We formulate and solve a robust dynamic pricing problem for an ambiguity-averse agent who faces an uncertain probabilistic law governing the realized demand for a single product. Specifically, the pricing problem is framed as a stochastic game that involves a maximizing player (the "agent") and a minimizing player ("nature") who promotes robustness by distorting the agent's beliefs within prescribed limits. Our methodology builds on the approach of Lim & Shanthikumar(2007), but can be used to generate a much more versatile class of uncertainty sets. We derive the optimal pricing strategy and the corresponding value function by applying stochastic dynamic programming and solving a version of the Bellman-Isaacs equation. The usefulness of our framework is illustrated by several special cases. Finally, the value of model robustness is exposed through a carefully designed numerical example.

**KEYWORDS**
model uncertainty, ambiguity, robust control, dynamic pricing, revenue management, stochastic games

1 | INTRODUCTION

Price-based revenue management has received a great deal of attention in the literature because varying prices is perhaps the most natural mechanism for demand management in business practices. Analytical models regarding how to make price adjustments to maximize revenues or profits
can date back to the seminal paper by Gallego and van Ryzin (1994). Since then various generalizations have been considered, including (among others) multi-product dynamic pricing (Gallego and Van Ryzin, 1997; Maglaras and Meissner, 2006), pricing with time-dependent price sensitivity (Li and Huh, 2012), pricing under customer choice (Talluri and Van Ryzin, 2004; Zhang and Cooper, 2005; Akçay et al., 2010), joint pricing and inventory control (Federgruen and Heching, 1999; Dong et al., 2009; Zhu and Thonemann, 2009), joint pricing and order fulfilment optimization (Lei et al., 2018), and pricing with demand learning (Besbes and Zeevi, 2009; Araman and Caldentey, 2009). See also the excellent surveys by Bitran and Caldentey (2003) and Elmaghraby and Keskinocak (2003).

Despite the fact that this collection of work has achieved a huge success, most studies assume that the probabilistic law governing the realized demands is “simple” and accurately represents the real-world scenario. In particular, demand arrivals are often assumed to follow a Poisson point process so that a Markov decision process (MDP) framework can be applied to derive optimal pricing strategies. In reality, demand can be more complex than what is implied by a Poisson model; e.g., demand data may exhibit autoregression (Graves, 1999). But still, a Poisson model may be used due to its analytical appeal, in which case model errors may arise, introducing bias into the decision-making process. To summarize, model misspecification can arise easily through simplifying assumptions such as “stationary” and “Markov.” Since model errors can have important operational consequences, developing robust pricing models will be valuable.

We lay no claim to uncovering this deficiency. In the past, some researchers have tried to address model uncertainty or ambiguity in making dynamic pricing decisions. Adida and Perakis (2006) advance a robust approach to the dynamic pricing and inventory control problem with multiple products using a demand-based fluid model. Lim and Shanthikumar (2007) formulate a robust counterpart to the single-product revenue management problem as in Gallego and van Ryzin (1994). In that paper, model uncertainty arises at the level of the point process distribution characterizing realized demand, and the authors adopt a max-min criterion with which nature is adversarial at every point in time. Our approach to handling model uncertainty is in the same spirit as Lim and Shanthikumar (2007) — the cloud of alternative probability measures is only vaguely specified (i.e., the pricing agent evaluates a decision strategy within a set of incompletely articulated probabilistic laws that are formed by perturbing a reference measure). However, there is a key distinction between their work and ours. Lim and Shanthikumar (2007) consider two robust control problems: a constraint problem and a penalty problem, both expressing the agent’s distrust in the reference measure through the notion of the relative entropy. Although our sole focus is a penalty problem, our formulation permits a broader category of uncertainty sets. To briefly explain, constructing an uncertainty set in our setting is tantamount to specifying a penalty rate function $f$, and the use of entropic uncertainty sets, as in Lim and Shanthikumar (2007), corresponds to letting...
\[ f(x) = \theta(x \ln x + 1 - x), \] where the multiplier \( \theta \) parameterizes the class of entropic ambiguities.

From a modeling perspective, our approach to dealing with model uncertainty pertains to the literature that permits ambiguous beliefs on sequential decision making by adding a set of perturbed measures surrounding a reference measure and an imaginary adversary who promotes robustness. This concept can be traced back to the early works by Petersen et al. (2000); Hansen and Sargent (2001); Hansen et al. (2006) and has found applications in a wide range of problem domains, including portfolio optimization (Maenhout, 2004), corporate investment (Nishimura and Ozaki, 2007), probability of lifetime ruin (Bayraktar and Zhang, 2015), optimal maintenance (Kim, 2016), among others. This body of studies typically uses the relative entropy to quantify the distance between probability measures. To our knowledge, few papers have gone beyond entropic constraint or penalty in formulating their robust optimization/control problems. While the use of entropy-based measures is conceptually appealing, allowing for a broader choice of uncertainty sets, as we do in this paper, can be theoretically interesting and practically valuable as it offers more flexibility in the modeling and pricing of actual products. It is also worth observing that different types of uncertainty sets may lead to further qualitatively different robust pricing strategies.

From a methodological point of view, the (standard) assumption that the decision maker is risk-neutral and ambiguity-averse would lead to a max-min or min-max optimization problem, which can be solved through the so-called Bellman-Isaacs equation. A Bellman-Isaacs equation is, in essence, a nonlinear differential equation. Addressing relevant fundamental questions (e.g., existence and uniqueness of solutions) associated with the equation often makes up the major technical hurdle in this body of work. Indeed, one major technical hurdle in Lim and Shanthikumar (2007), as the authors argue, is that the Bellman-Isaacs equation does not seem to satisfy global Lipschitz continuity conditions, preventing the application of the Picard-Lindelöf theorem to establish solvability directly. The authors of that paper manage to circumvent this barrier via an ingenuous nonlinear transformation that results in a set of related equations for which solvability can be established. Here, the adoption of general penalty rate functions prevents us from carrying out a similar transformation to our Bellman-Isaacs equation as we lose the nice properties of the penalty rate function entailed by the relative entropy. To overcome this difficulty, we devise some new tools which we believe carry methodological novelty.

To summarize, the paper makes two contributions to the extant literature. First, we develop a methodology for pricing a single perishable product when the pricing agent faces model uncertainty and is ambiguity-averse. In particular, this methodology generalizes the “entropic approach” in a meaningful way, enabling the design of more flexible uncertainty sets. Though the paper focuses on a single-product pricing problem, the idea may be adapted to handle other problems that can be modeled as a controlled Markov chain under model misspecification. Second, through a few carefully designed numerical experiments, we provide evidence that considering robustness
against model misspecification can indeed yield value. In particular, we demonstrate that good decisions can be made based on a misspecified model.

The remainder of this paper is organized as follows. In Section 2 we review the classical pricing problem and introduce our robust control problem faced by an economic agent who is pricing a single product. In Section 3 we state and prove our main results concerning the Bellman–Isaacs equation and its connection to the robust problem. In Section 4 we consider several special cases; those special cases inspire the modeling framework introduced in Section 2. Section 5 ends the paper by presenting numerical examples and two simulation programs to illustrate the “value of robustness”.

2 MODEL FORMULATION

Our starting point, as in Lim and Shanthikumar (2007), is the single-product dynamic revenue management problem proposed by Gallego and van Ryzin (1994). All relevant random quantities are carried by a filtered complete probability space \((\Omega, \mathcal{F}, \mathbb{F} := \{\mathcal{F}_t\}, \mathbb{P})\) satisfying the usual conditions.

2.1 A Classical Pricing Problem

There are \(c\) perishable items putting on sale over the finite time horizon \([0, T]\). \(N(t)\) is an \(\mathbb{F}\)-Poisson process that tracks the number of sales made up to time \(t\). Let \(p(t)\) denote the price posted at time \(t\). It follows that the total revenue over the time period \([0, T]\) can be expressed by \(\int_0^T p(t) dN(t)\).

We assume the demand rate to be Markov in the price; that is, \(N\) is a non-homogeneous Poisson process whose intensity at time \(t\) is \(\lambda(t) = \lambda(p(t))\) for some real-valued function \(\lambda(\cdot)\). We impose the following regularity conditions on the function \(\lambda\).

Assumption 1

(i) \(\lambda(p)\) is differentiable and strictly decreases with \(\lambda(0) = b < \infty\). Moreover, there exists a null price \(\bar{p}\) such that \(\lim_{p \to \bar{p}} \lambda(p) = 0\). (ii) \(h(p) := -\lambda'(p)/\lambda(p)\) is non-decreasing.

Assumption 1(i) conforms with those made in Gallego and van Ryzin (1994). Assumption 1(ii) is also standard. To see this, suppose that the demand rate model is specified by

\[\lambda(p) = b\mathbb{P}(V > p),\]

where \(b\) denotes the potential demand rate and \(V\) represents a customer’s evaluation of the merits of the product. Assumption 1(ii) is equivalent to requiring the customer valuation distribution to have a non-decreasing hazard rate, which is a standard assumption from revenue management guaranteeing the existence of a unique maximizer of the revenue rate function \(r(p) := p\lambda(p)\). It
is straightforward to verify that a function in the form of \( \lambda(p) = Ce^{-Dp} \) with \( C, D > 0 \) meets all the aforementioned requirements: It is continuously differentiable and strictly decreasing in \( p \); additionally, we have \( \bar{p} = \infty \) and \( 0 \leq \lambda(p) \leq C \) for all \( p \in [0, \infty) \). Finally, \( h(p) = D \) for all \( p \in [0, \infty) \) and hence it is a non-decreasing function.

A pricing policy \( p \) is said to be admissible if it is \( \mathcal{F}_t \)-predictable. Given that the pricing agent has full confidence in \( \mathbb{P} \), the optimal pricing problem is to choose an admissible pricing policy \( p \) to

\[
\max \mathbb{E} \left[ \int_0^T p(u) \, dN(u) \right] \quad \text{subject to} \quad N(T) \leq c,
\]

where \( \mathbb{E} \) denotes the expectation taken with respect to \( \mathbb{P} \). Since the demand rate is Markov in the price, one can focus on Markov strategies without loss of generality; that is, the pricing decision at time \( t \) depends solely on current time \( t \) and inventory level \( X(t) = c - N(t) \).

### 2.2 Model Uncertainty & Relative Entropy

We have assumed that each fixed pricing strategy \( p \) induces an arrival process \( N \) whose intensity at time \( t \) is \( \lambda(t) := \lambda(p(t)) \) under the measure \( \mathbb{P} \). In reality, the intensity process may not be \( \lambda \) but rather some other process, \( \beta \). In other words, the agent may be doubtful about the measure \( \mathbb{P} \) and therefore wish to consider alternative measures \( \mathbb{Q} \). To capture the idea that the pricing agent only sees \( \mathbb{P} \) as an approximation of the real-world scenario (i.e., a reference measure), one can surround \( \mathbb{P} \) with a cloud of measures that are difficult to distinguish with limited data and add a malevolent second agent (i.e., nature) whose role is to promote robustness. In essence, nature prompts the pricing agent to assess the fragility of the alternative decision rules for deviations from the reference measure in the actual world.

Lim and Shanthikumar (2007) puts the notion into practice by introducing two formulations of the robust pricing problem. Both formulations use a definition of the relative entropy (an expected log likelihood ratio) to constrain the gap between the reference measure and a perturbation to it. The first formulation, called the *constrained problem*, seeks to

\[
\max \min_q \mathbb{E}^q \left[ \int_0^T p(u) \, dN(u) \right] \quad \text{subject to} \quad \mathbb{E}^q [D(\mathbb{Q} \| \mathbb{P})] \leq \gamma \quad \text{and} \quad N(T) \leq c,
\]

where \( q \) denotes nature’s strategy, \( \mathbb{Q} \) denotes the real-world measure realized under the strategy \( q \), \( D(\mathbb{Q} \| \mathbb{P}) \) represents the relative entropy of \( \mathbb{Q} \) with respect to \( \mathbb{P} \) on \( \mathcal{F}_T \), and \( \gamma \) determines the size of the uncertainty set. The requirement that entropy be finite limits the form that model misspecification can occur. In particular, a finite entropy requires that admissible perturbations of the reference measure be absolutely continuous with respect to it over finite intervals. The second
formulation, referred to as the **penalty problem**, seeks to

$$\text{maximize } \min_{q} E^{q}\left[ \int_{0}^{T} p(u) dN(u) + \theta D(Q||P) \right] \quad \text{subject to } N(T) \leq c, \quad (1)$$

where $\theta$ is naturally interpreted as the Lagrange multiplier for the relative entropy constraint. As noted by Lim and Shanthikumar (2007), the parameter $\theta$ also measures the pricing agent’s confidence level in the nominal measure. A large value of $\theta$ imposes a large penalty on deviations from $P$ by nature, corresponding to pricing agent’s high level of confidence in $P$.

Both aforementioned formulations give rise to a two-player dynamic game wherein the pricing agent, who is also the maximizing player, takes nature’s reactions into account in contemplating pricing strategies. To express the dynamic game more explicitly, one can invoke the Girsanov theorem to portray a $P$-absolutely continuous measure $Q$ via its Radon-Nikodym derivative with respect to $P$:

$$\frac{dQ}{dP}_{\mathcal{F}_T} = \theta(T) = \exp\left( \int_{0}^{T} \ln \psi(u) dN(u) + \int_{0}^{T} (1 - \psi(u)) \lambda(u) du \right), \quad (2)$$

where $N$ is the demand arrival process introduced previously and $\psi := \{\psi(t); t \geq 0\}$ is an $\mathcal{F}$-predictable process satisfying proper regularity conditions. Moreover, the Girsanov theorem guarantees that $N$ is an $\mathcal{F}$-adapted Poisson process with intensity $\beta(t) = \psi(t) \lambda(t)$ under $Q$. This means that specifying an alternative measure $Q$ is tantamount to selecting an $\mathcal{F}$-predictable process $\psi$. In other words, one can think of $\psi$ as a probability distortion process that nature employs to cast a specific perturbation of the reference measure $P$. Therefore, we can parameterize the class of alternative measures by $\psi$. We will henceforth use $P^\psi$ to denote the measure induced by $\psi$.

In general, for two measures $P$ and $Q$, the relative entropy of $Q$ with respect to $P$ is defined as

$$D(Q||P) := \int \ln \left( \frac{dQ}{dP} \right) dQ. \quad (3)$$

Thus, for the penalty problem in particular, one can use (2) and (3) to express the relative entropy of $P^\psi$ with respect to $P$ on $\mathcal{F}_T$ by

$$D(P^\psi||P) = \int_{0}^{T} \lambda(p(u)) [\psi(u) \ln \psi(u) + 1 - \psi(u)] du.$$

In the spirit of the penalty formulation, we can restate the objective in (1) that the agent chooses
an admissible pricing strategy $p$ to maximize

$$\min_{\psi} \mathbb{E}[\psi] \left[ \int_0^T p(u) dN(u) + \theta \int_0^T \lambda(p(u)) \left[ \psi(u) \ln \psi(u) + 1 - \psi(u) \right] du \right] \quad \text{subject to} \quad N(T) \leq c,$$

(4)

where we add the superscript $\psi$ to the expectation operator to indicate the average value when nature employs the strategy $\psi$.

### 2.3 A General Framework for Promoting Robustness

The problem described by (4) suggests the following: By choosing $\psi(t)$ at time $t$, nature gets penalized at the rate $\theta \left[ \psi(t) \ln \psi(t) + 1 - \psi(t) \right]$ scaled by $\lambda(p(t))$. The fact that the penalty rate function does not have to take this particular form is a key observation that motivates the current research. In fact, many other functions can be used as the penalty rate function (when scaled by $\lambda(p)$). Therefore, one of the primary goals of this paper is to advance a robust decision framework that is more flexible than the entropic approach. Of course, our approach also assumes that the actual intensity process $\beta$, albeit not known precisely, is such that the corresponding measure governing the realized demands is absolutely continuous with respect to the reference measure $\mathbb{P}$ on $\mathcal{F}_T$.

Specifically, we define the penalty rate at time $t$ as $\lambda(p(t))f(\psi(t))$ where $f(\cdot)$ is a function that attains a minimum value of zero at $x = 1$. The intuition is that $\psi \equiv 1$ corresponds to the reference measure, and as a result, the penalty should be zero whenever nature sets $\psi$ to 1. (We will discuss the choice of $f$ in the next section). To continue, let us define

$$g(y) := \min_x \{ xy + f(x) \} = -\max_x \{ -xy - f(x) \} = -f^*(-y),$$

(5)

where we have used $f^*$ to denote the Legendre transformation of the function $f$. It is easy to check that $g(0) = 0$ due to our hypothesis that $f$ attains its minimum value zero at $x = 1$. Also, let $(-A, B] \cap \mathbb{R}$ denote the domain of $g$ with both $A$ and $B$ being greater than zero and possibly infinity. We will impose the following regularity conditions on $f$ through its transform, $g$.

**Assumption 2**

(i) There exists some positive constant $\eta$ such that $g$ is Lipschitz continuous on the interval $[-\eta, B] \cap \mathbb{R}$ with the Lipschitz constant being $L$. (ii) The function $\ell(y) := g'(y)/g(y)$ is continuous and strictly decreasing on $(0, B)$ with $\lim_{y \to 0} \ell(y) = \infty$ and $\lim_{y \to B} \ell(y) = 0$.

With these preparations, we can formally state our robust revenue maximization problem: The
agent chooses a pricing strategy $p$ from the set of admissible policies to

$$\text{maximize} \quad \min_\psi \quad \mathbb{E}^\psi \left[ \int_0^T (p(u) dN(u) + \lambda(p(u)) f(\psi(u)) du) \right] \quad \text{subject to} \quad N(T) \leq c,$$

where we have used $\mathbb{E}^\psi$ to denote expectation with respect to $\mathbb{P}^\psi$ (i.e., the measure induced by a probability distortion process $\psi$). An implicit assumption behind this formulation is that at time zero, both players commit to decision strategies whose time $t$ components are measurable functions of $F_t$: The minimizing player who chooses distorted beliefs $\psi$ takes $p$ as given while the maximizing player who chooses $p$ takes $\psi$ as given. It may be worth reiterating that the minimizing player’s malevolence serves as a useful tool which the maximizing player (i.e., the agent) can use to analyze the fragility of alternative decision rules. In addition, the function $f$ is chosen to determine both the “shape” and the size of the uncertainty set. In particular, a large (small) penalty implied by $f$ entails weak (strong) ambiguity about the intensity of arrival process $N$. This is because large penalties discourage nature from choosing a real-world measure that is far away from the reference measure $\mathbb{P}$. Let $V(t, n)$ denote the value function for the problem (6), where $n$ is understood to be the product stock level at time $t$. Noticing that

$$\mathbb{E}^\psi \left[ \int_0^T p(u) dN(u) \right] = \mathbb{E}^\psi \left[ \int_0^T p(u) \beta(u) du \right] = \mathbb{E}^\psi \left[ \int_0^T p(u) \psi(u) \lambda(u) du \right],$$

with reference to the general theory of Markov decision process, we conjecture that the value function $V(t, n)$, if it exists, will satisfy the following Bellman-Isaacs equation:

$$0 = V_t(t, n) + \max_p \min_\psi \{ \psi p + \psi [V(t, n - 1) - V(t, n)] + f(\psi) \},$$

$$V(T, n) = 0 \quad \text{for} \quad n = 0, 1, \ldots, c,$$

$$V(t, 0) = 0 \quad \text{for all} \quad t \in [0, T].$$

The properties concerning the Bellman-Isaacs equation, such as existence and uniqueness of solutions, will be the subject of the next section. Using the definition of the function $g$, we can “remove” the inner minimization problem in (7) and rewrite it as

$$0 = V_t(t, n) + \max_p \lambda(p) g(p + V(t, n - 1) - V(t, n)),$$

$$V(T, n) = 0 \quad \text{for} \quad n = 0, 1, \ldots, c,$$

$$V(t, 0) = 0 \quad \text{for all} \quad t \in [0, T].$$
We anticipate that the optimal pricing policy is obtained by finding the maximum \( p \) in (8); i.e.,

\[
p^\star(t) := p^\star(t, X(t-)) = \inf \arg \max_p \lambda(p) g(p + V(t, X(t-)) - V(t, X(t-)))
\]  

ought to be an optimal pricing rule under the max-min criterion as specified by (6). As we demonstrate below, under Assumptions 1(ii) and 2(ii), the right-hand side of (9) admits a unique maximum. Also, if letting

\[
\psi^\#(t, X(t-), p(t)) = \inf \arg \min_\psi \lambda(p(t)) \{\psi [p(t) + V(t, X(t-)) - 1) - V(t, X(t-))] + f(\psi)\}
\]

be a response by nature to a given pricing policy \( p \), we anticipate \( \psi^\star := \psi^\#(t, X(t-), p^\star(t)) \) to be nature’s best response to \( p^\star \).

As in Lim and Shanthikumar (2007), explicit expression of the pricing rule \( p^\star \) and that of the “best response” \( \psi^\star \) by nature can not be expected unless specific functional forms are assumed for the demand rate function \( \lambda(p) \) and the penalty rate function \( f \). Nonetheless, we will demonstrate that a solution to (8) not only exists but is unique under our model assumptions.

3 | MAIN RESULTS

This section is devoted to the statements and proofs of our key results concerning the Bellman-Isaacs equation as well as its connection to the robust revenue maximization problem (6). To begin with, we present the following lemma that asserts the existence and uniqueness of solutions to the equation given by (8).

Lemma 1 Suppose both Assumptions 1 and 2 hold. Then there exists a unique solution \( J(t, n) \) to (8).

Proof of Lemma 1.
Consider following differential equation

\[
U_t(t, n) = \max_p \lambda(p) g(p + U(t, n - 1) - U(t, n)),
\]

\[
U(0, n) = 0 \quad \text{for} \quad n = 0, 1, \ldots, c,
\]

\[
U(t, 0) = 0 \quad \text{for all} \quad t \in [0, T].
\]

Clearly, if (11) admits a unique solution \( U(t, n) \), then \( V(t, n) \) defined via \( V(t, n) = U(T - t, n) \) uniquely solves (8). Hence, it suffices to prove existence and uniqueness of (11).
To start, define

\[ m_1(x) := \max_{\lambda(p)g(p - x)} \lambda(p)g(p - x). \]  

(12)

For each fixed \( x \geq 0 \), let \( p^* := p^*(x) \) denote a maximizer of the right-hand side of (12). We can use the first-order condition to deduce that \( p^* \) ought to satisfy

\[ -\frac{\lambda'(p^*)}{\lambda(p^*)} = \frac{g'(p^* - x)}{g(p^* - x)} \quad \text{or} \quad h(p^*) = \ell(p^* - x), \]

if it is in the interior of the range \([0, \bar{p}]\). Existence and uniqueness of such \( p^* \) then follows from Assumption 1(ii) and Assumption 2(ii). Also, we have that \( p^* > x \), which in view of (12) implies that \( m_1 \) is non-negative on \([0, \infty)\). Next, we show that the function \( m_1 \) is Lipschitz continuous on \([0, \infty)\). By appealing to the identity \( \sup_x f_1(x) - \sup_x f_2(x) \leq \sup_x (f_1(x) - f_2(x)) \), we deduce, for \( 0 \leq x_1 < x_2 \) with \( |x_1 - x_2| \leq \eta \), that

\[
\begin{align*}
    m_1(x_1) - m_1(x_2) &= \max_{p \geq x_1} \lambda(p)g(p - x_1) - \max_{p \geq x_2} \lambda(p)g(p - x_2) \\
    &= \max_{p \geq x_1} \lambda(p)g(p - x_1) - \max_{p \geq x_1} \lambda(p)g(p - x_2) \\
    &\leq \sup_{p \geq x_1} [g(p - x_1) - g(p - x_2)] \\
    &\leq bL |x_1 - x_2|,
\end{align*}
\]

where the last inequality follows from Assumption 2(ii) plus the fact that \( |x_1 - x_2| \leq \eta \). In the same vein, we can argue (by switching the roles of \( x_1 \) and \( x_2 \)) that \( m_1(x_2) - m_1(x_1) \leq bL |x_1 - x_2| \); hence

\[ |m_1(x_1) - m_1(x_2)| \leq bL |x_1 - x_2|. \]

The desired Lipschitz continuity of \( m_1 \) then follows from the fact that \( x_1 \) and \( x_2 \) are chosen arbitrarily. Since the differential equation corresponding to \( n = 1 \) can be written as \( U_1(t, 1) = m_1(U(t, 1)) \), where \( m_1 \) has shown to be non-negative and Lipschitz continuous on \([0, \infty)\), existence and uniqueness of \( U(t, 1) \) follow by applying the Picard-Lindelöf theorem.

To establish existence and uniqueness results for \( n = 2 \), let

\[ m_2(t, x) := \max_{p} \lambda(p)g(p + U(t, 1) - x). \]  

(13)

Then the differential equation corresponding to \( n = 2 \) can be written as \( U_2(t, 2) = m_2(t, U(t, 2)) \). Existence and uniqueness of \( U(t, 2) \) will follow if \( m_2(t, x) \) can be shown to be non-negative and uniformly Lipschitz continuous in \( x \) for \( x \geq 0 \). With a slight abuse of notation, let \( p^* := p^*(t, x) \) be
a maximizer of the right-hand side of (13). Using the first-order condition, we know that \( p^* \) ought to be such that

\[
h(p^* + U(t, 1)) = \ell(p^* + U(t, 1) - x).
\]

Again, we can assert existence and uniqueness of \( p^* \) by appealing to Assumption 1(ii) and Assumption 2(ii); moreover, we have that \( p^* + U(t, 1) \geq x \), which implies that \( m_2(t, x) \) is non-negative for \( x \geq 0 \). Next, for \( 0 \leq x_1 < x_2 \) with \( |x_1 - x_2| \leq \eta \), we have that

\[
m_2(t, x_1) - m_2(t, x_2) = \max_{p \geq x_1 - U(t, 1)} \lambda(p)g(p + U(t, 1) - x_1) - \max_{p \geq x_2 - U(t, 1)} \lambda(p)g(p + U(t, 1) - x_2)
\]

\[
= \max_{p \geq x_1 - U(t, 1)} \lambda(p)g(p + U(t, 1) - x_1) - \max_{p \geq x_1 - U(t, 1)} \lambda(p)g(p + U(t, 1) - x_2)
\]

\[
\leq \sup_{p \geq x_1 - U(t, 1)} \lambda(p)[g(p + U(t, 1) - x_1) - g(p + U(t, 1) - x_2)]
\]

\[
\leq bL|x_1 - x_2|,
\]

where again the last inequality uses Assumption 2(i). By switching the roles of \( x_1 \) and \( x_2 \) in the above calculation, we conclude that \( m_2(t, x_2) - m_2(t, x_1) \leq bL|x_1 - x_2| \); hence

\[
m_2(t, x_1) - m_2(t, x_2) \leq bL|x_1 - x_2|.
\]

The desired uniform Lipschitz continuity of \( m_2 \) in \( x \) then follows by noting that \( x_1 \) and \( x_2 \) are chosen arbitrarily plus the fact that the preceding analysis is independent of \( t \). To encapsulate, we have shown that \( m_2(t, x) \) is non-negative and uniformly Lipschitz continuous on \( x \geq 0 \); therefore, existence and uniqueness of \( U(t, 2) \) follows from the Picard-Lindelöf theorem. The existence and uniqueness of (11) for every \( n = 3, \ldots, c \) can be established in an analogous way.

Before proceeding, we would like to comment on the technique used by Lim and Shanthikumar (2007) to establish the solvability of the Bellman-Isaacs equation. When \( f(x) = \theta(x \ln x + 1 - x) \), the transform \( g \) can be calculated as \( g(y) = \theta(1 - e^{-y/\theta}) \). Plugging this into (11) gives us

\[
U_1(t, n) = \theta \max_p \lambda(p) \left\{ 1 - \exp \left[ -\frac{1}{\theta} (p + U(t, n - 1) - U(t, n)) \right] \right\}.
\]

Because this equation fails to satisfy the global Lipschitz continuity condition, the Picard-Lindelöf theorem cannot be applied directly. The nonlinear transformation used by Lim and Shanthikumar (2007) to establish the existence and uniqueness of solutions to the Bellman-Isaac equation corresponds to letting \( Q(t, n) = -\ln U(t, n) \). With this transformation, we can translate the preceding
equation into

\[ Q_t(t, n) = \min_{\rho} \lambda(\rho) \left[ \rho + Q(t, n-1)e^{-\rho/\theta} - Q(t, n) \right], \]

which indeed satisfies the global Lipschitz continuity condition. However, in our model, the penalty rate is represented by a general function \( f \), and we are not able to identify such a nonlinear transformation because \( f \) lacks an explicit formula. Therefore, our method for establishing the existence and uniqueness of the optimality equation differs fundamentally from that used in Lim and Shanthikumar (2007).

Next, we show that the pricing rule \( p^* \) given by (9) is indeed optimal for the robust revenue management problem (8). This is guaranteed by the following verification theorem. The proof of this result is of the same nature as that of Theorem 4.1 in Lim and Shanthikumar (2007) but streamlines some of the steps taken in that paper.

**Theorem 1** Suppose both Assumptions 1 and 2 hold and let \( V(t, n) \) be the solution of (8).

(i) For every admissible \( p \), we have

\[ V(t, n) \geq \min_{\psi} \mathbb{E}^\psi \left[ \int_t^T (p(t)dN(t) + \lambda(p(t))f(\psi(t))dt) \bigg| X(t-) = n \right]. \]

(ii) For the pricing rule \( p^* \) given by (9), it holds that

\[ V(t, n) = \mathbb{E}^{\psi^*} \left[ \int_t^T (p^*(u)dN(u) + \lambda(p^*(u))f(\psi^*(u))du) \bigg| X(u-) = n \right] \]

\[ = \max_{\rho} \min_{\psi} \mathbb{E}^\psi \left[ \int_t^T (p(u)dN(u) + \lambda(p(u))f(\psi(u))du) \bigg| X(u-) = n \right], \]

where \( \psi^* \) is defined via \( \psi^*(t) := \psi^#(t, X(t-), p^*(t)) \) for \( \psi^# \) given by (10).

**Proof of Theorem 1.**

Recall from model description that \( X(t) \) represents the remaining stock level at time \( t \). Applying Ito’s formula for jump processes yields

\[ dV(t, X(t)) = V_t(t, X(t-))dt + [V(t, X(t-)) - V(t, X(t-))] dN(t). \]
Note that $\psi^\#$ defined via (10) is an admissible strategy for nature. Thus, integrating (15) from $t$ to $T$ and taking expectation with respect to the measure induced by $\psi^\#$ gives us

$$0 = V(t, n) + E^{\psi^\#} \left[ \int_t^T \left\{ V_t(u, X(u)) + \lambda(p(u))\psi^\#(u) \times [V(u, X(u) - 1) - V(u, X(u))] \right\} du \right]$$

$$= V(t, n) + E^{\psi^\#} \left[ \int_t^T \left\{ V_t(u, X(u)) + \lambda(p(u))g(p(u) + V(u, X(u) - 1) - V(u, X(u))) \right. \right.$$  

$$- \lambda\lambda(p(u))\psi^\#(u)\rho(u) - \lambda(p(u))f(\psi^\#(u)) \right) du \right]$$

$$\leq V(t, n) - E^{\psi^\#} \left[ \int_t^T \left\{ \lambda(p(u))\psi^\#(u)\rho(u) - \lambda(p(u))f(\psi^\#(u)) \right\} du \right],$$

where the last inequality is due to (8). Rearranging terms yields

$$V(t, n) \geq E^{\psi^\#} \left[ \int_t^T \left\{ \lambda(p(u))\psi^\#(u)\rho(u) - \lambda(p(u))f(\psi^\#(u)) \right\} du \right]$$

$$\geq \min_{\psi} E^{\psi} \left[ \int_t^T \left\{ \lambda(p(u))\psi^\#(u)\rho(u) - \lambda(p(u))f(\psi^\#(u)) \right\} du \right],$$

where the second inequality draws on the admissibility of $\psi^\#$. Part (i) of the theorem then follows by noting that $N$ is a non-homogeneous Poisson process with intensity $\lambda(p(t))\psi(t)$ under the measure induced by $\psi$.

Towards proving part (ii), note that the first equality in (14) follows from the observation that the last inequality in (16) holds as an equality by choosing $\rho = \rho^\ast$. To establish the second equality in (14), we need to demonstrate that $\psi^\ast$ is minimizing player's best response to the pricing policy $\rho^\ast$. To this end, we substitute $\rho^\ast$ back to the first line of (8) to get

$$0 = V_t(t, n) + \lambda(p^\ast(t))g(p^\ast(t) + V(t, n - 1) - V(t, n)),$$

which, in view of (5), is equivalent to

$$0 = V_t(t, n) + \min_{\psi} \lambda(p^\ast(t)) \left\{ \psi [p^\ast(t) + V(t, n - 1) - V(t, n)] + f(\psi) \right\}. \quad (17)$$

A standard verification argument involving a straightforward application of Ito's formula as in (15) would show that the solution to (17) with boundary conditions given in (8) coincides with the value function for nature's minimization problem with the pricing strategy fixed at $\rho^\ast$. As a result, $\psi^\ast$ is nature's best response to the pricing policy $\rho^\ast$. This completes the proof of part (ii). □
4 | SOME SPECIAL CASES

In this section, we solve the resultant robust dynamic pricing problem when specific functional forms are specified for the penalty rate function $f$. While there is a lot of flexibility with the choice of $f$, we are primarily interested in those designed based on divergence measures; these include the entropic penalty and Rényi-type penalty to be discussed in §4.1 and §4.2 respectively.

4.1 | Entropic Penalty

Recall from §2 that when the amount of penalties paid by nature over the time duration $[0, T]$ is proportional to the distance of $\bar{\psi}$ from $\bar{\psi}$ measured by the relative entropy, the objective of the agent becomes to choose an admissible pricing strategy $p$ to maximize (4), wherein $\theta$ is a constant introduced to represent the agent’s level of ambiguity aversion. Intuitively, larger (lower) values of $\theta$ correspond to heavier (lighter) penalty for nature, hence lesser (greater) ambiguity aversion of the agent. Upon close inspection, we see that the general formulation (6) specializes to (4) by setting

$$f(x) = \theta(x \ln x + 1 - x). \tag{18}$$

Accordingly, the transform $g$ associated with $f$ can be calculated as

$$g(y) = \min_x \{ xy + \theta(x \ln x + 1 - x) \} = \theta(1 - e^{-y/\theta}).$$

Clearly, the function has a domain $(-\infty, \infty)$ and satisfies part (i) of Assumption 2. Now, direct calculations yield

$$g'(y) = e^{-y/\theta} \quad \text{and} \quad \ell(y) = \frac{e^{-y/\theta}}{\theta(1 - e^{-y/\theta})} = \frac{1}{\theta(e^{y/\theta} - 1)}.$$ 

It is straightforward to check that $\ell$ is strictly decreasing on $(0, \infty)$ and satisfies $\lim_{y \to 0} \ell(y) = \infty$ and $\lim_{y \to \infty} \ell(y) = 0$. Thus, Assumption 2(ii) also holds. Having verified all the conditions in Assumption 2, we obtain the following result as a direct consequence of Theorem 1.

**Corollary 1** Suppose Assumption 1 holds and let $V(t, n)$ be the solution of (8) wherein $f$ is specified by (18). (i) For every admissible $p$, we have

$$V(t, n) \geq \min_{\psi} \mathbb{E}^{\psi} \left[ \int_t^T \{ p(u) dN(u) + \theta \lambda(p(u)) [\psi(u) \ln \psi(u) + 1 - \psi(u)] du \} \mid X(t-) = n \right].$$
(ii) For the pricing rule \( p^\star \) given by (9), it holds that
\[
V(t, n) = \mathbb{E}^\Psi \left[ \int_t^T \{ p^\star(u) dN(u) + \theta \lambda(p^\star(u)) [\psi^\star(u) \ln \psi^\star(u) + 1 - \psi^\star(u)] \} \right] \}
\]
\[
= \max_{\psi} \min_{\psi^\star} \mathbb{E}^\Psi \left[ \int_t^T \{ p(u) dN(u) + \theta \lambda(p(u)) [\psi(u) \ln \psi(u) + 1 - \psi(u)] \} \right] \}
\]
where \( \psi^\star \) is defined via \( \psi^\star(t) := \psi^\#(t, X(t-), p^\star(t)) \).

### 4.2 Rényi-Type Penalty

Given two probability measures \( \mathbb{P} \) and \( \mathbb{Q} \), Rényi divergence of order \( \alpha \)
\[
\mathcal{R}_\alpha(\mathbb{Q}||\mathbb{P}) := \frac{1}{\alpha - 1} \ln \left( \int \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right)^{\alpha - 1} d\mathbb{Q} \right). \tag{19}
\]
Given \( \mathbb{Q} \) and \( \mathbb{P} \) fixed, \( \mathcal{R}_\alpha \) is continuous in \( \alpha \) on \( (0, 1) \cup (1, \infty) \), and it tends to the relative entropy as \( \alpha \) approaches 1. This leads to defining Rényi divergence of order \( \alpha = 1 \) as the relative entropy, so one may as well assume \( \alpha \in (0, \infty) \). Moreover, \( \mathcal{R}_\alpha(\mathbb{Q}||\mathbb{P}) \) is convex in \( \mathbb{Q} \) when \( \alpha \in (0, 1] \) and quasi-convex in \( \mathbb{Q} \) when \( \alpha \in (1, \infty) \). Van Erven and Harremos (2014) can be consulted for detailed proofs of these results.

A straightforward yet tedious calculation involving (2), the definition (19) and the moment generating function of Poisson distribution allows us to obtain
\[
\mathcal{R}_\alpha(\mathbb{P}^\Psi||\mathbb{P}) = \frac{1}{(\alpha - 1)} \int_0^T \lambda(p(u))(\psi(u)^\alpha - \alpha \psi(u) + \alpha - 1) \, du \quad \text{for} \quad \alpha \neq 1. \tag{20}
\]
The proof sketch of this formulation is given in Appendix A.

Suppose now that in the two-player zero-sum game nature gets penalized in proportion to the distance of the perturbed measure \( \mathbb{P}^\Psi \) from \( \mathbb{P} \) gauged by Rényi divergence. Then the robust control problem for the agent is to choose an admissible pricing strategy \( p \) to maximize
\[
\min_{\psi} \mathbb{E}^\Psi \left[ \int_0^T p(u) dN(u) + \frac{\theta}{(\alpha - 1)} \int_0^T \lambda(p(u))(\psi(u)^\alpha - \alpha \psi(u) + \alpha - 1) \, du \right] \tag{21}
\]
subject to \( N(T) \leq c \), where again the constant \( \theta \) captures the agent’s level of ambiguity aversion.
This corresponds to choosing
\[
f(x) = \frac{\theta}{\alpha - 1} (x^\alpha - \alpha x + \alpha - 1) \tag{22}
\]
in our general formulation. To continue our verification process as done in §4.1, we distinguish
between two different cases, namely, \( \alpha \in (0, 1) \) and \( \alpha \in (1, \infty) \).

### 4.2.1 The case of \( \alpha \in (0, 1) \)

Using the definition of \( g \), we deduce that

\[
g(y) = \theta \left( 1 - \left( 1 - \frac{\zeta}{\theta} y \right)^{\frac{1}{\alpha}} \right),
\]

where \( \zeta := \zeta(\alpha) = (\alpha - 1)/\alpha < 0 \). Also, it is easy to see that \( g \) is defined on \((\theta/\zeta, \infty)\) and Lipschitz continuous on \([\theta/(2\zeta), \infty)\); hence by letting \(-\eta = \theta/(2\zeta)\) or \(\eta = -\theta/(2\zeta)\), Assumption 2(i) is satisfied. With regard to Assumption 2(ii), direct calculations give us

\[
g'(y) = \left( 1 - \frac{\zeta}{\theta} y \right)^{\frac{1}{\alpha} - 1} \quad \text{and} \quad \ell(y) = \frac{1}{\theta (1 - \frac{\zeta y}{\theta})^{1-1/\alpha}} - \frac{1}{1 - \frac{\zeta y}{\theta}}.
\]  

(24)

A close inspection reveals that \( \ell \) is strictly decreasing on \((0, \infty)\) and satisfies \( \lim_{y \to 0} \ell(y) = \infty \) and \( \lim_{y \to \infty} \ell(y) = 0 \). Hence all the conditions in Assumption 2 hold true for all \( \alpha \in (0, 1) \).

### 4.2.2 The case of \( \alpha \in (1, \infty) \)

For all \( \alpha \in (1, \infty) \), expressions for \( g \), \( g' \) and \( \ell \) are exactly the same as those in (23) and (24). However, rather than being defined on \((\theta/\zeta, \infty)\), the function \( g \) now lives on the domain \((-\infty, \theta/\zeta]\) and it is Lipschitz continuous on \([-\kappa, \theta/\zeta]\) for all finite \( \kappa > 0 \); hence we have verified Assumption 2(i). Furthermore, it is readily seen that \( \ell \) is strictly decreasing on \((0, \theta/\zeta]\) and satisfies \( \lim_{y \to 0} \ell(y) = \infty \) and \( \ell(\theta/\zeta) = 0 \). Therefore, all the requirements in Assumption 2 are satisfied for all \( \alpha \in (1, \infty) \).

We summarize the discussions in §4.2.1 and §4.2.2 as the following corollary which can be seen as a direct consequence of Theorem 1.

**Corollary 2** Suppose Assumption 1 holds and let \( V(t, n) \) be the solution of (8) wherein \( f \) is specified by (22). (i) For every admissible \( \rho \), we have

\[
V(t, n) \geq \min_{\psi} \mathbb{E}^\psi \left[ \int_t^T \left( \rho(u) \mathrm{d}N(u) + \frac{\theta \lambda(\rho(u))}{\alpha - 1} \left[ (\psi(u))^\alpha - \alpha \psi(u) + \alpha - 1 \right] \mathrm{d}u \right) \right] X(t-) = n.
\]

(ii) For the pricing rule \( \rho^* \) given by (9), it holds that

\[
V(t, n) = \mathbb{E}^\psi \left[ \int_t^T \left( \rho^*(u) \mathrm{d}N(u) + \frac{\theta \lambda(\rho^*(u))}{\alpha - 1} \left[ (\psi^*(u))^\alpha - \alpha \psi^*(u) + \alpha - 1 \right] \mathrm{d}u \right) \right] X(t-) = n
\]
\[
= \max_{\rho} \min_{\psi} \mathbb{E}^\psi \left[ \int_T^t \left( \rho(u) dN(u) + \frac{\theta \lambda(p(u))}{\alpha - 1} \left[ \psi(u)^\alpha - \alpha \psi(u) + \alpha - 1 \right] du \right) \bigg| X(t-) = n \right],
\]

where \( \psi^* \) is defined via \( \psi^*(t) := \psi^t(t, X(t-), \rho^*(t)) \).

5 | NUMERICAL STUDIES AND SIMULATION

This section serves two purposes. First, in §5.1, we present numerical studies providing insights and understanding of the effects of various factors on the optimal pricing strategy. Second, in §5.2, we demonstrate how the concept of robustness can add value to decision-making. Specifically, we show that our solution framework allows us to make “right” decisions based on a “wrong” model.

5.1 | Pricing Under Ambiguity

The first step of our numerical study is to numerically solve the value function \( V(t, n) \) for \( n = 1, \ldots, c \) in a recursive way (starting from \( n = 1 \)). Based on the value function, we can find the optimal pricing policy \( \rho(t, n) \) as well as nature’s best response \( \psi \). The algorithm we employ to solve the value function is given in Appendix B. We mainly focus on the class of uncertainty sets designed based on the notion of Rényi divergence. This means that we will solve our robust pricing model with \( f \) being in the form of (22). We also consider \( \alpha = 1 \), in which case the penalty rate function is considered to be in the form of (18). For each fixed \( \alpha \), we obtain uncertainty sets of different sizes by varying the value \( \theta \). In our numerical studies, we set \( c = 4, T = 20 \) and use exponential demand rate model \( \lambda(p) = Ce^{-Dp} \) with \( C = D = 1 \). Figures 1, 2 and 3 depict the optimal pricing policies obtained by using \( \alpha = 1/2, 1 \) and 2, respectively. Within each plot, pricing policies with different degrees of ambiguity are put in contrast. The value functions used to produce these pricing policies are displayed in Appendix C.

First, for fixed \( \theta, \alpha \) and \( n \), we observe that the pricing function \( \rho(t, n) \) is a decreasing function of \( t \). This is intuitive because, as the selling season approaches its end, a firm tends to increase its sales volume by cutting prices on the products. Second, by fixing the uncertainty set (i.e., fixing for \( \alpha \) and \( \theta \)), we see that all pricing functions converge to some certain value. This is because, as the end of the sales period approaches, the firm will seek to sell products at the lowest price they can accept. It is also evident that \( \rho(t, k_1) < \rho(t, k_2) \) if \( k_1 > k_2 \). The reason is that too many products in stock would put high pressure on the firm’s sales efforts, leading to low prices for large values of \( n \). Third, for each fixed \( \alpha \), the price tends to be higher when \( \theta \) becomes large. This suggests that when facing a high degree of model uncertainty, the firm tends to lower its price. Moreover, the effect is more pronounced at the beginning of the selling season, as evidenced by the wider gap
between the solid and dash curves for small \( t \). Finally, by comparing Figures 1, 2 and 3, we find that for a fixed \( \theta \), the optimal price at a given point of time increases in \( \alpha \). This is because Rényi divergence is non-decreasing as a function of its order \( \alpha \), meaning that as \( \alpha \) grows, the penalty imposed on nature becomes larger; hence the pricing solution becomes less conservative.

5.2 | Exposing the Value of Robustness

In this section, we will use two examples to explain the importance of considering robustness. First, we consider a case where demand arrivals follow a non-homogeneous Poisson process whose
intensity is a Cox–Ingersoll–Ross (CIR) process. To be specific, $\lambda(t)$ is the solution to the following stochastic differential equation

$$d\lambda(t) = k(\bar{\lambda}(t) - \lambda(t))dt + \sigma \sqrt{\lambda(t)} d\tilde{B}(t),$$

(25)

where $\tilde{B}$ is a standard Brownian motion independent of everything else, and $\bar{\lambda}(t) = Ce^{-Dp(t)}$, with $C$ and $D$ being positive constants. In the above, parameter $k$ corresponds to the speed of adjustment to the mean process $\bar{\lambda}$, and $\sigma$ measures the variability of the intensity process. To generate sample paths of (25), one can employ the so-called Euler scheme. It is assumed that the pricing agent knows the function $\bar{\lambda}$ and thinks of it as the demand intensity process.

We set $C = D = 1$, $k = 1$ and $\sigma = 1$, and let $c = 10$. Suppose the pricing agent uses a Rényi-type penalty to generate uncertainty sets. Thus, the pricing agent has everything s/he needs to construct a robust pricing model. As before, for each combination of $(\alpha, \theta)$, we can numerically solve the value function $V(t, n)$ as well as the pricing policy $p(t, n)$ for $n = 1, \ldots, 10$ in a recursive way (starting from $n = 1$).

Our first simulation uses the “actual” demand rate model described by (25). Specifically, from time zero to the first arrival epoch (denoted as $t_1$), the pricing function is $p(t, 10)$. After $t_1$ we plug the price function $p(t, 9)$ into the process (25) to generate the next arrival time $t_2$. Repeating this process, we can calculate the total revenue based on demands realized up to $T$. Our simulation program uses $r = 1000$ i.i.d. replications of the system observed over a fixed sales period of $T = 20$. For a random variable $X$, its mean $\mathbb{E}[X]$ is estimated by the sample averages of the $r$ values, which should be Gaussian. Hence, the 95% confidence interval can be constructed in the usual way with $t_{0.025}(r - 1)$. 

**FIGURE 3** Optimal pricing policy $p(t)$ under Rényi-type penalty with $\alpha = 2$, $\theta = 1$ and $50$
Figure 4 Expected revenues obtained from computer simulations with CIR processes under different Rényi-type penalties

In Figure 4 we draw the simulated average revenue with a different choice of $\theta$ under three different values of $\alpha$. It can be seen from the figure that the estimated average revenue first increases and then decreases in $\theta$ for all $\alpha$. It is worth pointing out that for very large $\theta$, the resultant pricing policy will be the one obtained by completely ignoring model uncertainty. Hence, the difference between the "limiting value" of each curve and the maximum value of that curve can be used to quantify the value of robustness. For example, when $\alpha = 2$, the value of robustness is roughly $6.155 - 5.801 = 0.354$, which is approximately 6.1% improvement relative to completely ignoring model uncertainty.

In the second example, nominal demand arrivals still follow a non-homogeneous Poisson process. However, the actual demand is dictated by one of the two demand functions, $\lambda_1(t)$ and $\lambda_2(t)$ at any given point in time. These two demand models serve as two “states” of a continuous-time Markov chain (CTMC). The demand function $\lambda_1(t)$ represents the optimistic scenario, whereas $\lambda_2(t)$ represents the pessimistic scenario. The weighted sum of $\lambda_1(t)$ and $\lambda_2(t)$ is our nominal demand function, $\tilde{\lambda}(t)$. To simulate the first arrival, we use the price function, $p(t, 10)$. However, whether the intensity $\lambda_1(p(t, 10))$ or $\lambda_2(p(t, 10))$ is used depends on the state of CTMC at time $t$.

Then, to identify which demand function should be used to simulate the second arrival, we must again track the state of CTMC at any given time prior to the second arrival. Repeating this process, we can calculate the total revenue up to time $T$.

Let $\tilde{\lambda}(t) = e^{-p(t)}$, $\lambda_1(t) = 3/2e^{-p(t)}$, and $\lambda_2(t) = 1/2e^{-p(t)}$. In the CTMC, let each state’s sojourn time have the same exponential distribution: $h_1(t) = h_2(t) = 50e^{-50t}$. We set the total selling period as $T = 20$ and the total number of products as $c = 10$, same as in the first example. In
Figure 5 we draw the simulated average revenue with different values of $\theta$. Again, we see that the estimated average revenue rises initially, then falls in $\theta$ for all $\alpha$. This time, the value of robustness for $\alpha = 2$ is approximately $6.032 - 5.578 = 0.454$, which is 8.1% improvement over fully ignoring model uncertainty.

The main insight of Figure 4 and Figure 5 is that using a robust optimization approach that accounts for uncertainty results in significant revenue increases. The decision-maker who adopts the nominal price strategy will be penalized since he or she did not adequately prepare for randomness. Constantly contemplating the worst-case scenario, on the other hand, will result in a loss of profit as well. The robust methodology can help to achieve a trade-off between these two extremes.

references


Appendix A: Proof Sketch of (20)

To quantify the distance between \( \mathbb{P} \) and \( \mathbb{P}^\psi \), consider \( \psi \) fixed, and divide the entire time period \( T \) into \( m \) small sections, where \( m \) is sufficiently large. We have that

\[
\mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \int_0^T \ln \psi(u) \, dN(u)} \right] = \mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \left( \int_0^1 \ln \psi(u) \, dN(u) + \int_{1}^2 \ln \psi(u) \, dN(u) + \ldots + \int_{m-1}^m \ln \psi(u) \, dN(u) \right)} \right] \\
= \mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \int_0^1 \ln \psi(u) \, dN(u)} \cdot e^{(\alpha-1) \int_1^2 \ln \psi(u) \, dN(u)} \cdot \ldots \cdot e^{(\alpha-1) \int_{m-1}^m \ln \psi(u) \, dN(u)} \right] \\
= \mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \int_0^1 \ln \psi(u) \, dN(u)} \right] \cdot \ldots \cdot \mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \int_{m-1}^m \ln \psi(u) \, dN(u)} \right].
\]

The third equality is due to independent increments of \( N \). Then, for every small time period, we consider piece-wise constant approximation of a non-homogeneous Poisson process. In particular, on the small time interval \([t_{i-1}, t_i] \), we have that

\[
\mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \int_{t_{i-1}}^t \ln \psi(u) \, dN(u)} \right] \approx \mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \ln \psi(t_i) \Delta N(t_i)} \right] = \exp \left\{ \psi(t_i) \lambda(t_i) (e^{(\alpha-1) \ln \psi(t_i)} - 1) \Delta t \right\},
\]

where the second inequality comes from the moment generating function for a Poisson random variable. In view of the above calculations, we see that \( \mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \int_0^T \ln \psi(u) \, dN(u)} \right] \) can be approximated by

\[
\exp \left\{ \sum_i \psi(t_i) \alpha \lambda(t_i) \Delta t - \sum_i \psi(t_i) \lambda(t_i) \Delta t \right\},
\]

which tends to

\[
\exp \left\{ \int_0^T \psi(u)^\alpha \lambda(u) \, du - \int_0^T \psi(u) \lambda(u) \, du \right\} \quad \text{as} \quad \Delta t \to 0.
\]

Using the definition of Rényi divergence, we have that

\[
\mathcal{R}_\alpha \left( \mathbb{P}^\psi || \mathbb{P} \right) = \frac{1}{\alpha - 1} \ln \mathbb{E}^{\mathbb{P}^\psi} \left[ e^{(\alpha-1) \int_0^T \ln \psi(u) \, dN(u)} \cdot e^{(\alpha-1) \int_0^T (1-\psi(u)) \lambda(u) \, du} \right] \\
= \frac{1}{\alpha - 1} \ln \exp \left\{ \int_0^T \psi(u)^\alpha \lambda(u) \, du - \int_0^T \psi(u) \lambda(u) \, du + (\alpha - 1) \int_0^T (1-\psi(u)) \lambda(u) \, du \right\} \\
= \frac{1}{\alpha - 1} \int_0^T \lambda(\rho(u)) \left( \psi(u)^\alpha - \alpha \psi(u) + \alpha - 1 \right) \, du,
\]
Appendix B: Numerical Algorithm for Solving $V(t, n)$

To find the solution of the optimality equation (21), we start with an initial guess of $V(t, 1)$, denoted as $V^0(t, 1)$, that solves

$$V^0_t(t, 1) = \theta \lambda(\bar{p}) \left( \frac{1}{\left( \frac{\xi}{\bar{p}} V^0(t, 1) + 1 - \frac{\xi \bar{p}}{\bar{p}} \right)^\frac{1}{\tau}} - 1 \right)$$

subject to the boundary condition $V^0(T, 1) = 0$. Notice that (26) is a first-order ordinary differential equation which we can solve numerically via finite difference method (FDM). Then for the value function $V^0(t, 1)$ we seek a pricing policy $p^0(t)$ that maximizes

$$\lambda(p^0(t)) \left( \frac{1}{\left( \frac{\xi}{\bar{p}} V^0(t, 1) + 1 - \frac{\xi \bar{p}}{\bar{p}} \right)^\frac{1}{\tau}} - \lambda(p^0(t)) \right).$$

The next step is to find $V^1(t, 1)$ such that

$$V^1_t(t, 1) = \theta \lambda(p^0(t)) \left( \frac{1}{\left( \frac{\xi}{\bar{p}} V^1(t, 1) + 1 - \frac{\xi \bar{p}}{\bar{p}} \right)^\frac{1}{\tau}} - 1 \right)$$

subject to the same boundary condition as mentioned previously. We can solve (27) by FDM again.

In general, for $n = 1$, using the $k$th estimate of $V(t, 1)$, denoted as $V^k(t, 1)$, we can find $p^k(t)$ that maximizes

$$\lambda(p^k(t)) \left( \frac{1}{\left( \frac{\xi}{\bar{p}} V^k(t, 1) + 1 - \frac{\xi \bar{p}}{\bar{p}} \right)^\frac{1}{\tau}} - \lambda(p^k(t)) \right),$$

and further solve the ordinary differential equation

$$V^{k+1}_t(t, 1) = \theta \lambda(p^k(t)) \left( \frac{1}{\left( \frac{\xi}{\bar{p}} V^{k+1}(t, 1) + 1 - \frac{\xi \bar{p}}{\bar{p}} \right)^\frac{1}{\tau}} - 1 \right)$$

subject to the boundary condition $V^{k+1}(T, 1) = 0$ by using FDM to get $V^{k+1}(t, 1)$, the $(k + 1)$th estimate of $V(t, 1)$. Repeating these steps we obtain an iterative procedure that generates a sequence $\{V^{k+1}(t, 1), p^k\}$ which is expected to converge to the optimal solution when $k \to \infty$. Although we do not attempt to rigorously prove the desired convergence result, our extensive numerical experiments suggest convergence happens after a few iterations. The algorithm terminates when the iteration error becomes sufficiently small. Denote the convergent value function as $V(t, 1)$.

Having obtained a numerical solution to $V(t, 1)$, we can compute $V(t, n)$ for $n \geq 2$ in a recursive
fashion. Take \( n = 2 \) for example, we can employ the same strategy to solve the equation

\[
V_{t}^{k+1}(t, 2) = \theta \lambda (\rho^k(t)) \left( \frac{1}{(\frac{1}{2}V^{k+1}(t, 2) - \frac{\zeta}{2}V(t, 1) + 1 - \frac{\zeta \rho^k(t)}{2})^{\frac{1}{2}}} - 1 \right)
\]

subject to the boundary condition \( V^{k+1}(T, 2) = 0 \) by FDM. Repeating the steps we obtain an iterative procedure that generates a sequence \( \{ V^{k+1}(t, 2), \rho^k \} \) which is expected to converge to the optimal solution when \( k \to \infty \). It is worth mentioning here the price policy \( \rho(t) \) when \( n = 2 \) is different from price policy when \( n = 1 \), however we can still start with \( \rho^0(t) = \bar{\rho} \) for all \( t \). The algorithm terminates when the iteration error becomes sufficiently small. This way, we can compute \( V(t, n) \) for any \( n \) we desire.

Appendix C: Additional Numerical Results

Figure 6, 7 and 8 provide the value functions mentioned in §5. We observe that with the growth of \( \alpha \), the gap between value functions for fixed \( n \) under different \( \theta \) decreases. This result is consistent with our pricing strategies. Furthermore, table 1 and 2 provide the confidence intervals (obtained from computer simulations) of all the point estimates shown in Figure 4 and 5 in the main paper.

**TABLE 1** Estimated average revenues obtained from computer simulations for CIR process

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( \alpha = 1/2 )</th>
<th>( \alpha = 1 )</th>
<th>( \alpha = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5.8501±1.4E-1</td>
<td>5.9263±1.3E-1</td>
<td>5.9640±1.4E-1</td>
</tr>
<tr>
<td>5</td>
<td>5.9105±1.4E-1</td>
<td>5.9641±1.4E-1</td>
<td>6.0060±1.4E-1</td>
</tr>
<tr>
<td>10</td>
<td>5.9304±1.5E-1</td>
<td>6.0231±1.4E-1</td>
<td>6.1552±1.4E-1</td>
</tr>
<tr>
<td>20</td>
<td>6.0264±1.4E-1</td>
<td>6.1202±1.4E-1</td>
<td>6.0279±1.4E-1</td>
</tr>
<tr>
<td>30</td>
<td>6.1128±1.5E-1</td>
<td>5.9570±1.4E-1</td>
<td>5.9587±1.4E-1</td>
</tr>
<tr>
<td>50</td>
<td>6.0493±1.5E-1</td>
<td>5.8710±1.5E-1</td>
<td>5.8846±1.5E-1</td>
</tr>
<tr>
<td>200</td>
<td>5.8279±1.5E-1</td>
<td>5.8117±1.5E-1</td>
<td>5.8012±1.4E-1</td>
</tr>
</tbody>
</table>

**TABLE 2** Estimated average revenues obtained from computer simulations for CTMC

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( \alpha = 1/2 )</th>
<th>( \alpha = 1 )</th>
<th>( \alpha = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5.4376±1.3E-1</td>
<td>5.5089±1.3E-1</td>
<td>5.6060±1.3E-1</td>
</tr>
<tr>
<td>5</td>
<td>5.6705±1.3E-1</td>
<td>5.6965±1.4E-1</td>
<td>5.7899±1.4E-1</td>
</tr>
<tr>
<td>10</td>
<td>5.7678±1.4E-1</td>
<td>5.7854±1.3E-1</td>
<td>5.8947±1.3E-1</td>
</tr>
<tr>
<td>20</td>
<td>5.8850±1.4E-1</td>
<td>5.9173±1.4E-1</td>
<td>6.0315±1.4E-1</td>
</tr>
<tr>
<td>30</td>
<td>5.9132±1.4E-1</td>
<td>5.9609±1.4E-1</td>
<td>5.8421±1.5E-1</td>
</tr>
<tr>
<td>50</td>
<td>5.8177±1.4E-1</td>
<td>5.8439±1.4E-1</td>
<td>5.8070±1.4E-1</td>
</tr>
<tr>
<td>200</td>
<td>5.5903±1.5E-1</td>
<td>5.6172±1.4E-1</td>
<td>5.5781±1.4E-1</td>
</tr>
</tbody>
</table>
FIGURE 6  Value function under Rényi-type penalty with $\alpha = 1/2, \theta = 1$ and 50

FIGURE 7  Value function under Rényi-type penalty with $\alpha = 1, \theta = 1$ and 50

FIGURE 8  Value function under Rényi-type penalty with $\alpha = 2, \theta = 1$ and 50