizes the long-run average cost among all adaptive controls that the decision-maker can possibly take. To verify this, motivated by Proposition 5 and its proof, we aim to find a function $v$ that is twice differentiable almost everywhere, has a bounded and continuous first-order derivative, and satisfies

$$\min \left\{ \frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_{i^*}, \min_i \inf_{0 \leq z \leq w} \left[ v(w - z) + \phi_i(z) \right] - v(w) \right\} \geq 0 \quad (31)$$

with $v'(0) = 0$. In fact, in Theorem 1 below, we establish that if such a function $v$ can be found, then the control band policy $(i^*, q_{i^*}, s_{i^*})$ is overall optimal under the min-max criterion.

**Theorem 1.** If some function $v$, which is twice differentiable almost everywhere and has a bounded first-order derivative, satisfies (31) with $v'(0) = 0$, then the control band policy $(i^*, q_{i^*}, s_{i^*})$ is average cost optimal under the minimax criterion among all adaptive controls.

To ensure the practical value of the result presented in Theorem 1, it is necessary to demonstrate the construction of a function $v$ that satisfies all the specified properties. In order to achieve this, we let $\ell_* := \min_i \ell_i$, and define $(s_*, \ell_*)$ as the coordinates of the point where the curve of $\pi(\cdot, \eta_{i^*})$ intersects the constant function $\ell_*$ for the second time. By making these preparations, we can now state Proposition 6, which guarantees the existence of a function $v$ possessing the desired properties.

**Proposition 6.** Let $v$ be such that its first-order derivative is equal to $\pi(\cdot, \eta_{i^*})$ on the interval $[0, s_*$), and $v(w) = v(s_*) + (w - s_*)\ell_*$ for $w \geq s_*$. Then $v$ satisfies (31).

The detailed proofs of Theorem 1 and Proposition 6 can be found §EC.2. We will utilize these results to develop practical policy recommendations in §6.4. In closing this subsection, we provide a remark concerning the fixed outsourcing costs.

**Remark 2.** When all fixed costs are zero, outsourcing controls become singular controls, and one can read off the cheapest product to outsource when the backlog of work is considered excessive. While this simplified problem resembles the one in Cohen (2019) that also involves a right boundary playing a similar role to $s^*$, our right boundary $s^*$ is a policy parameter that has yet to be optimized. In contrast, the right boundary in Cohen (2019) is predetermined, corresponding to a buffer that leads to job rejection. When $L_i$ equals $L$ for all $i$, which is typically the case when all outsourcing goes to the same contract manufacturer, it is easy to see that the class $i$ with the smallest $\ell_i$ value will be the cheapest class to outsource. Let $i^*$ denote the cheapest outsourcing class. Then, in order to find the solution to the workload problem, it suffices to solve (30) with $i = i^*$. 
6.4. Policy Recommendations

We propose an implementable control rule based on the solution to the workload problem. Here we take both $\alpha$ and $\gamma$ as given and fixed. Using the method described in §6.3.2, one can calculate $(i^*, q_{i^*}, s_{i^*})$ and $\pi(\cdot, \eta_{i^*})$. For ease of implementation, if $q_{i^*}$ and $s_{i^*}$ are not integers, they should be rounded to the nearest integers. A procedure regarding how to select $\gamma$ for a given $\alpha$ in a practical setting efficiently will be the focus of Section 8.

To avoid introducing new notation, we overload the notation $W(t)$ to let it now denote the workload of the actual system at time $t$, so that $W(t) = \sum_i m_i Q_i(t)$, where $Q_i(t)$ is the actual number of class $i$ orders awaiting processing at time $t$ and $m_i$ is the mean class $i$ service time.\(^1\) The proposed control has two components, as described below.

**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 enough $o_{i^*}$ 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 $(x_{i^*}^\star)$ to the optimization problem (16) 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 $(x_{i^*}^\star)$ 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 $(x_{i^*}^\star)$. 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 $t$.

7. Discussion

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

7.1. Connections to Harrison’s Framework

Here, we briefly explain how to scale different model primitives if we were to directly adopt the framework pioneered by Harrison (1988), which requires introducing a sequence of problems that are

---

\(^1\) Previously, we have used $W$ to denote the approximate workload process.
indexed by a scaling parameter, denoted as $n$. In our study, the model primitives include long-run average arrival rates $\bar{\lambda}_i$, class-specific service rates $\mu_i$, outsourcing cost parameters $L_i$ and $\ell_i$, as well as cost rate functions $c_i$ and $r$ as given in (7).

In our context, one can think of the scaling parameter $n$ as the sum of the nominal arrival rates $\bar{\lambda}_i$, and let it grow to infinity so as to create a sequence of models while ensuring condition (8) holds so that we need each $\mu_i$ to grow proportionally to $n$. We also need to scale the cost parameters associated with outsourcing operations and the cost rate functions $c_i$ and $r$. For the decision-maker, the crux is the trade-off between waiting costs and outsourcing costs. To ensure this trade-off is nontrivial, these two costs need to be of the same order of magnitude. This can be achieved by leaving the fixed outsourcing costs unscaled but scaling down the proportional outsourcing costs by a factor of $\sqrt{n}$. Meanwhile, we let the waiting cost rate function in the $n$-th system, denoted as $c^n_i$, be scaled so that $c^n_i(\cdot) = \hat{c}(\cdot/\sqrt{n})$, where $\hat{c}$ is a baseline function independent of the scaling parameter $n$. Further, let nature’s penalty rate function in the $n$-th system, denoted as $r^n$, be scaled in such a way that $r^n(\cdot) = \hat{r}(\sqrt{n}\cdot)/n$, where $\hat{r}$ is a baseline function independent of the scaling parameter.

The scaling condition imposed on the penalty rate function implies that nature’s perturbations are moderate, with an order of $1/\sqrt{n}$. This, in turn, implies that the system will remain critically loaded for a sufficiently large value of $n$, because of the condition (8) and that nature’s perturbations are moderate. Then standard heavy-traffic approximation theory suggests that the queue lengths will be of the order of $\sqrt{n}$, which is consistent with the scaling condition imposed on the waiting cost rate functions. Moreover, according to the dynamic equation (1), the outsourcing batch size in the $n$-th system, denoted as $\xi^n$, should be of the same order of magnitude as the queue length. This means that the outsourcing batch size will also be around $\sqrt{n}$, which is compatible with the scaling condition imposed on the proportional outsourcing costs.

### 7.2. About the Proposed Control Rule

The derivation of the SDG assumes a critically loaded system with high demand and service rates. Under this assumption, the robust control problem simplifies significantly, reducing the state space dimension from $I+1$ to one. This enables our solution to decompose over two time scales.

On the faster time scale, we optimize the distribution of workload among classes based on the current total workload. While instantaneous redistribution is not practical (unlike in the SDG), adjustments to queue lengths can be made relatively quickly, at a rate of approximately $\sqrt{n}$ (if we were to borrow the terminology from §7.1). If queue lengths fall short of targets, the corresponding classes temporarily lose server access, causing queues to only receive inflow from demand arrivals. As demand arrivals occur at a rate of order $n$, significantly higher than the queue lengths of order $\sqrt{n}$, actual queue lengths rapidly return to targets at a rate of order $\sqrt{n}$. By the same token, classes...
whose actual queue lengths exceed the targets will have exclusive access to the total service capacity, thereby seeing “an underloaded system” due to the critical-loading condition (8) and the fact that nature’s distortions are moderate. In particular, as the net outflow (the actual outflow from these queues minus the inflow) is of order $n$ for these classes, their queue lengths will also quickly return to their targets, at a rate of order $\sqrt{n}$.

On the slower time scale, where the total workload evolves, we solve an impulse control problem to determine when and to what extent to reduce the workload, giving rise to the proposed outsourcing rule. This rule allows for the postponement of outsourcing until a sufficient number of class $i^*$ orders are received, even if the total workload has already exceeded $s_{i^*}$. The delay caused by postponement is expected to be short if the threshold is small. Larger thresholds may result in longer delays, potentially defeating the purpose of reducing an excessive workload, as orders in the virtual queue will continue to accrue holding costs. However, in heavy traffic, long delays are rare even with large thresholds.

To briefly explain, the control parameters $q_{i^*}$ and $s_{i^*}$ should be of order $1/\sqrt{n}$, as they are used to regulate the workload, which is also of order $1/\sqrt{n}$. Furthermore, the definition of $o_{i^*}$ implies that it is of order $\sqrt{n}$, given that $m_{i^*}$ is of order $1/n$. Since the virtual queue faces a demand rate of roughly $\bar{\lambda}_{i^*}$, which is of order $n$, it follows that the virtual queue will be filled up relatively quickly to the desired target $o_{i^*}$ at a rate of order $\sqrt{n}$.

One caveat is that accumulating enough outsourceable orders may be slow if the cheapest class to outsource has very low demand. In such cases, this class is deemed to have a “thin arrival,” as described in Ata (2006)—the usual assumption that the demand rate for each class is of order $n$ does not hold. Consequently, when formulating the diffusion approximation and the corresponding workload problem, the role of this class would be limited to contributing to the drift rate in the diffusion approximation, as discussed in Ata (2006). As a result, this class is not considered a “legitimate class” for outsourcing, potentially resolving the issue.

7.3. Selecting the Uncertainty Set

One primary motivation for robust control is to strike a balance between tractability and practicality. This is accomplished by employing a simple nominal model and establishing an uncertainty set around it. The selection of an appropriate uncertainty set is crucial because a set that is too small may exclude the true model, while an excessively large set can result in overly conservative solutions. In our framework, specifying the uncertainty set involves determining the suitable $\alpha$, which defines the type of uncertainty set, and $\gamma$, which controls its size.

To aid in the selection of the uncertainty set, we draw inspiration from regularized regression in machine learning. In regularized regression, the norm for the penalty term and the regularization parameter must be specified. The choice of norm depends on the desired properties, such as using $L1$
norm for sparsity or L2 norm for continuity. Combining these norms (Elastic Net) is also an option, introducing an additional tuning parameter that shares conceptual similarities with the role of $\alpha$ in our study. Hence, we regard the selection of $\alpha$ as more subjective. In §EC.7, we provide examples illustrating how decision-makers can subjectively choose the value of $\alpha$ based on their prior knowledge of the actual demand rates.

In regularized regression, the regularization parameter is often considered more critical and is typically chosen through cross-validation. In our framework, the vector $\gamma$ plays a role analogous to the regularization parameter in penalized regression, influencing the resulting control rule. This motivates the need to propose a mechanism for selecting $\gamma$ similar to cross-validation. By varying $\gamma$, we can generate a set of control rules and assess their effectiveness through computer simulation. The best uncertainty set corresponds to the $\gamma$ vector that yields the lowest estimated cost from the simulation.

In Section 8, we propose a simulation-based method for locating the “best” $\gamma$ vector.

### 8. A Simulation-based Method for Uncertainty Set Selection

We now demonstrate how to select the parameter $\gamma$. As alluded to previously, we will reduce the significance of the “shape” parameter $\alpha$ by treating it as a fixed value. However, it is important to note that changing the value of $\alpha$ allows for greater flexibility in designing the uncertainty set. In §EC.7, we discuss this matter further, explaining why our simulation-based uncertainty set selection method mainly focuses on choosing the value of $\gamma$.

Although the real-world system may be complex, it often allows for a simplification that can serve as the nominal model used in the formulation of the robust control problem. With each fixed vector $\gamma$, the robust control problem leads to a Bellman-Isaacs equation that can be solved to obtain a robust control rule, denoted as $P(\gamma)$. Assume optimistically that when the control rule $P(\gamma)$ is applied to the “true” model, a long-run average cost, represented by $C(\gamma)$, is obtained. The best uncertainty set is then identified by choosing a vector $\gamma$ that minimizes $C(\gamma)$. If the expression for $C(\gamma)$ is known, a natural approach is to use a gradient-descent algorithm. However, finding such an expression can be challenging. On the other hand, the function value of $C$ can be evaluated efficiently using computer simulations. This motivates our proposed algorithm, which is similar to a standard gradient-descent 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 illustrates how to find the best value of $\gamma$ for a fixed $\alpha$. In the algorithm, $\kappa$ represents the learning rate, and $\gamma^{(j)} := \{\gamma_1^{(j)}, \ldots, \gamma_I^{(j)}\}$ represents the $\gamma$ value at the end of the $j$th outer-iteration.

---

2 Here, we take an optimistic view because, without a formal specification of the “true” demand model, there is no guarantee that the “true” model, along with the control derived from the SDG with the penalty parameter $\gamma$, has a well-defined long-run average cost. This contrasts with settings where the “true” demand model is explicitly defined, such as Poisson or some Markov process. In those cases, it becomes possible to formulate the problem as an MDP for which established theory exists, including regularity conditions that can ensure the well-posedness of the problem.
\( \hat{C}(\gamma) \) denotes the simulation estimate of \( C(\gamma) \). By selecting a small \( \delta \in \mathbb{R}_+ \), we approximate the \( i \)th component of the gradient of \( C \) at the point \( \gamma^{(j)} \) using

\[
\nabla C_i(\gamma^{(j)}) := \hat{C}(\gamma^{(j)} + \delta e_i) - \hat{C}(\gamma^{(j)} - \delta e_i) / 2\delta
\]

where \( e_i \) denotes the unit vector whose \( i \)th component is one and the remaining components are zero. We present two 2-class examples in §9.2 and a 4-class example in §EC.5 that demonstrate the efficiency of our simulation-based method.

Algorithm 1 Quasi-Gradient-Descent Algorithm

1: initialize: \( \kappa, \gamma^{(0)}, \gamma^{(1)}, \epsilon \)

2: while \( \gamma^{(j+1)} \neq \gamma^{(j)} \) do

3: \( \tilde{\gamma}^{(j)} \leftarrow \gamma^{(j)} \)

4: for \( i = 1 : I \) do

5: while \( \left| \nabla C_i(\tilde{\gamma}^{(j)}) \right| > \epsilon \) do

6: \( V^{(j)} \leftarrow \tilde{\gamma}^{(j)} \)

7: \( \tilde{\gamma}^{(j)}_i \leftarrow \max\left\{ \tilde{\gamma}^{(j)}_i - \kappa \nabla C_i(\tilde{\gamma}^{(j)}) , 0.02 \right\} \) \( \triangleright \) Maximum ensures positiveness.

8: \( \kappa \leftarrow \frac{\left| (V^{(j)} - \tilde{\gamma}^{(j)}) \right|}{\left( \nabla C_i(V^{(j)}) - \nabla C_i(\tilde{\gamma}^{(j)}) \right)^2} \)

9: end while

10: end for

11: \( \gamma^{(j+1)} \leftarrow \{ \tilde{\gamma}^{(j)}_1, \ldots, \tilde{\gamma}^{(j)}_I \} \)

12: end while

The algorithm for finding the best \( \gamma \) relies on simulations, and the warm-up period is a crucial parameter that requires careful selection to ensure reliable and accurate results. The warm-up period represents the initial simulation period needed for the system to reach a steady state. A general guideline is to choose a warm-up period that is sufficient for achieving a steady state but not excessively long, consuming a significant portion of the total simulation time. In our numerical studies, we determine the appropriate warm-up period through a naive “sensitivity analysis.” This involves comparing statistical estimates of the desired output variable (long-run average cost) under two candidate warm-up periods, \( T_1 \) and \( T_2 \), where \( T_1 < T_2 \). If there is no noticeable difference between the estimates, we adopt \( T_2 \) as the warm-up period.

The number of replications, which determines how many times the simulation model is executed, is another important parameter requiring judicious selection. A higher number of replications improves accuracy but increases computational time. One approach to determine the appropriate number
of replications is by measuring the variability of the output variable. After each replication, the confidence interval of the desired output variable can be computed, and the iteration can be stopped when the width of the confidence interval falls below a predetermined threshold. The threshold should be chosen small enough to ensure sufficient accuracy in the final result.

Our framework consists of two main components: an analytical component and a simulation-based evaluation component. The analytical component, developed through Sections 4–6, involves a nominal model, an uncertainty set, and an efficient approach for computing candidate control rules. The simulation-based evaluation component assesses the performance of these candidate control rules in a high-fidelity environment and aims to identify the rule with the lowest long-run average cost based on simulation outputs. Compared to a pure simulation-based optimization method, our framework offers three key advantages. First, the analytical component provides a clear set of control rules as input for the simulation-based evaluation, giving the simulation step a specific objective. In contrast, a direct simulation-based optimization may lack a clear focus on candidate policies. Second, the control rules generated by the analytical component possess interpretable structural properties, providing insights into the underlying mechanisms. These insights can be challenging to obtain from a pure simulation-based approach. Lastly, a robust control rule is less sensitive to the calibration of the “true” model, unlike a solution obtained directly from simulation-based optimization, especially for complex system models.

9. Numerical Studies
This section presents various numerical studies. Throughout this section, by mentioning the solution to the SDG, we mean the solution to the workload problem. In §9.1, we compare the solution obtained from the original penalty problem with that from the SDG. In §9.2, we present examples demonstrating the “value of robustness.” In §9.3, we compare the optimal cost attained using the actual demand model (assuming the “true” model is known to the decision-maker) with the optimal robust cost achieved by our proposed method.

9.1. Solution Comparison Between the Penalty Problem and the SDG
In this subsection, we numerically solve the original penalty problem and its diffusion approximation and compare these solutions to demonstrate the accuracy and reliability of the diffusion approximation. Specifically, for the SDG, we find its solution via the associated optimality equation (24)–(25). Implementation details can be found in §EC.3. For the original penalty problem, whose state-descriptor can be found in §EC.1, we adopt the “strategy iteration” approach. In each iteration, we fix the strategy of one player and compute the best response of the opponent in the resulting one-player game (MDP); this can be achieved via standard numerical schemes for solving MDPs, and we use value iteration to find the best response of the opponent. Then the strategy we originally fixed is
improved based on the response of the opponent. As an initialization, we solve the nominal model, which assumes that all arrival rates are exactly equal to their nominal values, hence no participation from nature, allowing us to obtain a strategy for the decision-maker to start with. While we do not intend to establish a formal theory guaranteeing the convergence of the strategy iteration algorithm, our experiments with numerous instances suggest that the algorithm shows quick convergence.

We consider a two-class model with the following parameters: $\bar{\lambda}_1 = 30$, $\bar{\lambda}_2 = 40$, $\mu_1 = 60$, and $\mu_2 = 80$. The cost data includes fixed outsourcing costs $L_1 = 5$ and $L_2 = 15$, proportional outsourcing costs $\ell_1 = \ell_2 = 1.0$, and quadratic holding cost rates $a_1 = 0.4$ and $a_2 = 0.5$. Figure 1 illustrates the optimal sequencing rules under Rényi divergence with $\alpha = 1/2$. For the original problem, the decision-maker would prioritize class 1 (class 2) when the system is in states represented by a blue circle (a green square). For the SDG, the decision-maker should prioritize class 2 if the state is above the gray line and class 1 if it is below the gray line. The result shows that the sequencing rules derived from the two different approaches are very close, validating the near-optimality of the generalized $c\mu$ rule.

![Figure 1](image)

(a) Class 1 is in service
(b) Class 2 is in service

**Figure 1** Sequencing strategies derived from the penalty problem and the SDG with $\gamma_1 = \gamma_2 = 30$ when $\alpha = 1/2$

Results from Figure 2 pertain to outsourcing. 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. Each horizontal or vertical arrow in the figures connects the “origin” (the state right before outsourcing) and the “destination” (the state right after outsourcing) associated with an outsourcing operation. The decision-maker is inclined to outsource orders of class 1, unless queue 2 becomes exceedingly large, which is in alignment with the control strategy prescribed by the SDG. The green lines represent the outsourcing thresholds, $q_1$ and $s_1$, derived from the diffusion approximation and indicate that the decision-maker should always outsource
product 1. Hence, we can conclude that the outsourcing rule derived from the diffusion approximation is reasonably close to that obtained from the penalty problem.

In §EC.4.1, we present an additional example involving KL divergence ($\alpha = 1$). Therein we also compare the long-run average cost generated by the original penalty problem ($\eta_{\text{exact}}$) and that from the SDG ($\eta_{\text{SDG}}$) under varying levels of model uncertainty for both $\alpha = 1/2$ and $\alpha = 1$. The results show that the long-run average costs derived from the SDG are very close to the values attained from the original penalty problem, highlighting the reliability of the diffusion approximation.

9.2. Exposing the Value of Robustness

In this subsection, we showcase the importance of incorporating robustness into decision-making. Specifically, the system in §9.2.1 assumes the demand rate of each product to follow an auto-regressive integrated moving average (ARIMA) model, whereas the system in §9.2.2 assumes the demand rate of each product to follow a continuous-time Markov chain (CTMC). Simulations are conducted over a time interval of 2,000 for the stochastic system. In order to estimate the performance of the system, we test initial warm-up periods with $T_1 = 50$ and $T_2 = 100$. After comparing the statistical estimates of the long-run average simulated cost, we adopt $T_2 = 100$ as the warm-up period. In addition, to determine the appropriate number of replications, we set the desired width of the confidence interval to 0.2. After conducting experiments, we observe that each example requires approximately 100 replications to achieve the desired width of the confidence interval.

9.2.1. ARIMA Intensity

We demonstrate the simulated average cost $\hat{C}(\gamma)$, assuming a real-world demand model in which the demand rate of each product follows a non-homogeneous Poisson process with an ARIMA model for intensity. We study a two-class make-to-order system with the

![Figure 2](image-url)
following parameters: $\lambda_1 = 80$, $\lambda_2 = 60$, $\mu_1 = 100$, and $\mu_2 = 300$. The cost data includes outsourcing costs $L_1 = 0.5, L_2 = 0.8, \ell_1 = \ell_2 = 0.2$, and quadratic holding cost rates $a_1 = 0.01$ and $a_2 = 0.02$. In addition, for both products, we set the two “ARLags” to 0.8 and −0.8 respectively, the MA error coefficient to 0.4 and the variance to 100.

The quadratic holding costs give rise to a queue-ratio type sequencing rule: if $a_1 \mu_1 Q_1 > (<) a_2 \mu_2 Q_2$, priority should be given to producing product 1 (or product 2 if the inequality is reversed), with ties being broken arbitrarily. 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 $(q_1, s_1)$ for different pairs of $(\gamma_1, \gamma_2)$ can be found in Table EC.2. Figure 3 depicts the results of our algorithm, with $\hat{C}(\gamma)$ presented under both Rényi ($\alpha = 3$) and KL divergence. The black arrows indicate each step taken during the inner-iterations, while the blue arrows correspond to the final step of each inner-iteration. The orange arrow denotes the final step of the whole algorithm, while the termination point is marked with an orange circle, indicating the optimal value of $\gamma$ and the associated simulated cost. It is worth noting that when $\gamma_1$ and $\gamma_2$ are very large (thus, $1/\gamma_1$ and $1/\gamma_2$ are very small), the corresponding control rule can be viewed as one that completely ignores model uncertainty. This implies that the difference between the “limiting value” of each plot and the minimum value on that plot can be seen as the value of robustness. For instance, for $\alpha = 3$, the value of robustness is approximately $2.5174 - 2.3879 = 0.1295$, which is around 5.42% better than completely ignoring model uncertainty. Similarly, for $\alpha = 1$, the value of robustness is approximately
0.0859, which is roughly 3.54% better than ignoring model uncertainty. The comparison between the two values of robustness explicitly demonstrates that a flexible $\alpha$ can lead to cost savings. Table EC.3 in §EC.4.2 displays confidence intervals derived from all computer simulations of the point estimates in this example. Nevertheless, we employed our quasi-gradient-descent algorithm and discovered that it is capable of locating the minimum value. This minimum value closely corresponds to the lowest cost presented in Table EC.3.

**9.2.2. CTMC Intensity** We now explore an arrival model in which the demand rate for each order adheres to a non-homogeneous Poisson process, and its intensity is represented as a CTMC. Specifically, for each product, the arrival intensity is modeled as a CTMC with two distinct states, denoted by $\tilde{\lambda}_i$ and $\hat{\lambda}_i$. The demand rate $\tilde{\lambda}_i$ corresponds to an optimistic scenario, whereas $\hat{\lambda}_i$ represents a pessimistic scenario. The nominal demand $\bar{\lambda}_i$ is computed as the weighted sum of the two demand rates. The system parameters are the same as described in §9.2.1. In addition, in the CTMC, each state’s sojourn time is exponentially distributed with a rate of 10. We further set $\tilde{\lambda}_1 = 150$ and $\tilde{\lambda}_2 = 110$, while $\hat{\lambda}_1 = \hat{\lambda}_2 = 10$. Note that, unlike in §9.2.1, we purposely amplify the fluctuations of real-world demand rates beyond what heavy-traffic scaling would entail. This deliberate choice aims to showcase the versatility of our framework in different operating regimes. Although we could have set the parameters to yield moderate fluctuations, thereby adhering closely to heavy-traffic scaling requirements, as will be demonstrated numerically in §9.3, our proposed approach remains effective even when confronted with substantial fluctuations in real-world demand rates.

![Figure 4](image-url) Quasi-gradient-descent algorithm for simulations with CTMC intensity
In this example, the optimal decision rules (outsourcing and sequencing) are also the same as in the previous example. The path taken by the algorithm to locate the minimum value is illustrated in Figure 4. The value of robustness for $\alpha = 3$ is approximately 0.7608, which represents a 9.07% improvement over fully ignoring model uncertainty. For $\alpha = 1$, the value of robustness is approximately 0.5137, which is approximately 5.99% better than ignoring model uncertainty. Again, these results demonstrate that the flexibility offered by $\alpha$ can result in additional cost savings compared to using KL divergence alone. Table EC.4 in §EC.4 provides confidence intervals obtained from all computer simulations.

The main takeaway from Figures 3 and 4 is that employing a robust control formulation that accounts for ambiguity can result in significant cost savings. Completely ignoring model uncertainty and following the policy based on the nominal model can result in penalties for inadequate preparation for model misspecification. Conversely, an excessively conservative solution may be produced by overemphasizing the worst-case scenario, which can occur if the uncertainty set is chosen too large. A proper level of ambiguity, meaning an appropriate choice of the uncertainty set, can safeguard against model errors without generating overly conservative solutions.

### 9.3. Cost Comparison Between Actual and Robust Models

In this subsection, we compare the cost of the “best” robust control policy, denoted by $\hat{C}(\gamma^*)$, with the actual optimal cost $C^*$ where we assume that the decision-maker has perfect knowledge of the “true” model. For the “true” model, we use a model setting similar to the one outlined in §9.2.2 (CTMC intensity). To determine the optimal cost $C^*$, we must solve the “true” model, which is essentially an MDP, through value iteration. However, solving this problem with value iteration is computationally intensive, as the two-class example results in a five-dimensional MDP, including the two queue lengths, the index of the class being served, and the current arrival rates of the two classes. Therefore, we opt to address a single-class problem instead.

We set $\bar{\lambda} = \mu = 100$ and consider the following cost parameters: outsourcing parameters $L = 5, \ell = 1.0$, and quadratic holding cost rate $a = 0.4$. In the CTMC, each state’s sojourn time is exponentially distributed with a rate of 10. Additionally, we set $\bar{\lambda} = 150$ and $\hat{\lambda} = 50$. In Figure 5, the black straight line at the bottom denotes the actual optimal cost $C^*$, obtained by solving the “true” model through value iteration. The blue curve illustrates the long-run average robust cost $\hat{C}(\gamma)$ for varying $\gamma$. Remarkably, the difference between $\hat{C}(\gamma^*)$ and $C^*$ is only 2.45%. If the decision-maker were to ignore model uncertainty, he would incur a cost that is 6.95% higher than $C^*$. Therefore, our robust control formulation, accompanied by the parameter-determining method, delivers a noticeable value.

For detailed information on the outsourcing thresholds $(q,s)$ derived from the SDG, as well as the simulated costs in the above example and an additional example, please refer to §EC.4.3.
Figure 5  Actual cost and cost obtained from the “best” robust model with $\tilde{\lambda} = 150$ and $\hat{\lambda} = 50$

10. 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. We present one main formulation that involves nature as a second player that promotes robustness to model misspecification. The formulation can be interpreted as a two-player, zero-sum game. By considering the system in a suitable operating regime, we reach an approximating problem that can be further turned into a one-dimensional stochastic game. By solving the one-dimensional stochastic game, we find that the optimal control strategy for the decision-maker is of a control-band form, whereas the optimal strategy for nature is a state-dependent drift-rate control of the workload process. This, combined with simulation programs, leads to a procedure that can produce a useful joint outsourcing and sequencing rule. The effectiveness of the procedure is validated through numerical work.

Several limitations of this study should be acknowledged. First, our solution approach is based on an asymptotic framework that assumes heavy traffic conditions (where total demand is approximately equal to production capacity). Therefore, the analysis, results, and insights primarily apply to systems that satisfy heavy-traffic conditions. Exploring effective solutions for “overloaded” systems, where nominal demand significantly exceeds capacity, would be valuable. In Section EC.6.3, we present a formulation expected to handle slightly overloaded systems and provide insights on whether solutions derived from heavy-traffic conditions can be effective for controlling significantly overloaded systems with a single class. Second, we do not consider the incorporation of online learning to update or refine
the uncertainty set based on new information as the system evolves over time. Combining online learning with a robust control framework can be valuable, especially when model simplification arises due to a lack of historical data rather than tractability considerations. Recent works, such as Kim (2016), have embraced this idea. Third, it is unclear to what extent the proposed framework can be applied to more complex systems that are potentially difficult to simulate. For example, simulating a system with many potentially heterogeneous servers could pose practical challenges compared to the make-to-order system considered in this paper, where simulating one server is sufficient. Further investigation into this matter would be worthwhile.

While our modeling approach enables real-time decision-making using a simplified model, it is important to note that there are other methods employing meaningful “model simplifications” to solve complex decision-making problems. For instance, in the context of performance optimization for assembly-to-order systems, which share similarities with make-to-order systems, Nadar et al. (2018) have advanced a novel “state aggregation” method. This method significantly reduces the computational burden while still producing approximate solutions with very small optimality gaps. It would be interesting to investigate how our approach fares in comparison to such methods.

References


**E-Companion**

The e-companion is organized as follows: In §EC.1, we provide the proof of Proposition 2. In §EC.2, we prove other main results. In §EC.3, we describe our numerical scheme used for solving the Bellman-Isaacs equation. We supply additional data on the numerical experiments in §EC.4 and present the numerical results for a 4-class example in §EC.5. In §EC.6.3, we present extensions for slightly overloaded or underloaded systems. We offer some high-level guidance on the choice of the value of $\alpha$ in §EC.7 and provide proofs for some auxiliary results in §EC.8.

**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 nature’s decision problem as a discrete-time 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 $\mathcal{S}$, but also allows for the division of $\mathcal{S}$ into two sets: $\tilde{\mathcal{S}}$, which includes all the states that do not trigger outsourcing, and $\bar{\mathcal{S}}$, which includes all the states that will trigger outsourcing. Clearly,

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

Moreover, for each $x \in \tilde{\mathcal{S}}$, 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 $x := (q, j) \in \tilde{\mathcal{S}}$, we can define two sets:

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

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 $\tilde{\mathcal{S}}$. To spell out the transition law of this discrete-time, suppose $X(n) = (q, j)$ 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 $1 - \frac{\mu_j + \sum_{i=1}^I \lambda_i(n)}{\nu}$,

$$X(n + 1) = (q, j).$$

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)):

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

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:

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

where $x = (q, j)$. Therefore, nature’s problem can be cast into a CMDP that seeks $\lambda$ to maximize

$$C_{ca}(X, \lambda) := \limsup_{m \to \infty} \frac{1}{m} \mathbb{E} \left[ \sum_{n=1}^m \tilde{c}(X(n), \lambda(n)) \right]$$
subject to

\[ D_i(\lambda) := \limsup_{m \to \infty} \frac{1}{m} \mathbb{E} \left[ \sum_{n=1}^{m} d_i^n(\lambda(n)) \right] \leq \beta_i, \quad i = 1, \ldots, I, \quad \text{(EC.1)} \]

where

\[ d_i^n(\lambda) := \begin{cases} \frac{\lambda_i}{\nu(\alpha-1)} \left\{ \left( \frac{\lambda_i}{\lambda} \right)^\alpha - \alpha \left( \frac{\lambda_i}{\lambda} \right) - 1 \right\} & \text{for } \alpha \neq 1, \\ \frac{1}{\nu} \left\{ \lambda_i \ln \left( \frac{\lambda_i}{\lambda} \right) - \lambda_i + \bar{\lambda}_i \right\} & \text{for } \alpha = 1. \end{cases} \]

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 \( \lambda \) and the set of stationary deterministic \( \lambda \).

**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

\[ \max_{\lambda \in \Phi} C_{ea}(X,\lambda) = \max_{\lambda \in \Phi_S} \min_{\gamma \geq 0} J_{\gamma}^{ea}(X,\lambda) = \min_{\gamma \geq 0} \max_{\lambda \in \Phi_D} J_{\gamma}^{ea}(X,\lambda), \quad \text{(EC.2)} \]

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 2002, Chapter 11) ensures that \( \Phi_S \) is a dominating class among all admissible \( \lambda \), 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 2002, 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 \( \bar{\epsilon} \in \mathbb{R} \),

\[ \{ x \in \tilde{S} : \max_{\lambda \in \Lambda} \bar{c}(x,\lambda) > \bar{\epsilon} \} \text{ is finite.} \quad \text{(EC.3)} \]

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

**Condition 2** (Boundedness Condition; Condition 11.1 in Altman (1999)): \( \bar{c} \) is bounded from above and for each \( i = 1, 2, \ldots, I \), \( d_i^\alpha \) 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 \( \bar{c} \) is bounded from above. In addition, we see that for all \( \alpha \),

\[ \min_{\lambda \in \Lambda} d_i^\alpha(\lambda) = d_i^\alpha(\bar{\lambda}) = 0, \quad \text{(EC.4)} \]
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 $\max_{\lambda \in \Phi_D} C_{ea}(X, \lambda)$, 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

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

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

$$C_{\text{constraint}}^*(\beta) := \min_{(T, \Psi)} \max_{\lambda \in \Phi_D} C_{ea}(X, \lambda) = \min_{\gamma > 0} \min_{(T, \Psi)} \max_{\lambda \in \Phi_D} J_{ea}(X, \lambda)$$  

$$= \min_{\gamma > 0} \left[ C_{\text{penalty}}^*(\gamma) + \langle \beta, \gamma \rangle \right],$$

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 3, Proposition 5 and Theorem 1.

**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

$$\psi_i^{\bar{\alpha}}(t) = \exp \left\{ \bar{\alpha} \int_0^t \ln(1 + \theta_i(u))dA_i(u) \right\} \cdot \exp \left\{ -\bar{\alpha} \int_0^t \bar{\lambda}_i \theta_i(u)du \right\}. \quad (EC.8)$$

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

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

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

$$\Xi := \mathbb{E} Q_i \left[ \exp \left\{ \sum_k \bar{\alpha} \ln(1 + \theta_i(u_k))(A_i(u_{k+1}) - A_i(u_k)) \right\} \right].$$
Then
\[
\mathbb{E}^{Q_i} \left[ \exp \left\{ \tilde{\alpha} \int_0^T \ln(1 + \theta_i(u)) dA_i(u) \right\} \right] = \mathbb{E}^{Q_i} \left[ \lim_{\Delta \to 0} \exp \left\{ \sum_k \tilde{\alpha} \ln(1 + \theta_i(u_k)) (A_i(u_{k+1}) - A_i(u_k)) \right\} \right]
\]
\[
= \lim_{\Delta \to 0} \mathbb{E}^{Q_i} \left[ \exp \left\{ \sum_k \tilde{\alpha} \ln(1 + \theta_i(u_k)) (A_i(u_{k+1}) - A_i(u_k)) \right\} \right]
\]
where 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
\[
\mathbb{E}^{Q_i} \left[ \exp \left\{ \tilde{\alpha} \int_0^T \ln(1 + \theta_i(u)) dA_i(u) \right\} \right] = \mathbb{E}^{Q_i} \left[ \lim_{\Delta \to 0} \exp \left\{ \sum_k \tilde{\alpha} \ln(1 + \theta_i(u_k)) (A_i(u_{k+1}) - A_i(u_k)) \right\} \right]
\]
\[
= \lim_{\Delta \to 0} \mathbb{E}^{Q_i} \left[ \exp \left\{ \sum_k \tilde{\alpha} \ln(1 + \theta_i(u_k)) (A_i(u_{k+1}) - A_i(u_k)) \right\} \right]
\]
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
\[
R^*_i(t) := \frac{1}{\alpha - 1} \ln \mathbb{E}^{Q_i} [\psi_i(t)^{\alpha - 1}] = \frac{1}{\alpha - 1} \left\{ \int_0^T \tilde{\lambda}_i [(1 + \theta_i(u))^{\alpha} - (1 + \theta_i(u))] du - \tilde{\alpha} \int_0^T \tilde{\lambda}_i \theta_i(u) du \right\},
\]
which, after further simplification, leads to the desired result. \( \square \)

**Proof of Proposition 3.** Note that \( \hat{\xi}_i(k) := m_i \xi_i(k) \) for \( k \geq 0 \) and \( i = 1 \ldots , I \). Given an admissible control \( (\hat{Y}, \hat{\Psi}) \) for the SDG, then \( B(t) \) is a one-dimensional Brownian motion with specified parameters. Since \( (\hat{Y}, \hat{\Psi}) \) is admissible, we know that \( W(t) \geq 0 \) and \( U(t) \) is non-decreasing with \( U(0) = 0 \). By our construction of the effective holding cost rate function \( h \), the cost rate function \( r^* \) and the new outsourcing cost function \( \tilde{\phi} \), it is easy to verify that the long-run average cost for the SDG and its workload version are identical along almost every sample path. To prove the second part, given an admissible control \( (U, \hat{\Psi}) \) for the workload problem, let \( \hat{\xi}_i(k) = \xi_i(k)/m_i \) for \( k \geq 0 \) and \( i = 1 \ldots , I \). We also let \( \hat{Z}_i \) be independent Brownian motions where \( \sum_{i=1}^I m_i \hat{Z}_i(t) = B(t) \). In addition, let \( \sum_{i=1}^I m_i \hat{Q}_i(t) = W(t), \sum_{i=1}^I \rho_i \theta_i(t) = \zeta(t) \) and \( \sum_{i=1}^I \hat{Y}_i(t) = U(t) \). We can write down that
\[
\mu_i \hat{Y}_i(t) = \hat{Q}_i(t) - \hat{Q}_i(0) - \hat{Z}_i(t) - \int_0^t \tilde{\lambda}_i \theta_i(u) du + \sum_{k=0}^{N_i(t)} \hat{\xi}_i(k), \quad i = 1, \ldots , I.
\]
Note that because \( (U, \hat{\Psi}) \) is admissible, we conclude that (11)–(13) hold. And by our construction of \( h \) and \( r^* \) (minimization), \( (\hat{Y}, \hat{\Psi}) \) is an admissible policy for the SDG whose cost is greater than or equal to that for the workload problem. This completes the proof of Proposition 3. \( \square \)
**Proof of Proposition 5.** To begin, we introduce the following supporting lemmas that are critical for proving this proposition. Below, unless stated otherwise, $\pi'(w, \eta)$ is used to denote the first-order partial derivative of $\pi$ with respect to its first argument. Lemma EC.1 asserts non-negativity and Lipschitz continuity of the function $g$.

**Lemma EC.1.** $g(x)$ is non-negative and Lipschitz continuous in $x \in \mathbb{R}$, i.e., for any $x_1$ and $x_2$, we have

$$|g(x_1) - g(x_2)| \leq M |x_1 - x_2|. \quad \text{(EC.10)}$$

Note that the properties claimed by Lemma EC.1 provide the standard (sufficient) condition for equations like (27) to have a unique solution (see, for example, chapter 3 in David et al. (2018)).

**Lemma EC.2.** (i) For any $\eta \in \mathbb{R}$, the ordinary differential equation (27) has a unique continuously differentiable solution $\pi(w, \eta)$. (ii) $\pi(w, \eta)$ is continuous in $\eta \in \mathbb{R}$, and $\pi'(w, \eta)$ is continuous in $w \in \mathbb{R}_+$, and $\eta \in \mathbb{R}$, respectively.

**Lemma EC.3.** For fixed $w > 0$, $\pi(w, \eta)$ is strictly increasing in $\eta \in \mathbb{R}$ and

$$\lim_{\eta \to \pm \infty} \pi(w, \eta) = \pm \infty. \quad \text{(EC.11)}$$

**Lemma EC.4.** There exists an upper bound $\bar{\eta}$ with $\bar{\eta} > 0$ such that the following results hold:

(i) If $\eta \leq 0$, then $\pi(w, \eta)$ is strictly decreasing in $w \in [0, \infty)$ and

$$\lim_{w \to \infty} \pi(w, \eta) = -\infty. \quad \text{(EC.12)}$$

(ii) If $\eta \geq \bar{\eta}$, then $\pi(w, \eta)$ is strictly increasing in $w \in [0, \infty)$ and

$$\lim_{w \to \infty} \pi(w, \eta) = \infty. \quad \text{(EC.13)}$$

(iii) If $0 < \eta < \bar{\eta}$, then there exists a unique number $w^*(\eta)$ such that $\pi(w, \eta)$ is strictly increasing in $[0, w^*(\eta)]$ and strictly decreasing in $[w^*(\eta), \infty)$, where

$$w^*(\eta) := \inf \{ w \geq 0 : \pi'(w, \eta) \leq 0 \}.$$ 

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

Lemma EC.4 divides the value of $\eta$ into three segments, separated by two cut-off points: 0 and $\bar{\eta}$. In particular, the presence of the cut-off point 0 is due to the condition $g(0) = 0$. Mathematically, if $g(0) \neq 0$, then the cut-off point 0 should be modified to $g(0)$. We can then use the previous results to find the unique parameters $q_i$ and $s_i$ for each $i$. In Lemma EC.5 we show that there exist unique $q_i$, $s_i$ and $\eta_i$ with $\eta_i \in (0, \bar{\eta})$ and $0 < q_i < s_i$ such that
\[
\pi(q_i, \eta_i) = \pi(s_i, \eta_i) = \tilde{\ell}_i, \tag{EC.14}
\]

and
\[
\int_{q_i}^{s_i} \left[ \pi(w, \eta_i) - \tilde{\ell}_i \right] dw = L_i. \tag{EC.15}
\]

**Lemma EC.5.** (i) There exists a finite number \(\eta^* \in (0, \bar{\eta})\) such that for any \(\eta \in (\eta^*, \bar{\eta})\), there exist two unique numbers \(q(\eta)\) and \(s(\eta)\) with \(0 < q(\eta) < w^*(\eta) < s(\eta)\) satisfying
\[
\pi(q(\eta), \eta) = \pi(s(\eta), \eta) = \tilde{\ell}_i.
\]

(ii) There exists a unique finite number \(\eta_i \in (\eta^*, \bar{\eta})\) such that
\[
\tilde{f}(\eta_i) = L_i, \tag{EC.16}
\]

where \(\tilde{f}(\eta) := \int_0^{\infty} (\pi(w, \eta) - \tilde{\ell}_i)^+ dw\) is strictly increasing in \(\eta \in \mathbb{R}\).

**Remark EC.1.** Letting \(q_i = q(\eta_i)\) and \(s_i = s(\eta_i)\), Lemma EC.5 directly implies (EC.14) and (EC.15).

We are now going to prove part (i) of Proposition 5. Let \(\pi_i = \pi(w, \eta_i)\). Recall that \(\pi_i\) is a continuously differentiable solution to (27) with the initial condition (28), so (EC.14) and (EC.15) ensure the boundary condition (29). Also from Lemma EC.4(iii) we know that \(0 < q_i < s_i < \infty\). Thus we have completed the proof of part (i).

To prove part (ii) of the proposition, it suffices to show
\[
\frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_i \begin{cases} = 0 & \text{for } w \in (0, s_i) \\ \geq 0 & \text{for } w \geq s_i \end{cases} \tag{EC.17}
\]
and
\[
\inf_{0 \leq z \leq w} \left\{ v(w - z) + \tilde{\ell}_i z + L_i \right\} - v(w) \begin{cases} \geq 0 & \text{for } w \in (0, s_i) \\ = 0 & \text{for } w \geq s_i \end{cases} \tag{EC.18}
\]
for any function \(v\) defined as in the proposition. By the definition of \(v\), we know that
\[
\frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_i = 0 \quad \text{for } w \in (0, s_i).
\]

For \(w \geq s_i\), we have
\[
\frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_i \geq \frac{1}{2} \sigma^2 v''(s_i-) + g(\tilde{\ell}_i) + h(s_i-) - \eta_i = 0.
\]

This proves (EC.17). To show (EC.18), it suffices to establish the following equivalence:
\[
\sup_{0 \leq y \leq w} \int_y^w \left[ v'(z) - \tilde{\ell}_i \right] dz \begin{cases} \leq L_i & \text{for } w \in (0, s_i) \\ = L_i & \text{for } w \geq s_i \end{cases}
\]
which holds true by the definition of $v$ and the fact that
\[
\pi(w, \eta_i) = \begin{cases} < \ell_i & \text{for } w < q_i \\ \geq \ell_i & \text{for } w \in [q_i, s_i] \\ < \ell_i & \text{for } w > s_i \end{cases}
\]
due to Lemma EC.4(iii) and Lemma EC.5. We thus complete the proof for part (ii) of the proposition.

Towards proving part (iii) of Proposition 5, let us define
\[
J_i(w) := \max_{\zeta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\mathcal{C} \left[ \int_0^t h(W(u)) du - \int_0^t r^*(\zeta(u)) du + \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right]_{W(0) = w},
\]
and let
\[
\delta_i(k) := v(W(\tau_i(k))) - v(W(\tau_i(k)^-)).
\]
From (30) it follows that $v(y) - v(x) \leq \tilde{\phi}_i(y - x)$ for $y > x$, and so
\[
-\delta_i(k) \leq \tilde{\phi}_i(W(\tau_i(k)^-)) - v(W(\tau_i(k))) = \tilde{\phi}_i(\tilde{\xi}_i(k)) \quad \text{for } k = 0, 1, 2, \ldots. \tag{EC.19}
\]
On the other hand, applying the generalized Itô’s formula, we obtain, for $t \geq 0$,
\[
\mathbb{E}^\mathcal{C}[v(W(t))] = v(w) + \mathbb{E}^\mathcal{C} \left[ \int_0^t \left( \frac{\sigma^2}{2} v''(W(u)) + \zeta(u) v'(W(u)) \right) du \right] + \mathbb{E}^\mathcal{C} \left[ \sum_{k=0}^{N_i(t)} \delta_i(k) \right].
\]
On substituting (EC.19) into above identity and using (30), we deduce
\[
v(w) \leq \mathbb{E}^\mathcal{C} \left[ \int_0^t \left( h(W(u)) - r^*(\zeta(u)) - \eta_i \right) du \right] + \mathbb{E}^\mathcal{C} \left[ \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right] \tag{EC.20}
\]
\[
+ \mathbb{E}^\mathcal{C}[v(W(t))] + \mathbb{E}^\mathcal{C} \left[ \int_0^t \left( g(v'(W(u))) + r^*(\zeta(u)) - \zeta(u) v'(W(u)) \right) du \right].
\]
Now, consider a special drift-rate control $\zeta^#(W)$, defined as
\[
\zeta^#(W) := \inf \max_{\zeta} \{ v'(W) \zeta - r^*(\zeta) \}. \tag{EC.21}
\]
Clearly $\zeta^#(\cdot)$ is an adaptive control satisfying
\[
g(v'(W)) + r^*(\zeta^#(W)) - \zeta^#(W) v'(W) = 0. \tag{EC.22}
\]
On combining (EC.20) and (EC.22), we see that
\[
v(w) \leq \mathbb{E}^{\mathcal{C}^#} \left[ \int_0^t \left( h(W(u)) - r^*(\zeta(u)) - \eta_i \right) du \right] + \mathbb{E}^{\mathcal{C}^#} \left[ \sum_{k=0}^{N_i(t)} \tilde{\phi}_i(\tilde{\xi}_i(k)) \right] + \mathbb{E}^{\mathcal{C}^#}[v(W(t))]. \tag{EC.23}
\]
Now, dividing both sides of (EC.23) 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
\[
\eta_i \leq J_i(w) + \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^{\mathcal{C}^#}[v(W(t))]. \tag{EC.24}
\]
If \( \limsup_{t \to \infty} (1/t) \mathbb{E}^{\zeta^\#} [v(W(t))] \leq 0 \), then \( \eta_i \leq J_i(w) \) holds trivially as a result of (EC.24). 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 that \( \eta_i \leq J_i(w) \). To do so, we adopt the ingenious argument used by Ormeci et al. (2008) in their optimality proof. To begin with, put \( a := \limsup_{t \to \infty} (1/t) \mathbb{E}^{\zeta^\#} [v(W(t))] \). Then there exists some constant \( \tilde{t} > 0 \) such that \( (1/t) \mathbb{E}^{\zeta^\#} [v(W(t))] > a/2 \) for \( t \geq \tilde{t} \). 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.25)

Taking expectation on both sides of (EC.25), 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

\[
\mathbb{E}^{\zeta^\#} [W(t)] \geq \frac{1}{t} \left[ -v(w) + ta/2 \right] - w = l_1 t + l_2 \quad \text{for} \quad t \geq \tilde{t},
\]

where \( l_1 > 0 \) and \( l_2 \in \mathbb{R} \) are two fixed constants. Thus,

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

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 \( v_m \in C^2(0, s_i) \) 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

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

Comparing these with (24) and (25), we immediately conclude that \( v_m = v \) and \( \eta_m = \eta_i \). Therefore, the control rule \( \zeta^\# \) defined by (EC.21) 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 4) completes the proof. \( \square \)
Proof of Theorem 1. The proof of this theorem follows closely the steps in the proof of part (iii) in Proposition 5. Thus, we only highlight the key differences. To start, let
\[ J(w) := \max_{\zeta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\zeta \left[ \int_0^t h(W(u)) \, du - \int_0^t r^*(\zeta(u)) \, du + \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \bar{\phi}_i(\xi_i(k)) \right] \]
and define \( \delta_i(k) \) in the same way as we did in the proof of Proposition 5. Now using (31), we conclude that, for all \( i \), \( v(y) - v(x) \leq \bar{\phi}_i(y - x) \); thus for all \( k = 1, 2, \ldots \) and \( i = 1, \ldots, I \), we have
\[ -\delta_i(k) \leq \bar{\phi}_i(\xi_i(k)). \] (EC.26)
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.26) into above identity and using (31), we deduce
\[ v(w) \leq \mathbb{E}^\zeta \left[ \int_0^t \left( h(W(u)) - r^*(\zeta(u)) - \eta_k \right) \, du \right] + \mathbb{E}^\zeta \left[ \sum_{i=1}^I \sum_{k=0}^{N_i(t)} \bar{\phi}_i(\xi_i(k)) \right] + \mathbb{E}^\zeta [v(W(t))] + \mathbb{E}^\zeta \left[ \int_0^t \left( g(v'(W(u))) + r^*(\zeta(u)) - \zeta(u) v'(W(u)) \right) \, du \right]. \]

The rest of the proof proceeds in exactly the same fashion as the proof of Proposition 5. First, by choosing \( \zeta = \zeta^\# \) with \( \zeta^\# \) given as in (EC.21), one can formally show that \( \eta_{i^*} \leq J(w) \). Second, one can easily argue that \( \zeta^\# \) is the maximizer's best response: when the decision-maker chooses \((i^*, q^*, s^*)\), the maximizer will follow \( \zeta^\# \) and never deviate from it. The desired optimality of the control band policy \((i^*, q^*, s^*)\) then follows immediately from these two points plus the fact that the “lower bound” \( \eta_{i^*} \) is attained with the policy \((i^*, q^*, s^*)\) as demonstrated by Proposition 5. \( \square \)

Proof of Proposition 6. We need to show that
\[ \frac{1}{2} \sigma^2 v''(w) + g(v'(w)) + h(w) - \eta_\star \geq 0 \quad \text{and} \]
\[ \inf_{0 \leq z \leq w} \left[ v(w - z) + \bar{\phi}_i(z) \right] - v(w) \geq 0 \quad \text{for all} \ i \] (EC.27) (EC.28)
for any \( v \) defined as in the proposition. The verification for (EC.27) is very similar to that for (EC.17), so we will only focus on proving (EC.28). To this end, it is sufficient to show that
\[ \sup_{0 \leq y \leq w} \int_y^w \left[ v'(z) - \bar{\ell}_i \right] \, dz \leq L_i \quad \text{for all} \ i. \] (EC.29)
Recall that Lemma EC.4(iii) demonstrates that for every \( i \), the function graph of \( \pi(\cdot, \eta_{i^*}) \) either does not intersect, intersects once (“touch”), or intersects twice with the horizontal line \( \bar{\ell}_i \). Thus, if
\[ \sup_{0 \leq y \leq w} \int_y^w \left[ v'(z) - \bar{\ell}_k \right] \, dz > L_k \quad \text{for some} \ k, \]
then by the definition of \( v \), function graph of \( \pi(\cdot, \eta_*) \) must intersect the horizontal line \( \tilde{\ell}_k \) twice and we have

\[
\int_q^s \left[ v'(z) - \tilde{\ell}_k \right] \, dz = \sup_{0 \leq y \leq w} \int_y^w \left[ v'(z) - \tilde{\ell}_k \right] \, dz,
\]

where

\[
q := \inf \left\{ w \geq 0 : v'(w) = \tilde{\ell}_k \right\} \quad \text{and} \quad s := \sup \left\{ w \geq 0 : v'(w) = \tilde{\ell}_k \right\}.
\]

It follows that

\[
\int_q^s \left[ \pi(z, \eta_*) - \tilde{\ell}_k \right] \, dz > L_k,
\]

which in turn implies that \( \eta_k < \eta_* \). This is a contradiction by the definition of \( \eta_* \). Therefore, we can conclude that (EC.29) must be satisfied. The proof is thus complete. \( \square \)

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

To find the solution of the optimality equation (24), we start with an initial guess of \( v \), denoted as \( v_0 \), that solves

\[
\frac{1}{2} \sigma^2 v''_0(w) + h(w) = \eta_i, \quad w \in (0, s_i) \tag{EC.30}
\]

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.30) is a second-order linear ordinary differential equation, so we can solve it analytically. Then for each \( w \in (0, s_i) \), we seek \( \zeta_0(w) \) that maximizes \( \{ \zeta_0(w) v'_0(w) - r^*(\zeta_0(w)) \} \). The next step is to find \( v_1 \) such that

\[
\frac{1}{2} \sigma^2 v''_1(w) + \zeta_0(w)v'_1(w) + h(w) - r^*(\zeta_0(w)) = \eta_i, \quad w \in (0, s_i) \tag{EC.31}
\]

subject to the same boundary conditions and necessary optimality conditions as mentioned previously. We can solve (EC.31) numerically via the finite difference method (FDM).

In general, using the \( k \)th estimate of \( v \), denoted as \( v_k \), we can find \( \zeta_k(w) \) that maximizes \( \{ \zeta_k(w) v'_k(w) - r^*(\zeta_k(w)) \} \), and further solve the ordinary differential equation

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

subject to the set of boundary conditions and necessary optimality conditions by using FDM to get \( v_{k+1}(w) \), the \((k + 1)\)th estimate of \( v \). Repeating these steps we obtain an iterative procedure that generates a sequence \( \{v_{k+1}(w), \zeta_k(w)\} \) 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 \( ||v_{k+1}(w) - v_k(w)|| \) and \( ||\zeta_{k+1}(w) - \zeta_k(w)|| \) become sufficiently small, for \( w \in (0, s_i) \).
EC.4. Additional Numerical Results for Two-product Models
This section presents additional numerical results. In §EC.4.1, we present another example where we compare the solution to the original penalty problem with that to the SDG. In §EC.4.2, we provide the optimal thresholds used in simulations and the corresponding simulated costs. In §EC.4.3, we offer an additional example comparing the optimal cost attained using the actual demand model with the optimal robust cost achieved through our proposed method.

EC.4.1. Solution Comparison Between the Penalty Problem and the SDG
Figures EC.1 and EC.2 show the optimal sequencing and outsourcing rules attained by solving the original penalty problem and the SDG under the KL divergence. We observe that the value of the parameter $\alpha$ also influences the optimal outsourcing rule, such that some states that trigger outsourcing under KL divergence may not trigger outsourcing under the Rényi divergence with $\alpha = 1/2$.

Table EC.1 and Figure EC.3 compare the optimal costs obtained from the original penalty problem to those derived from the SDG (workload problem) for various $(\gamma_1, \gamma_2)$ pairs. We observe a remarkable similarity between the long-run average costs obtained from the SDG and the values obtained from the original penalty problem.

![Figure EC.1](image1.png)
(a) Class 1 is in service
(b) Class 2 is in service

Figure EC.1  Sequencing strategies derived from the penalty problem and the SDG with $\gamma_1 = \gamma_2 = 30$ when $\alpha = 1$

EC.4.2. Optimal Thresholds and the Simulated Costs in §9.2
We present supplementary numerical results obtained in §9.2 in this subsection. Table EC.2 reports the optimal outsourcing thresholds used in simulations, while Tables EC.3 and EC.4 report the estimated long-run average costs generated from simulations in §9.2.1 and §9.2.2, respectively.
(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 $\alpha = 1$

Figure EC.3 Long-run average costs of the original penalty problem and the SDG with $\alpha = 1/2$ and 1

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\eta_{\text{exact}}$</th>
<th>$\eta_{\text{SDG}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>35.1853</td>
<td>35.9829</td>
</tr>
<tr>
<td>1</td>
<td>35.1853</td>
<td>35.9829</td>
</tr>
<tr>
<td>1/2</td>
<td>22.1417</td>
<td>21.0459</td>
</tr>
<tr>
<td>1</td>
<td>22.1417</td>
<td>21.0459</td>
</tr>
<tr>
<td>1/2</td>
<td>20.7016</td>
<td>19.7286</td>
</tr>
<tr>
<td>1</td>
<td>20.7016</td>
<td>19.7286</td>
</tr>
<tr>
<td>1/2</td>
<td>18.6757</td>
<td>17.9884</td>
</tr>
<tr>
<td>1</td>
<td>18.6757</td>
<td>17.9884</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\eta_{\text{exact}}$</th>
<th>$\eta_{\text{SDG}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>35.1853</td>
<td>35.9829</td>
</tr>
<tr>
<td>1</td>
<td>35.1853</td>
<td>35.9829</td>
</tr>
<tr>
<td>1/2</td>
<td>22.1417</td>
<td>21.0459</td>
</tr>
<tr>
<td>1</td>
<td>22.1417</td>
<td>21.0459</td>
</tr>
<tr>
<td>1/2</td>
<td>20.7016</td>
<td>19.7286</td>
</tr>
<tr>
<td>1</td>
<td>20.7016</td>
<td>19.7286</td>
</tr>
<tr>
<td>1/2</td>
<td>18.6757</td>
<td>17.9884</td>
</tr>
<tr>
<td>1</td>
<td>18.6757</td>
<td>17.9884</td>
</tr>
</tbody>
</table>
Table EC.2  Optimal outsourcing parameters for $\alpha = 3$ and $\alpha = 1$

<table>
<thead>
<tr>
<th>$(\gamma_1, \gamma_2)$</th>
<th>$q_1$</th>
<th>$s_1$</th>
<th>$(\gamma_1, \gamma_2)$</th>
<th>$q_1$</th>
<th>$s_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.02, 0.02)</td>
<td>0.012</td>
<td>0.246</td>
<td>(0.1, 0.1)</td>
<td>0.033</td>
<td>0.193</td>
</tr>
<tr>
<td>(0.02, 0.03)</td>
<td>0.012</td>
<td>0.242</td>
<td>(0.1, 1)</td>
<td>0.034</td>
<td>0.192</td>
</tr>
<tr>
<td>(0.02, 0.1)</td>
<td>0.012</td>
<td>0.239</td>
<td>(0.1, 5)</td>
<td>0.035</td>
<td>0.191</td>
</tr>
<tr>
<td>(0.02, 0.5)</td>
<td>0.013</td>
<td>0.238</td>
<td>(0.1, 10)</td>
<td>0.035</td>
<td>0.191</td>
</tr>
<tr>
<td>(0.02, 0.100)</td>
<td>0.013</td>
<td>0.240</td>
<td>(0.1, 100)</td>
<td>0.035</td>
<td>0.192</td>
</tr>
<tr>
<td>(0.05, 0.02)</td>
<td>0.019</td>
<td>0.213</td>
<td>(1.0, 1)</td>
<td>0.050</td>
<td>0.188</td>
</tr>
<tr>
<td>(0.05, 0.05)</td>
<td>0.020</td>
<td>0.211</td>
<td>(1, 1)</td>
<td>0.052</td>
<td>0.187</td>
</tr>
<tr>
<td>(0.05, 0.1)</td>
<td>0.021</td>
<td>0.208</td>
<td>(1.5)</td>
<td>0.053</td>
<td>0.189</td>
</tr>
<tr>
<td>(0.05, 0.5)</td>
<td>0.021</td>
<td>0.207</td>
<td>(1.10)</td>
<td>0.054</td>
<td>0.189</td>
</tr>
<tr>
<td>(0.05, 1.0)</td>
<td>0.026</td>
<td>0.198</td>
<td>(5.0, 1)</td>
<td>0.073</td>
<td>0.200</td>
</tr>
<tr>
<td>(0.1, 0.02)</td>
<td>0.029</td>
<td>0.194</td>
<td>(5.1)</td>
<td>0.077</td>
<td>0.201</td>
</tr>
<tr>
<td>(0.1, 0.1)</td>
<td>0.031</td>
<td>0.193</td>
<td>(5.5)</td>
<td>0.079</td>
<td>0.203</td>
</tr>
<tr>
<td>(0.1, 0.5)</td>
<td>0.032</td>
<td>0.190</td>
<td>(5.10)</td>
<td>0.080</td>
<td>0.204</td>
</tr>
<tr>
<td>(0.1, 1.00)</td>
<td>0.035</td>
<td>0.190</td>
<td>(5.100)</td>
<td>0.080</td>
<td>0.204</td>
</tr>
<tr>
<td>(0.5, 0.02)</td>
<td>0.044</td>
<td>0.187</td>
<td>(10.0, 1)</td>
<td>0.077</td>
<td>0.202</td>
</tr>
<tr>
<td>(0.5, 0.05)</td>
<td>0.054</td>
<td>0.188</td>
<td>(10.1)</td>
<td>0.080</td>
<td>0.205</td>
</tr>
<tr>
<td>(0.5, 0.1)</td>
<td>0.059</td>
<td>0.191</td>
<td>(10.5)</td>
<td>0.083</td>
<td>0.207</td>
</tr>
<tr>
<td>(0.5, 0.5)</td>
<td>0.064</td>
<td>0.193</td>
<td>(10.10)</td>
<td>0.084</td>
<td>0.207</td>
</tr>
<tr>
<td>(0.5, 1.00)</td>
<td>0.066</td>
<td>0.194</td>
<td>(10.100)</td>
<td>0.084</td>
<td>0.207</td>
</tr>
<tr>
<td>(10, 0.02)</td>
<td>0.055</td>
<td>0.190</td>
<td>(100, 0.1)</td>
<td>0.081</td>
<td>0.203</td>
</tr>
<tr>
<td>(10, 0.05)</td>
<td>0.071</td>
<td>0.197</td>
<td>(100, 1)</td>
<td>0.083</td>
<td>0.206</td>
</tr>
<tr>
<td>(10, 0.1)</td>
<td>0.078</td>
<td>0.203</td>
<td>(100, 5)</td>
<td>0.085</td>
<td>0.208</td>
</tr>
<tr>
<td>(10, 0.5)</td>
<td>0.085</td>
<td>0.209</td>
<td>(100, 10)</td>
<td>0.086</td>
<td>0.209</td>
</tr>
<tr>
<td>(10, 1.00)</td>
<td>0.088</td>
<td>0.210</td>
<td>(100, 100)</td>
<td>0.088</td>
<td>0.210</td>
</tr>
</tbody>
</table>

Table EC.3  $\hat{C}(\gamma)$ generated from simulations with ARIMA intensity

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\hat{C}(\gamma)$ $\alpha = 3$</th>
<th>$\gamma$</th>
<th>$\hat{C}(\gamma)$ $\alpha = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.02, 0.02)</td>
<td>2.750 ± 1E−1 (0.1, 0.1)</td>
<td>2.626 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.02, 0.03)</td>
<td>2.782 ± 1E−1 (0.1, 1)</td>
<td>2.534 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.02, 0.1)</td>
<td>2.703 ± 1E−1 (0.1, 5)</td>
<td>2.549 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.02, 0.5)</td>
<td>2.680 ± 1E−1 (0.1, 10)</td>
<td>2.547 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.02, 1.00)</td>
<td>2.760 ± 1E−1 (0.1, 100)</td>
<td>2.648 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.05, 0.02)</td>
<td>2.689 ± 1E−1 (1.0, 1)</td>
<td>2.570 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.05, 0.05)</td>
<td>2.646 ± 1E−1 (1.1)</td>
<td>2.551 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.05, 0.1)</td>
<td>2.604 ± 1E−1 (1.5)</td>
<td>2.412 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.05, 0.5)</td>
<td>2.508 ± 1E−1 (1.10)</td>
<td>2.461 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.05, 1.00)</td>
<td>2.591 ± 1E−1 (1.100)</td>
<td>2.570 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.1, 0.02)</td>
<td>2.693 ± 1E−1 (5.0, 1)</td>
<td>2.576 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.1, 0.05)</td>
<td>2.595 ± 1E−1 (5.1)</td>
<td>2.473 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.1, 0.1)</td>
<td>2.619 ± 1E−1 (5.5)</td>
<td>2.534 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.1, 0.5)</td>
<td>2.546 ± 1E−1 (5.10)</td>
<td>2.422 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.1, 1.00)</td>
<td>2.494 ± 1E−1 (5.100)</td>
<td>2.488 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.5, 0.02)</td>
<td>2.556 ± 1E−1 (10.0, 1)</td>
<td>2.539 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.5, 0.05)</td>
<td>2.444 ± 1E−1 (10.1)</td>
<td>2.528 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.5, 0.1)</td>
<td>2.387 ± 1E−1 (10.5)</td>
<td>2.457 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.5, 0.5)</td>
<td>2.404 ± 1E−1 (10.10)</td>
<td>2.472 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(0.5, 1.00)</td>
<td>2.496 ± 1E−1 (10.100)</td>
<td>2.565 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(10, 0.02)</td>
<td>2.639 ± 1E−1 (100, 0.1)</td>
<td>2.562 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(10, 0.05)</td>
<td>2.506 ± 1E−1 (100, 1)</td>
<td>2.593 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(10, 0.1)</td>
<td>2.503 ± 1E−1 (100, 5)</td>
<td>2.538 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(10, 0.5)</td>
<td>2.487 ± 1E−1 (100, 10)</td>
<td>2.496 ± 1E−1</td>
<td></td>
</tr>
<tr>
<td>(10, 1.00)</td>
<td>2.517 ± 1E−1 (100, 100)</td>
<td>2.508 ± 1E−1</td>
<td></td>
</tr>
</tbody>
</table>
where we can see that the difference between $\hat{\lambda}$ each state is exponentially distributed with a rate of 2. The system parameters include $\bar{\lambda}$ arrival model of each product follows a CTMC, similar to demonstrate the effectiveness of our quasi-gradient-descent method in enhancing performance. The this subsection.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\tilde{C}(\gamma)$</th>
<th>$\gamma$</th>
<th>$\hat{C}(\gamma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.02, 0.02)</td>
<td>8.908 ± 1E-1</td>
<td>(0.1, 0.1)</td>
<td>8.599 ± 1E-1</td>
</tr>
<tr>
<td>(0.02, 0.05)</td>
<td>8.586 ± 1E-1</td>
<td>(0.1, 1)</td>
<td>8.568 ± 1E-1</td>
</tr>
<tr>
<td>(0.02, 0.1)</td>
<td>8.632 ± 1E-1</td>
<td>(0.1, 5)</td>
<td>8.588 ± 1E-1</td>
</tr>
<tr>
<td>(0.02, 0.5)</td>
<td>8.709 ± 1E-1</td>
<td>(0.1, 10)</td>
<td>8.608 ± 1E-1</td>
</tr>
<tr>
<td>(0.02, 0.10)</td>
<td>8.656 ± 1E-1</td>
<td>(0.1, 100)</td>
<td>8.644 ± 1E-1</td>
</tr>
<tr>
<td>(0.05, 0.02)</td>
<td>8.630 ± 1E-1</td>
<td>(1, 0.1)</td>
<td>8.641 ± 1E-1</td>
</tr>
<tr>
<td>(0.05, 0.05)</td>
<td>8.584 ± 1E-1</td>
<td>(1, 1)</td>
<td>8.671 ± 1E-1</td>
</tr>
<tr>
<td>(0.05, 0.1)</td>
<td>8.382 ± 1E-1</td>
<td>(1, 5)</td>
<td>8.686 ± 1E-1</td>
</tr>
<tr>
<td>(0.05, 0.5)</td>
<td>8.545 ± 1E-1</td>
<td>(1, 10)</td>
<td>8.697 ± 1E-1</td>
</tr>
<tr>
<td>(0.05, 0.10)</td>
<td>8.547 ± 1E-1</td>
<td>(1, 100)</td>
<td>8.727 ± 1E-1</td>
</tr>
<tr>
<td>(0.1, 0.02)</td>
<td>8.608 ± 1E-1</td>
<td>(5, 0.1)</td>
<td>8.848 ± 1E-1</td>
</tr>
<tr>
<td>(0.1, 0.05)</td>
<td>8.478 ± 1E-1</td>
<td>(5, 1)</td>
<td>8.789 ± 1E-1</td>
</tr>
<tr>
<td>(0.1, 0.1)</td>
<td>8.487 ± 1E-1</td>
<td>(5, 5)</td>
<td>8.877 ± 1E-1</td>
</tr>
<tr>
<td>(0.1, 0.5)</td>
<td>8.543 ± 1E-1</td>
<td>(5, 10)</td>
<td>8.925 ± 1E-1</td>
</tr>
<tr>
<td>(0.1, 1)</td>
<td>8.609 ± 1E-1</td>
<td>(5, 100)</td>
<td>8.960 ± 1E-1</td>
</tr>
<tr>
<td>(0.5, 0.02)</td>
<td>8.643 ± 1E-1</td>
<td>(10, 0.1)</td>
<td>8.881 ± 1E-1</td>
</tr>
<tr>
<td>(0.5, 0.05)</td>
<td>8.460 ± 1E-1</td>
<td>(10, 1)</td>
<td>8.878 ± 1E-1</td>
</tr>
<tr>
<td>(0.5, 0.1)</td>
<td>8.654 ± 1E-1</td>
<td>(10, 5)</td>
<td>8.926 ± 1E-1</td>
</tr>
<tr>
<td>(0.5, 0.5)</td>
<td>8.618 ± 1E-1</td>
<td>(10, 10)</td>
<td>9.012 ± 1E-1</td>
</tr>
<tr>
<td>(0.5, 1)</td>
<td>8.770 ± 1E-1</td>
<td>(10, 100)</td>
<td>9.022 ± 1E-1</td>
</tr>
<tr>
<td>(1, 0.02)</td>
<td>8.503 ± 1E-1</td>
<td>(100, 0.1)</td>
<td>8.977 ± 1E-1</td>
</tr>
<tr>
<td>(1, 0.05)</td>
<td>8.891 ± 1E-1</td>
<td>(100, 1)</td>
<td>8.910 ± 1E-1</td>
</tr>
<tr>
<td>(1, 0.1)</td>
<td>9.046 ± 1E-1</td>
<td>(100, 5)</td>
<td>8.949 ± 1E-1</td>
</tr>
<tr>
<td>(1, 0.5)</td>
<td>9.060 ± 1E-1</td>
<td>(100, 10)</td>
<td>9.015 ± 1E-1</td>
</tr>
<tr>
<td>(1, 1)</td>
<td>9.143 ± 1E-1</td>
<td>(100, 100)</td>
<td>9.082 ± 1E-1</td>
</tr>
</tbody>
</table>

### EC.4.3. Comparison Between Actual and Robust Models

In this subsection, we present an additional example with the aim of comparing the cost of the “best” robust control policy to the actual optimal cost. We use the same cost parameters as the one in §9.3, but with $\hat{\lambda} = 160$ and $\bar{\lambda} = 40$, representing a larger variation of the realized demand rates. Each state’s sojourn time is again exponentially distributed with a rate of 10. Results are reported in Figure EC.4, where we can see that the difference between $\hat{C}(\gamma^*)$ and $C^*$ is 3.44%, which is arguably an acceptable difference. If completely ignoring model uncertainty, the decision-maker would incur a cost that is 8.77% higher than the actual optimal cost $C^*$. Table EC.5 further shows the optimal outsourcing thresholds and the corresponding simulated costs for the single-class examples discussed in §9.3 and this subsection.

### EC.5. Numerical Results for a Multi-product Model

This section presents a make-to-order system with four classes, which serves as an example to demonstrate the effectiveness of our quasi-gradient-descent method in enhancing performance. The arrival model of each product follows a CTMC, similar to §9.2.2. In the CTMC, the sojourn time of each state is exponentially distributed with a rate of 2. The system parameters include $\bar{\lambda}_1 = 90$, $\bar{\lambda}_2 = 60$, $\bar{\lambda}_3 = 50$, $\bar{\lambda}_4 = 50$, $\mu_1 = 150$, $\mu_2 = 300$, $\mu_3 = 500$, and $\mu_4 = 500$. Furthermore, we set $\lambda_1 = 150$, $\lambda_1 = 30$, $\lambda_2 = 100$, $\lambda_2 = 20$, $\lambda_3 = 80$, $\lambda_3 = 20$, and $\lambda_4 = 80$, $\lambda_4 = 20$. The cost data includes fixed outsourcing
Actual cost and cost obtained from the “best” robust model with $\tilde{\lambda} = 160$ and $\hat{\lambda} = 40$

Table EC.5 Outsourcing thresholds and simulated costs for the single-class example

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$q$</th>
<th>$s$</th>
<th>$\tilde{\mathcal{C}}(\gamma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.018</td>
<td>0.205</td>
<td>$3.855 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>0.3</td>
<td>0.025</td>
<td>0.190</td>
<td>$3.808 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>0.5</td>
<td>0.035</td>
<td>0.182</td>
<td>$3.816 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>0.7</td>
<td>0.040</td>
<td>0.182</td>
<td>$3.868 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>1</td>
<td>0.046</td>
<td>0.184</td>
<td>$3.841 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>2</td>
<td>0.053</td>
<td>0.188</td>
<td>$3.937 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>5</td>
<td>0.058</td>
<td>0.192</td>
<td>$3.915 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>10</td>
<td>0.060</td>
<td>0.193</td>
<td>$3.959 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>20</td>
<td>0.061</td>
<td>0.193</td>
<td>$3.971 \pm 3E^{-2}$</td>
</tr>
<tr>
<td>1000</td>
<td>0.061</td>
<td>0.194</td>
<td>$3.975 \pm 3E^{-2}$</td>
</tr>
</tbody>
</table>

cost parameters $L_1 = 2$, $L_2 = 3$, $L_3 = 4$, and $L_4 = 5$, proportional outsourcing cost parameters $\ell_1 = 0.2$ and $\ell_2 = \ell_3 = \ell_4 = 0.5$, and quadratic holding cost rates $a_1 = 0.01$, $a_2 = 0.1$, $a_3 = 0.2$, and $a_4 = 0.2$. The total duration of the experiment is 2000. We again set $T_1 = 50$ and $T_2 = 100$ as the warm-up periods. Statistical estimates of the long-run average simulated cost are compared, and no significant difference is observed between the estimates under the two candidate warm-up periods. Therefore, $T_2 = 100$ is selected as the warm-up period. The desired width of the confidence interval in this example is set to 0.2.

In the 4-class example, when $\alpha = 2$, the value of robustness is approximately $9.6714 - 9.1028 = 0.5686$, which is 6.25% better than completely ignoring model uncertainty. Additionally, the performance of our method is highlighted in Figure EC.5, showing that near-minimum value can be achieved within about 20 iterations. On the other hand, if we were to use the exhaustive method, taking only 5 values for each $\gamma_i$, we would need to conduct $5^4$ simulations, which requires significantly more intensive
computation. Thus, we believe that the quasi-gradient algorithm can deliver significant practical value.

EC.6. A Few Direct Extensions

EC.6.1. The Choice of Discrepancy Measure

The Rényi divergence provides a versatile family of uncertainty sets that represents model uncertainty by the single function \( r \), dictating how nature is penalized based on her actions. Importantly, this approach to representing model uncertainty does not require a specific penalty rate function. The decision-maker can specify any penalty form, as long as the resulting uncertainty set captures model misspecification concerns.

Let \( p(\cdot) \) be a penalty function mapping from \( \mathbb{R}^I \) to \( \mathbb{R}_+ \), such that when nature selects \( \theta(t) \) at time \( t \), the decision-maker incurs a penalty at the rate of \( p(\theta(t)) \). This motivates a general robust control formulation, described below. The robust control problem for the penalty function \( p(\cdot) \) can be formulated as follows: the decision-maker seeks an adapted strategy \((T, \Psi)\) that minimizes

\[
\max_{\theta} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E}^\theta \left[ \int_0^t \left( \sum_{i=1}^I c_i(Q_i(u)) - p(\theta(u)) \right) \, du + \sum_{i=1}^{I} \sum_{k=0}^{N_i(t)} \phi_i(\xi_i(k)) \right].
\]
Replacing \( p \) with \( r \) recovers problem (7). Following the development in §6.1, we arrive at an approximating SDG, in which the decision-maker chooses \((
abla, \hat{\Psi})\) to minimize

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

subject to constraints (11) – (13).

By following similar lines of argument to those in §6.2, we can obtain the corresponding workload problem, in which the decision-maker seeks an adaptive control \((U, \hat{\Psi})\) that minimizes

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

subject to constraints (19) – (21),

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 critical to the proofs of Proposition 5 and Theorem 1. First, \( r^* \) attains its minimum value at \( z = 0 \) with \( r^*(0) = 0 \). Second, its convex conjugate is non-negative and Lipschitz continuous. Consequently, all the analytical results established in the preceding section apply to all \( p^* \) possessing these two properties.

**EC.6.2. General Service Times**

In the main paper, we have assumed that service times are exponentially distributed. As far as diffusion analysis is concerned, this assumption can be effortlessly relaxed to allow for general service time distributions without affecting the main results established in Section 6. Indeed, with general service times, the process \( S_i(t) \) that 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 (10) becomes a Brownian motion with zero drift and variance parameter \( \bar{\lambda}_i \nu_i^2 \), whereas \( \hat{Z}_i \) in (11) becomes a Brownian motion with zero drift and infinitesimal variance \( \sigma_i^2 = \bar{\lambda}_i(1 + \nu_i^2) \).

**EC.6.3. (Slightly) “Imbalanced” Systems**

In this subsection, we demonstrate how to expand our analysis to a more general scenario by relaxing the assumption of “critical-loading” in Equation (8), allowing for an “imperfectly balanced” system. Specifically, we examine a situation where the capacity does not match the supply and nominal demand level exactly. This can be expressed as:

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

where the constant \( \omega \) measures the extent to which the capacity exceeds the nominal demand volume. However, for the purpose of employing diffusion approximation, we require that \( \omega \) be a value of the
order of $1/\sqrt{n}$, where $n$ is a parameter that reflects the system scale. (In the main paper, we mention that $n$ can be taken as $\sum_i \lambda_i$.) This ensures that the “imbalance” between capacity and demand, albeit present, is moderate at best.

With this relaxation, the corresponding SDG is modified to minimize (14) subject to (11), (13), and the constraint that

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

in place of (12). As a result, the workload process, again denoted as $W$, satisfies

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

where $B, \zeta,$ and $O$ are defined as in (15). Consequently, the Bellman-Isaacs equation associated with the workload problem becomes finding $(v, \eta)$ that satisfies

$$\min \left\{ \frac{1}{2} \sigma^2 v''(w) + g(v'(w)) - \omega v'(w) + h(w) - \eta, \min_i \inf_{z \geq 0} \left[ v(w - z) + \tilde{\phi}_i(z) \right] - v(w) \right\} \geq 0 \quad (EC.32)$$

subject to $v'(0) = 0$.

It is noteworthy that we do not require a complete overhaul of the entire analysis to establish the well-posedness of the new Bellman-Isaacs equation. To demonstrate this, we can define $\tilde{g}(\cdot) := g(\cdot) - \omega \cdot$. Consequently, Equation (EC.32) will resemble that in (31), and $\tilde{g}$ will possess all the essential properties of $g$'s that are necessary for the mathematical proofs to hold.

On a side note, a control policy based on a diffusion approximation may remain practically useful even when the actual operating regime deviates significantly from the critical-loading assumption required to justify the use of a diffusion approximation. To illustrate this point, we present a single-class example without model uncertainty, where key system parameters are set to $\lambda = 300$ and $\mu = 100$, and cost parameters include a fixed outsourcing cost parameter $L$ of 0.5, a proportional outsourcing cost parameter $\ell$ of 0.1, and a quadratic holding cost rate function $c(x) = 0.01x^2$. The choice of system parameters implies that the system is overloaded ($\omega = -2$) rather than critically loaded. Absent the role of nature, the SDG simplifies to a diffusion control problem (DCP), which we can solve for the control band thresholds for the queue length denoted as $q$ and $s$ (by a slight abuse of notation). On the other hand, as the system is overloaded, the existing literature on queuing approximation and control suggests that a deterministic (fluid-like) approximation may also work well. Specifically, we expect the queue length to increase approximately linearly at a rate of $\lambda - \mu$, which suggests an “EOQ” formula as a result of this deterministic approximation. Indeed, a direct application of the “EOQ” formula (with a slight yet straightforward modification) can give us another upper boundary, at which the decision-maker would like to push the queue length (through outsourcing) to a lower
barrier. As outsourcing operations are assumed to occur without delay, the application of the “EOQ” formula should produce a lower barrier of zero. In Figure EC.6, the blue line represents the sample path of the queue length \( Q(t) \), while the green and red lines indicate the control band parameters, \( q \) and \( s \), respectively, obtained from the DCP. The dash-dotted gray line represents the control band parameter computed from the “EOQ” formula. The plot suggests that the two solutions, one based on diffusion approximation and the other on deterministic approximation, match up, suggesting that a control policy derived from a diffusion approximation may still be practically valuable even if the conditions required to justify the use of diffusion approximation are not met.

**EC.7. Further Discussion on the Choice of Parameter \( \alpha \)**

In the main paper, we primarily vary \( \gamma \) to define the uncertainty set because it directly impacts the control rule derived for the decision-maker. However, to provide guidance on selecting an appropriate value of \( \alpha \), which plays a crucial role in defining the shape of the uncertainty set, we perform two additional numerical experiments in this section. Specifically, we evaluate different arrival processes in the real-world demand model.

The system parameters are set to \( \bar{\lambda}_1 = 200, \mu_1 = 250, \bar{\lambda}_2 = 100, \) and \( \mu_2 = 500 \), while the cost parameters are fixed outsourcing costs of \( L_1 = 0.5 \) and \( L_2 = 0.8 \), proportional outsourcing costs of \( \ell_1 = \ell_2 = 0.2 \), and quadratic holding cost rates of \( a_1 = 0.01 \) and \( a_2 = 0.02 \). Similar to §9.2.2 in the main paper, we use an arrival model where the demand rate for each order follows a non-homogeneous Poisson process with a CTMC intensity. Below, we present two examples: one where the actual demand rate of class 1 is moderately higher than the average value of \( \bar{\lambda}_1 = 200 \) most of the time but can occasionally drop to a very low value, and another where the actual demand rate of class 1 is moderately lower than the average value of \( \bar{\lambda}_1 = 200 \) most of the time but can occasionally rise to a very high value.
In the first example, we set $\tilde{\lambda}_1 = 260$ and assume the sojourn time is exponentially distributed with a rate of 5. We also set $\hat{\lambda}_1 = 20$ in the CTMC and assume the sojourn time is exponentially distributed with a rate of 15. Furthermore, we set $\tilde{\lambda}_2 = \hat{\lambda}_2 = 100$, so that the demand rate of class 2 stays at the average value $\bar{\lambda}_2 = 100$ at all times. The optimal outsourcing rule is always to outsource product 1, and the optimal sequencing rule is the generalized $c\mu$ rule. Under this model setting, if the decision-maker does not account for model uncertainty and uses the outsourcing thresholds obtained from the nominal model, we observe that the long-run average simulated cost is approximately $12.394 \pm 1E-1$. However, if the decision-maker incorporates KL divergence by varying $\gamma$, the optimal simulated cost is around $11.753 \pm 1E-1$. If the decision-maker chooses to use Rényi divergence with $\alpha = 5$, the optimal simulated cost is approximately $11.581 \pm 1E-1$. In conclusion, for this specific example, we note that using a Rényi divergence with a larger value of $\alpha$ seems to deliver a slightly better performance.

In the second example, we set $\tilde{\lambda}_1 = 380$ with the sojourn time exponentially distributed with a rate of 15, while $\hat{\lambda}_1 = 140$ with the sojourn time exponentially distributed with a rate of 5. Similarly, we set $\tilde{\lambda}_2 = \hat{\lambda}_2 = 100$, so that the demand rate of class 2 stays at the average value $\bar{\lambda}_2 = 100$ at all times. The optimal control rule in this case remains the same as the previous example since the SDG remains unchanged. In this case, if the decision-maker does not consider model uncertainty and uses the outsourcing thresholds obtained from the nominal model, the long-run average simulated cost is approximately $11.306 \pm 1E-1$. By incorporating KL divergence and varying $\gamma$, the optimal simulated cost is around $10.486 \pm 1E-1$. If, however, the decision-maker chooses to use Rényi divergence with $\alpha = 1/2$, the optimal simulated cost is approximately $10.269 \pm 1E-1$. Therefore, for this example, using Rényi divergence with a smaller $\alpha$ seems to deliver a slightly better performance.

The examples discussed above illustrate that the parameter $\alpha$ allows for potential cost savings when the real-world model deviates from the nominal model. However, in comparison to choosing the value of $\gamma$, the choice of $\alpha$ appears to have limited potential for improving cost savings. This explains, in part, why our main paper focuses on selecting an appropriate uncertainty set through the choice of $\gamma$. Nevertheless, the two examples provide valuable insights. In the first example, when the actual arrival rate is moderately higher than the average value for most of the time but significantly lower for a small portion of the time, selecting a larger value of $\alpha$ in the Rényi divergence benefits the decision-maker, albeit slightly. This is because a larger value of $\alpha$ penalizes the right tail of the demand rate distribution (i.e., values that are greater than the nominal demand rate) more heavily than the left tail (i.e., values that are smaller than the nominal demand rate). In contrast, a smaller value of $\alpha$ would penalize the left tail more heavily, reflecting the decision-maker’s prior belief that larger demand rates are more likely than smaller ones. The second example demonstrates this effect, as the actual arrival rate for class 1 is moderately lower than its nominal value for most of the time.
but significantly higher for a small portion of the time. In this case, the decision-maker benefits from choosing a smaller value for $\alpha$.

**EC.8. 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 bounded. Using properties of conjugate functions (see, e.g., Ex. 3.40 in Boyd et al. (2004)), we have

$$g'(x) = \arg \max_{\zeta \in \mathcal{Z}} x\zeta - r^*(\zeta), \quad \text{(EC.33)}$$

where $\mathcal{Z} := [\rho^\top a, \rho^\top b]$. The left-hand side of (EC.33) 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.33) 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 $\alpha$ for convenience). We intend to show that

$$r^*(\zeta_3) < \alpha r^*(\zeta_1) + (1 - \alpha) r^*(\zeta_2).$$

To that end, we note that for each $\zeta$, $r^*(\zeta)$ 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 $r^*(\zeta_3)$. Thus

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

where the second inequality is due to the strict convexity of $r$. This demonstrates that $r^*$ is strictly convex, implying that the right-hand side of (EC.33) is a singleton for all $x \in \mathbb{R}$. The desired Lipschitz continuity thus follows.

*Proof of Lemma EC.2.* For part (i), since $g$ is Lipschitz continuous and $h$ is continuous, we can invoke the Picard–Lindelöf theorem to conclude that there exists a unique continuous solution $\pi(w, \eta)$ to (27) on the interval $[0, \infty)$.

For part (ii), to show the continuity of $\pi(w, \eta)$ in $\eta \in \mathbb{R}$ and the continuity of $\pi'(w, \eta)$ in $w \in \mathbb{R}^+$ and $\eta \in \mathbb{R}$, we can refer to Lemma 5 in Cao and Yao (2018), along with part (i) of the present lemma and the continuity of $h$, $g$, and $\pi$. \qed
Proof of Lemma EC.3. We first argue that, if $\eta_1 < \eta_2$, then $\pi(w, \eta_1) < \pi(w, \eta_2)$ for any fixed $w > 0$. To that end, suppose for the sake of contradiction that $\pi(w, \eta_1) > \pi(w, \eta_2)$ for some $w > 0$. By a slight abuse of notation, we define:

$$f(w) := \pi(w, \eta_2) - \pi(w, \eta_1) \quad \text{and} \quad w_0 := \inf \{w > 0 : f(w) \leq 0\}.$$  

It follows from the definition and continuity of $\pi$ that $f(w_0) = 0 = f(0)$ and $f(w) > 0$ for all $w \in (0, w_0)$. By the continuity of $f(w)$ around $w_0$, there exist two real numbers $w_1, w_2 \in (0, w_0)$ with $w_1 < w_2$ such that

$$f(w_1) > f(w_2) \quad \text{and} \quad Mf(w) < \eta_2 - \eta_1 \quad \text{for all} \quad w \in [w_1, w_2]. \quad (EC.34)$$

It is clear from (27) that

$$\frac{1}{2}\sigma^2 f'(w) + g(\pi(w, \eta_2)) - g(\pi(w, \eta_1)) = \eta_2 - \eta_1. \quad (EC.35)$$

Integrating (EC.35) from $w_1$ to $w_2$ yields that

$$(\eta_2 - \eta_1)(w_2 - w_1)$$

$$= \frac{1}{2}\sigma^2 (f(w_2) - f(w_1)) + \int_{w_1}^{w_2} [g(\pi(w, \eta_2)) - g(\pi(w, \eta_1))] \, dw$$

$$< \int_{w_1}^{w_2} [g(\pi(w, \eta_2)) - g(\pi(w, \eta_1))] \, dw \quad (EC.36)$$

$$\leq \int_{w_1}^{w_2} Mf(w) \, dw$$

$$< (\eta_2 - \eta_1)(w_2 - w_1),$$

where the first and last inequalities follow from (EC.34) and the second inequality is due to (EC.10). Equation (EC.36) yields a contradiction. Therefore, $\pi(w, \eta_1) < \pi(w, \eta_2)$ holds for any fixed $w > 0$, if $\eta_1 < \eta_2$.

Next, we show that $\lim_{\eta \to \infty} \pi(w, \eta) = \infty$ for any given $w > 0$. Note that there must exist a number $\hat{\eta}$ (dependent on $w$) such that for all $\eta \geq \hat{\eta}$,

$$\eta > h(w). \quad (EC.37)$$

We claim that for any fixed $y \in (0, w)$,

$$\pi(y, \eta) \geq 0 \quad \text{for all} \quad \eta \geq \hat{\eta}. \quad (EC.38)$$

To prove this claim, suppose for the sake of contradiction that (EC.38) is not true. Then, there must exist some $z \in (0, w)$ such that $\pi(z, \eta) = 0$ and $\pi'(z, \eta) < 0$. It follows that

$$0 > \frac{1}{2}\sigma^2 \pi'(z, \eta) = \eta - h(z) > \eta - h(w) > 0,$$
which is a contradiction. Therefore, (EC.38) holds, which in particular implies that for all \( y \in (0, w) \)
\[
\frac{1}{2} \sigma^2 \pi'(y, \eta) + M \pi(y, \eta) \geq \eta - h(y) \quad \text{for all } \eta \geq \hat{\eta}.
\]
It follows that
\[
\pi(w, \eta) \geq \frac{2}{\sigma^2} \int_0^w \left[ \eta - h(y) \right] e^{-\frac{2M}{\sigma^2} y} dy \quad \text{for all } \eta \geq \hat{\eta}.
\]
Letting \( \eta \to \infty \) in the inequality above allows us to conclude that
\[
\lim_{\eta \to \infty} \pi(w, \eta) = \infty.
\]
The proof of \( \lim_{\eta \to -\infty} \pi(w, \eta) = -\infty \) is similar and thus is omitted.

Proof of Lemma EC.4. Note that \( h(w_1) < h(w_2) \) for \( w_1 < w_2 \). This, along with (27), allows us to conclude that there do not exist two numbers \( w_1 < w_2 \) such that
\[
\pi(w_1, \eta) = \pi(w_2, \eta) \quad \text{and } \pi'(w_1, \eta) \leq 0 \leq \pi'(w_2, \eta).
\]
Therefore, (a) \( \pi(w, \eta) \) cannot have a local minimizer in \( w \in (0, \infty) \), and (b) \( \pi(w, \eta) \) cannot be a constant in any interval in \( (0, \infty) \). We will then employ properties (a) and (b) to prove parts (i)–(iii).

We first prove part (i). Note that when \( \eta \leq 0 \), we must have \( \pi'(0, \eta) < 0 \). The continuity of \( \pi'(0, \eta) \) and properties (a) and (b) immediately imply that \( \pi(w, \eta) \) is strictly decreasing in \( w \) for \( w > 0 \). Next, we show that \( \lim_{w \to \infty} \pi(w, \eta) = -\infty \). We can prove this by contradiction. If the statement is not correct, there must exist a finite number \( \pi \) such that \( \lim_{w \to \infty} \pi(w, \eta) = \pi \) and \( \lim_{w \to \infty} \pi'(w, \eta) = 0 \). Letting \( w \to \infty \) in (27) yields that \( \lim_{w \to \infty} h(w) = \eta - g(\pi) \), which contradicts the fact that \( \lim_{w \to \infty}(w) = \infty \) thanks to Assumption 1.

For (ii), we define
\[
\bar{\eta} := \sup \{ \eta \in \mathbb{R} : \text{there exists a } w > 0 \text{ such that } \pi'(w, \eta) < 0 \}.
\]
Note that \( \bar{\eta} \) is well defined since \( \bar{\eta} > 0 \). If \( \eta \geq \bar{\eta} \), by the definition of \( \bar{\eta} \) and the continuity of \( \pi'(w, \eta) \) in \( \eta \), we can conclude that \( \pi(w, \eta) \) is strictly increasing in \( w \). The proof of \( \lim_{w \to \infty} \pi(w, \eta) = \infty \) is similar to that of (EC.12) and thus is omitted.

For (iii), we begin by claiming that for each \( \eta \in (0, \bar{\eta}) \), there exists a number \( w \) such that \( \pi'(w, \eta) < 0 \). If not, we have \( \pi'(w, \eta) > 0 \) for all \( w > 0 \). Similar to the proof of (i), we can obtain that
\[
\lim_{w \to \infty} \pi(w, \eta) = \infty.
\]
On the other hand, the definition of \( \bar{\eta} \) and the continuity of \( \pi(w, \eta) \) indicate that there exist a \( \eta^l > \eta \) such that \( \pi'(w^l, \eta^l) < 0 \) for some \( w^l \). For \( w > w^l \), \( \pi(w, \eta^l) \) is decreasing and
\[
\lim_{w \to \infty} \pi(w, \eta^l) = -\infty.
\]
However, (EC.41) and (EC.42) contradict Lemma EC.3 with $\eta^1 > \eta$.

By the continuity of $\pi'(w, \eta)$ in $w$ and the definition of $w^*(\eta)$, we can conclude that $\pi'(w^*(\eta), \eta) = 0$. Furthermore, properties (a) and (b) imply that $\pi(w, \eta)$ is strictly increasing in $[0, w^*(\eta)]$ and strictly decreasing in $[w^*(\eta), \infty)$. The proof of $\lim_{w \to \infty} \pi(w, \eta) = -\infty$ is very similar to that of (EC.12) and thus is omitted.

\textbf{Proof of Lemma EC.5.} For (i), define
\[
\eta^\dagger := \inf \left\{ \eta \in (0, \bar{\eta}) : \pi(w^*(\eta), \eta) \geq \tilde{\ell}_i \right\}.
\]
It follows from Lemma EC.4(iii) that
\[
\pi(w^*(\eta), \eta) = \max_{w \geq 0} \pi(w, \eta),
\]
when $\eta \in (0, \bar{\eta})$. Recall that Lemma EC.3 implies that $\pi(w^*(\eta), \eta)$ is increasing in $\eta$. Furthermore, Lemma EC.4 shows that
\[
\lim_{\eta \downarrow 0} \pi(w^*(\eta), \eta) = 0 < \tilde{\ell}_i \quad \text{and} \quad \lim_{\eta \uparrow \bar{\eta}} \pi(w^*(\eta), \eta) = \infty.
\]
Hence, $\eta^\dagger$ is well-defined, and we also have
\[
\pi(w^*(\eta^\dagger), \eta^\dagger) = \tilde{\ell}_i,
\]
and
\[
\pi(w^*(\eta^\dagger), \eta) > \tilde{\ell}_i, \quad \text{for} \quad \eta \in (\eta^\dagger, \bar{\eta}).
\]
Furthermore, let
\[
q(\eta) := \inf \left\{ w \geq 0 : \pi(w, \eta) = \tilde{\ell}_i \right\} \quad \text{and} \quad s(\eta) := \sup \left\{ w \geq 0 : \pi(w, \eta) = \tilde{\ell}_i \right\}.
\]
It follows from Lemma EC.4(iii) that both $q(\eta)$ and $s(\eta)$ are well-defined, finite, and unique.

For part (b), we can rewrite the function $\tilde{f}(\eta)$ as
\[
\tilde{f}(\eta) = \begin{cases} 
0 & \text{for } \eta \in (-\infty, \eta^\dagger], \\
\int_{q(\eta)}^{s(\eta)} [\pi(w, \eta) - \tilde{\ell}_i]dw & \text{for } \eta \in (\eta^\dagger, \bar{\eta}), \\
\infty & \text{for } \eta \in [\bar{\eta}, \infty).
\end{cases}
\]
Also, $\tilde{f}(\eta)$ is strictly increasing in $\eta \in (\eta^\dagger, \bar{\eta})$. We can thus obtain that
\[
\lim_{\eta \downarrow \eta^\dagger} \tilde{f}(\eta) = 0 \quad \text{and} \quad \lim_{\eta \uparrow \bar{\eta}} \tilde{f}(\eta) = \infty.
\]
Combining this with the continuity of $\tilde{f}$ in $\eta \in (\eta^\dagger, \bar{\eta})$, we conclude that there exists a unique $\eta_i \in (\eta^\dagger, \bar{\eta})$ such that (EC.16) holds. \qed