

On the optimality of threshold control in queues with model uncertainty

A. Jain · A.E.B. Lim · J.G. Shanthikumar

Received: 19 November 2009 / Revised: 8 March 2010 / Published online: 15 April 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract We consider a single-stage queuing system where arrivals and departures are modeled by point processes with stochastic intensities. An arrival incurs a cost, while a departure earns a revenue. The objective is to maximize the profit by controlling the intensities subject to capacity limits and holding costs. When the stochastic model for arrival and departure processes are completely known, then a threshold policy is known to be optimal. Many times arrival and departure processes can not be accurately modeled and controlled due to lack of sufficient calibration data or inaccurate assumptions. We prove that a threshold policy is optimal under a max–min robust model when the uncertainty in the processes is characterized by relative entropy. Our model generalizes the standard notion of relative entropy to account for different levels of model uncertainty in arrival and departure processes. We also study the impact of uncertainty levels on the optimal threshold control.

Keywords Threshold control · Relative entropy · Robust optimization · Intensity control

This work is supported in part by an NSF *CAREER* Award DMI-0348746 (Lim) and the NSF Grant DMI-0500503 (Lim and Shanthikumar). Nevertheless, the opinions, findings, conclusions and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Support from the Coleman Fung Chair in Financial Modelling (Lim) is also acknowledged.

A. Jain · A.E.B. Lim (✉)

IEOR Department, University of California, Berkeley, 4174 Etcheverry Hall, Berkeley, CA 94720, USA

e-mail: lim@ieor.berkeley.edu

A. Jain

e-mail: ankit@ieor.berkeley.edu

J.G. Shanthikumar

Krannert School of Management, Purdue University, West Lafayette, IN 47906, USA

e-mail: shanthikumar@purdue.edu

Mathematics Subject Classification (2000) 60K25**1 Introduction**

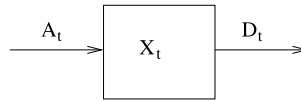
We consider a general single-stage queuing system, in which the input (arrival) and output (service completion) processes are modeled by point processes with dynamically controlled stochastic intensities. An entering job incurs a cost, \tilde{c} , and a job completion produces revenue, \tilde{p} . In addition there is a holding cost which is linearly proportional to the number of jobs in the system at a given time. The problem is to dynamically control both the input and output intensities so as to maximize discounted profit.

Problems of this type have been studied, for example, by Chen and Yao [3], where it is shown that a threshold policy for both the input and output processes is optimal under the assumption that the stochastic model for arrival and departure processes is accurate and known. (See [7, 14] and [12] for similar results.) In many applications, however, arrival and departure intensities can not be accurately modeled due to complexities of the real-world system or lack of sufficient calibration data. This raises natural questions including (i) what is the impact of model uncertainty on the “optimal” operating policies for the system, and (ii) are threshold policies still “optimal”? In this paper we account for model errors by formulating a max-min robust control version of this problem in which model uncertainty is incorporated using the notion of relative entropy. Within this framework we show that threshold policy is optimal for the robust control problem, and study the impact of the level of model uncertainty on the optimal threshold level.

While the use of relative entropy to account for model uncertainty in stochastic optimization problems has a relatively long history ([5, 8, 10] and [11]), one feature of this paper which departs from the standard approach is that we generalize the standard notion of relative entropy in order to allow for different levels of model uncertainty for the arrival as well as the departure processes (see also Lim, Shanthikumar and Watwai [9] for similar ideas in the context of dynamic pricing). Aside from being realistic—for example, it is likely to be the case that the system operator is substantially more knowledgeable about the service system he/she is controlling (since it is internal) than the customer arrival process, which is typically much more complicated and subject to many external factors—this also allows us to study (say) the impact of the level of model uncertainty in the arrival process on the service control policy.

The outline of this paper is as follows. In Sect. 2 we recall the model from Chen and Yao [3] and formulate the robust version of this problem. The robust version involves an extension of the notion of discounted relative entropy from Hansen, Sargent, Turmuhambetova and Williams [5] in order to handle different levels of model uncertainty for the arrival and departure processes. Dynamic programming equations for the robust control problem are derived in Sect. 3, and the impact of the level of model uncertainty on the threshold control levels is studied in Sect. 4.

Fig. 1 Queuing system



2 Model formulation

In this section we first introduce the standard model which is similar to [3] before formulating the robust model in Sect. 2.2. The robust model extends the notion of discounted relative entropy from [5] in order to handle different level of uncertainties in arrival and departure rates.

2.1 Nominal model

Consider a single-stage queuing system as shown in Fig. 1. Let X_t be the state of the system that denotes the number of jobs in process at time t . X_t takes values on nonnegative integers and is of the form

$$X_t = x_0 + A_t - D_t, \tag{1}$$

where $x_0 \geq 0$ is the state at time $t = 0$ of the system and A_t and D_t are the arrival and departure processes respectively. A_t and D_t denote the cumulative number of arrivals and departures until time t .

We assume that A_t and D_t are simple point processes. Let \mathcal{F}_t be the sigma field generated by X_t , i.e., $\mathcal{F}_t = \sigma(X_s, s \leq t)$. Also let A_t and D_t admit \mathcal{F}_t predictable intensities β_t and α_t . The rates α_t and β_t are subjected to the following capacity constraints:

$$\begin{aligned} 0 \leq \beta_t \leq y, \quad \forall t \geq 0, \\ 0 \leq \alpha_t \leq z, \quad \forall t \geq 0. \end{aligned} \tag{2}$$

If there is no ambiguity in the arrival or departure process, i.e., if we can exactly control the arrival and departure intensities, then our objective is to find a control $u = \{\beta_t, \alpha_t, t \geq 0\}$ to maximize the following discounted value function:

$$V(x_0, u) = \mathbb{E}_{x_0} \int_0^\infty e^{-\delta t} (\tilde{p} dD_t - \tilde{c} dA_t - hX_t dt). \tag{3}$$

\mathbb{E}_{x_0} denotes the conditional expectation given $X_0 = x_0$, δ is the discount factor, \tilde{p} is the revenue obtained by selling one unit of output, \tilde{c} is the cost of acquiring one unit of input and h is the unit holding cost for work-in-process inventory. Substituting X_t from (1) in (3) we get:

$$V(x_0, u) = \mathbb{E}_{x_0} \int_0^\infty e^{-\delta t} \left(\left(\tilde{p} + \frac{h}{\delta} \right) dD_t - \left(\tilde{c} + \frac{h}{\delta} \right) dA_t \right) - \frac{hx_0}{\delta}. \tag{4}$$

Defining $p = \tilde{p} + \frac{h}{\delta}$ and $c = \tilde{c} + \frac{h}{\delta}$ we have

$$V(x_0, u) = \mathbb{E}_{x_0} \int_0^\infty e^{-\delta t} (p dD_t - c dA_t) - \frac{hx_0}{\delta}. \tag{5}$$

We can drop the last term in (5) for the purpose of finding optimal control as it is a constant term. From the definition of stochastic intensity [2]

$$\begin{aligned} \mathbb{E}_{x_0} \int_0^\infty c dA_t &= \mathbb{E}_{x_0} \int_0^\infty c\beta_t dt, \\ \mathbb{E}_{x_0} \int_0^\infty p dD_t &= \mathbb{E}_{x_0} \int_0^\infty p\alpha_t dt. \end{aligned} \tag{6}$$

Rewriting the value function in (5) using (6) and dropping the constant term we have

$$V(x_0, u) = \mathbb{E}_{x_0} \int_0^\infty e^{-\delta t} (p dD_t - c dA_t) = \mathbb{E}_{x_0} \int_0^\infty e^{-\delta t} (p\alpha_t - c\beta_t) dt. \tag{7}$$

The problem formulation with unambiguous arrival rate is:

$$\max_u V(x_0, u) = \max_u \mathbb{E}_{x_0} \int_0^\infty e^{-\delta t} (p\alpha_t - c\beta_t) dt. \tag{8}$$

2.2 Robust model

Let $(\Omega, \mathcal{F}_t, \mathcal{F})$ be the underlying measurable space for arrival and departure processes, A_t and D_t respectively. A_t and D_t are counting processes and admit intensities. A complete specification of intensity λ_t of the process A_t and of intensity μ_t of the process D_t induces a measure P over \mathcal{F} . The nominal model is based on the assumption that the decision maker is able to set arrival and departure intensities precisely subject to capacity constraints. The objective then is to find (λ_t, μ_t) which are optimal.

In reality the real-world intensity processes are unlikely to be (λ_t, μ_t) . For example, the arrival rate, λ_t , might be a function of the price an arriving customer pays for the service being offered while μ_t could depend on the number of workers assigned to the customer in service, and the assumption in the nominal model is the decision maker knows the exact relationship between pricing decisions and the arrival rate λ_t , as well as the number of workers assigned and the departure rate μ_t , so that the arrival and departure rates can be set to the precise values that the decision maker desires. In practice, the relationship between the pricing decision and λ_t and also the number of assigned workers and the service rate μ_t may be difficult to characterize. The arrival intensity might be a complicated non-stationary function of the price and also of other factors such as amount of advertising. This makes it impossible to precisely calibrate intensities.

More generally we have a situation where the decision maker on the basis of her model thinks she is setting the arrival and departure rates at levels (λ_t, μ_t) but in reality the rates might be something different (say (β_t, α_t)). Our objective in this section is to incorporate the possibility of such model uncertainty into the formulation of the problem.

Suppose the real-world \mathcal{F}_t -predictable intensity processes β_t and α_t induces a measure Q over \mathcal{F} . We assume that the real-world intensity processes, while not known accurately, satisfy certain minimal conditions with respect to the intensity

processes λ_t and μ_t , which are precisely known to the decision maker. Let P_t and Q_t be restrictions of P and Q respectively to \mathcal{F}_t . In particular we assume that for all t , Q_t is absolutely continuous with respect to P_t , i.e.,

$$P_t(A) = 0 \Rightarrow Q_t(A) = 0 \quad \forall A \in \mathcal{F}_t.$$

The distribution Q is said to be absolutely continuous over finite intervals with respect to P if Q_t is absolutely continuous with respect to P_t for all t . This definition of absolute continuity captures the idea that two models are impossible to distinguish with certainty over a finite interval [5].

Let $\{\gamma_t, t \geq 0\}$ be a stochastic process such that for every t , γ_t is Radon–Nikodym derivative [4] of Q_t with respect to P_t . γ_t is a positive martingale and is adapted to filtration \mathcal{F}_t . It follows from [6] that there are F_t -predictable processes κ_t and η_t such that:

$$\gamma_t = \exp\left(\int_0^t (\ln(\kappa_s) dA_s + \ln(\eta_s) dD_s) + \int_0^t ((1 - \kappa_s)\lambda_s + (1 - \eta_s)\mu_s) ds\right). \tag{9}$$

The following result is a version of the Girsanov Theorem for point processes as stated in Bremaud [2].

Theorem 1 (Girsanov Theorem) *Let A_t and D_t be \mathcal{F}_t -adapted point processes with F_t -predictable intensities λ_t and μ_t respectively under the probability measure P . Suppose that γ_t is a positive F_t -martingale under P and that the Radon–Nikodym density of Q_t with respect to P_t is given by*

$$\frac{dQ_t}{dP_t} = \gamma_t = \exp\left(\int_0^t (\ln(\kappa_s) dA_s + \ln(\eta_s) dD_s) + \int_0^t ((1 - \kappa_s)\lambda_s + (1 - \eta_s)\mu_s) ds\right), \tag{10}$$

then A_t and D_t are \mathcal{F}_t -adapted point processes with intensities $\beta_t = \kappa_t\lambda_t$ and $\alpha_t = \eta_t\mu_t$ respectively under Q .

Theorem 1 allows us to parameterize the real-world model $Q = (\beta_t, \alpha_t, t \geq 0)$ through the processes κ_t and η_t .

2.3 Relative entropy

Relative entropy or KL divergence is a measure of difference between two probability measures. In this paper we use a weaker notion, called *Discounted Relative Entropy* [5] to measure the discrepancy between two measures over an infinite horizon.

The weaker notion requires that the two measure being compared put positive probability on all of the same events, except tail events. The discounted relative entropy is defined as:

$$\tilde{\mathbb{R}}(Q|P) = \delta \int_0^\infty \exp(-\delta t) \left(\int \ln\left(\frac{dQ_t}{dP_t}\right) dQ_t \right) dt, \tag{11}$$

where $\frac{dQ_t}{dP_t}$ is the Radon–Nikodym derivative of Q_t with respect to P_t .

This measure of relative entropy is convex in Q as shown in [5]. It should be noted here that even if the discounted measure of entropy is finite the standard relative entropy measure of distance between P and Q can be infinite, i.e., it can happen that:

$$\int \log\left(\frac{dQ}{dP}\right)dQ = +\infty. \tag{12}$$

If (12) holds but the discounted relative entropy (11) is finite, then it means that a statistician would be able to distinguish between the probability measures P and Q with a continuous record of data on an infinite interval while it is impossible to do so by recording finite length time interval data. As an example, if under P the arrival rate is constant λ and under Q the arrival rate is constant β , $\beta \neq \lambda$, then the relative entropy of P and Q is infinite but the discounted relative entropy between Q and P is finite.

Returning to our discussion on point processes, it follows from Theorem 1 that our measure of discounted relative entropy (11) transforms into:

$$\begin{aligned} \tilde{\mathbb{R}}(Q|P) &= \delta \int_0^\infty e^{-\delta t} \left(\int \ln \frac{dQ_t}{dP_t} dQ_t \right) dt \\ &= \delta \int_0^\infty e^{-\delta t} \left(\int_0^t (\lambda_s(\kappa_s \ln \kappa_s + 1 - \kappa_s) + \mu_s(\eta_s \ln \eta_s + 1 - \eta_s)) ds \right) dt \\ &= \delta \int_0^\infty (\lambda_s(\kappa_s \ln \kappa_s + 1 - \kappa_s) + \mu_s(\eta_s \ln \eta_s + 1 - \eta_s)) ds \left(\int_s^\infty e^{-\delta t} dt \right) \\ &= \int_0^\infty e^{-\delta s} \lambda_s(\kappa_s \ln \kappa_s + 1 - \kappa_s) ds + \int_0^\infty e^{-\delta s} \mu_s(\eta_s \ln \eta_s + 1 - \eta_s) ds, \end{aligned} \tag{13}$$

where the third equality is justified by Fubini’s theorem [4] as the integrand is positive. The first term $\tilde{\mathbb{R}}_1(Q|P) = \int_0^\infty e^{-\delta s} \lambda_s(\kappa_s \ln \kappa_s + 1 - \kappa_s) ds$ can be interpreted as a measure of ambiguity in the arrival process. Similarly the second term $\tilde{\mathbb{R}}_2(Q|P) = \int_0^\infty e^{-\delta s} \mu_s(\eta_s \ln \eta_s + 1 - \eta_s) ds$ measures the ambiguity in the departure process.

Our robust control problem corresponding to (8) is as follows:

$$\begin{aligned} \max_{u \in U} \min_Q \mathbb{E}_Q \left[\int_0^\infty e^{-\delta t} (p\alpha_t - c\beta_t) dt \right] \\ \text{Subject to: } \tilde{\mathbb{R}}(Q|P) \leq \eta. \end{aligned} \tag{14}$$

Here the control is $u = \{\lambda_t, \mu_t, \lambda_t \leq y, \mu_t \leq z, t \geq 0\}$.

The robust control problem is a two-player game between ‘nature’ and decision maker. Given the control u , nature chooses a “worst-case” measure Q from the class of measures defined by the convex discounted relative entropy constraint. The constant $\eta \geq 0$ is a measure of our confidence in the nominal measure P and restricts the amount that Q (or the real-world intensity processes β_t and α_t) can deviate from P (resp. λ_t and μ_t). A large value of η allows Q to deviate further from our nominal probability measure P while a small value of η is chosen when we have a high

degree of confidence in our nominal model. Putting $\eta = 0$ reduces the robust control problem to a standard one.

Alternatively, we may consider the following problem:

$$\max_{u \in U} \min_Q \left(\mathbb{E}_Q \left[\int_0^\infty e^{-\delta t} (p\mu_t - c\beta_t dt) \right] + \theta \tilde{\mathbb{R}}(Q|P) \right). \tag{15}$$

The constant $\theta > 0$ may be seen as the Lagrange multiplier for the relative entropy constraint in (14) and solving (14) is equivalent to solving (15) for an appropriate choice of θ . Alternatively, the parameter θ can represent our confidence in the nominal model. A large value of θ denotes high confidence in the model as the penalty of deviation from the model is large.

Note that the discounted relative entropy in (13) is the sum of two terms. The terms individually can be interpreted as measure of uncertainties in arrival and departure processes respectively. In formulation (15), as both the terms in discounted relative entropy expansion are multiplied by the same constant θ , the confidence levels in arrival and departure processes are assumed to be the same. If we have reason to believe in varying levels of confidence in arrival and departure processes the formulation (15) can be modified as:

$$\max_{u \in U} \min_Q \left(\mathbb{E}_Q \left[\int_0^\infty e^{-\delta t} (p\alpha_t - c\beta_t dt) \right] + \theta_A \tilde{\mathbb{R}}_1(Q|P) + \theta_D \tilde{\mathbb{R}}_2(Q|P) \right), \tag{16}$$

where θ_A and θ_D denotes the confidence in arrival and departure processes respectively. Hence Model (16) differs from the standard robust model (15) which assumes the same level of uncertainty for all parts of the model.

Substituting the value of discounted relative entropy for point processes from (13) to (15), our robust formulation is:

$$\max_{u \in U} \min_{\kappa, \eta} \mathbb{E}_{x_0} \left[\int_0^\infty e^{-\delta t} (p\eta_s \mu_s - c\kappa_s \lambda_s + \theta_A \lambda_s (1 - \kappa_s + \kappa_s \ln \kappa_s) + \theta_D \mu_s (1 - \eta_s + \eta_s \ln \eta_s)) ds \right]. \tag{17}$$

3 Characterization of optimal policy

Suppose we first restrict ourselves to the policies which are Markov in the state (the number of items that are currently in service). In other words, we can replace λ_t and μ_t by $\lambda(X_t)$ and $\mu(X_t)$ respectively. Further assume that nature is restricted to choose among a set of Markovian policy only, i.e., κ and η are only functions of X . In this case the formulation (17) reduces to:

$$\max_{\lambda, \mu} \min_{\kappa} \mathbb{E}_{x_0} \left[\int_0^\infty e^{-\delta t} (p\eta(X_s)\mu(X_s) - c\kappa(X_s)\lambda(X_s) + \theta_A \lambda(X_s)(1 - \kappa(X_s) + \kappa(X_s) \ln \kappa(X_s)) + \theta_D \mu(X_s)(1 - \eta(X_s) + \eta(X_s) \ln \eta(X_s))) ds \right]. \tag{18}$$

The Hamiltonian–Jacobi–Bellman (HJB) equation corresponding to the above formulation is:

$$\delta V(x) = \max_{\lambda(x), \mu(x)} \min_{\kappa(x), \eta(x)} \left[\lambda(x)\kappa(x)(-c + \theta_A(-1 + \ln \kappa(x)) + \Delta V(x)) + \lambda(x)\theta_A + \mu(x)\eta(x)(p + \theta_D(-1 + \ln \eta(x)) - \Delta V(x - 1)) + \mu(x)\theta_D \right], \tag{19}$$

where

$$\Delta V(x) = V(x + 1) - V(x) \tag{20}$$

and

$$V(x) = \max_{\lambda(x), \mu(x)} \min_{\kappa(x), \eta(x)} \mathbb{E}_x \left[\int_0^\infty e^{-\delta t} (p\mu(X_s) - c\kappa(X_s)\lambda(X_s) + \theta_A\lambda(X_s)(1 - \kappa(X_s) + \kappa(X_s)\ln \kappa(X_s)) + \theta_D\mu(X_s)(1 - \eta(X_s) + \eta(X_s)\ln \eta(X_s))) ds \right]. \tag{21}$$

The solution of the (unconstrained convex) inner minimization (with respect to κ and η) problem in (19) is characterized by the first order conditions and yields the following:

$$\begin{aligned} \kappa^*(x) &= \exp\left(-\frac{1}{\theta_A}(\Delta V(x) - c)\right), \\ \eta^*(x) &= \exp\left(-\frac{1}{\theta_D}(p - \Delta V(x - 1))\right). \end{aligned} \tag{22}$$

Substituting back the value of κ^* and $\eta^*(x)$ from (22) to (19) we get the following after some manipulation:

$$\begin{aligned} \delta V(x) &= \max_{\lambda(x), \mu(x)} \left[\theta_A\lambda(x) \left(1 - \exp\left(-\frac{1}{\theta_A}(\Delta V(x) - c)\right) \right) \right. \\ &\quad \left. + \theta_D\mu(x) \left(1 - \exp\left(-\frac{1}{\theta_D}(p - \Delta V(x - 1))\right) \right) \right]. \end{aligned} \tag{23}$$

As the above equation is linear in $\lambda(x)$ and $\mu(x)$ we get the following characterization of the optimal policy:

$$\lambda^*(x) = \begin{cases} y & \text{if } \Delta V(x) \geq c \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

$$\mu^*(x) = \begin{cases} z & \text{if } \Delta V(x - 1) \leq p, \quad x \geq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

This proves that the optimal policy would either allow arrivals at full force or not to allow arrivals at all. The same structure holds for production. We either produce

at full force or do not produce at all. In order to guarantee that the optimal policy is threshold we need to prove the existence of a number b such that:

$$\begin{cases} c \leq \Delta V(x) \leq p & \text{for } x \leq b \\ \Delta V(x) < c & \text{for } x > b. \end{cases} \tag{26}$$

As it does not make sense to stop the production if there is a positive inventory due to discounting and the holding cost, it is obvious that the optimal output policy should be of the following form:

$$\mu^*(x) = \begin{cases} z & \text{if } x \geq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{27}$$

Now consider the following policy for arrivals. At every stage there is a choice between setting the arrival intensity to zero or setting it equal to its maximum value of y . The value function if we follow this *binary* policy, which we have already proved to be optimal in the case when both nature and decision maker are restricted to the class of Markovian policies, is:

$$\begin{aligned} V(x) = \max_{\lambda(\cdot) \in \{0, y\}} \mathbb{E}_x & \left[\int_0^\infty e^{-\delta t} (p\mu^*(X_s)\eta^*(X_s) - c\kappa^*(X_s)\lambda(X_s) \right. \\ & + \theta_A\lambda(X_s)(1 - \kappa^*(X_s) + \kappa^*(X_s) \ln \kappa^*(X_s)) \\ & \left. + \theta_D\mu(X_s)(1 - \eta^*(X_s) + \eta^*(X_s) \ln \eta^*(X_s)) \right) ds \Big]. \end{aligned} \tag{28}$$

$\mu^*(x)$ is as described in (27), and $\kappa^*(x), \eta^*(x)$ are as in (22). Now suppose we can find a finite constant v such that $v \geq (y\kappa^*(x) + z\eta^*(x)), \forall x$. The existence of such a v is guaranteed if we look at the expression (22) as it is possible to get upper and lower bounds on $V(x)$.¹ Given such a v we can write the following dynamic programming equation (see Bertsekas [1, Chap. 5])

$$\begin{aligned} V(x) = \frac{1}{\delta + v} & [p\mu^*(x)\eta^*(x) + \mu^*(x)\theta_D(1 - \eta^*(x) + \eta^*(x) \ln \eta^*(x)) \\ & + (v - \mu^*(x))V(x) + \mu^*(x)V(x - 1) \\ & + \max(-cy\kappa^*(x) + y\theta_A(1 - \kappa^*(x) + \kappa^*(x) \ln \kappa^*(x)) \\ & + y\kappa^*(x)(V(x + 1) - V(x)), 0)]. \end{aligned} \tag{29}$$

Without loss of generality we can assume that $\delta + v = 1$ as it is possible to scale upper bounds z and y appropriately. Substituting the value of $\kappa^*(x)$ and $\eta^*(x)$ from (22) in (29) and simplifying we get:

$$\begin{aligned} V(x) = \mu^*(x)\theta_D & \left(1 - e^{-\frac{1}{\theta_D}(p - \Delta V(x-1))} \right) + vV(x) \\ & + y\theta_A \max \left(1 - e^{-\frac{1}{\theta_A}(\Delta V(x) - c)}, 0 \right). \end{aligned} \tag{30}$$

¹Zero is a lower bound. An upper bound is the value function of an unambiguous problem which can be uniformly upper bounded.

To prove the structural properties of $V(x)$ consider the following value-iteration algorithm:

$$\begin{aligned}
 V_{n+1}(x) = & \mu^*(x)\theta_D \left(1 - e^{-\frac{1}{\theta_D}(p - \Delta V_n(x-1))}\right) + vV_n(x) \\
 & + y\theta_A \max\left(1 - e^{-\frac{1}{\theta_A}(\Delta V_n(x)-c)}, 0\right). \tag{31}
 \end{aligned}$$

Such a value-iteration algorithm corresponding to a stochastic game can be shown to converge to the true value function (see [13]).

The set of value-iteration equations can be written more explicitly in the following form:

$$V_{n+1}(x) = z\theta_D \left(1 - e^{-\frac{1}{\theta_D}(p - \Delta V_n(x-1))}\right) + vV_n(x) + y\theta_A \max\left(1 - e^{-\frac{1}{\theta_A}(\Delta V_n(x)-c)}, 0\right) \tag{32}$$

where by convention $\Delta V_n(-1) = p$ for all n so that $e^{-\frac{1}{\theta_D}(p - \Delta V_n(-1))} = 1$.

Theorem 2 *Suppose we initialize $V_0(x) = 0$ for all x . If we iterate according to (32) then the following holds true for every n :*

- (a) $\Delta V_n(x) \leq p$.
- (b) $V_n(x)$ is increasing in x , i.e., $\Delta V_n(x) \geq 0$.
- (c) $V_n(x)$ is concave in x , i.e., $\Delta V_n(x)$ is decreasing in x .

Proof Proof is by induction. By construction the hypothesis holds true for $n = 0$. We now suppose that it holds for $n = k$ and show that it holds for $n = k + 1$.

(a)

$$\begin{aligned}
 \Delta V_{k+1}(x) = & z\theta_D \left(\left(1 - e^{-\frac{1}{\theta_D}(p - \Delta V_k(x))}\right) - \left(1 - e^{-\frac{1}{\theta_D}(p - \Delta V_k(x-1))}\right) \right) \\
 & + v(\Delta V_k(x)) + y\theta_A \left(\max\left(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x+1)-c)}, 0\right) \right. \\
 & \left. - \max\left(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x)-c)}, 0\right) \right) \\
 \leq & z\theta_D \left(1 - e^{-\frac{1}{\theta_D}(p - \Delta V_k(x))}\right) + v\Delta V_k(x) \\
 \leq & z(p - \Delta V_k(x)) + v\Delta V_k(x) = zp + (v - z)p \leq zp + (v - z)p = p.
 \end{aligned}$$

We have used the following facts: $\max(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x+1)-c)}, 0) \leq \max(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x)-c)}, 0)$ as $\Delta V_k(x)$ is a decreasing function of x ; $(1 - e^{-\frac{1}{\theta_D}(p - \Delta V_k(x-1))}) \geq 0$ as $\Delta V_k(x - 1) \leq p$ for all x and $1 - e^{-s} \leq s$ when $x \geq 0$.

(b)

$$\begin{aligned} \Delta V_{k+1}(x) &= z\theta_D \left(e^{-\frac{1}{\theta_D}(p-\Delta V_k(x-1))} - e^{-\frac{1}{\theta_D}(p-\Delta V_k(x))} \right) + v(\Delta V_k(x)) \\ &\quad + y\theta_A \left(\max \left(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x+1)-c)}, 0 \right) \right. \\ &\quad \left. - \max \left(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x)-c)}, 0 \right) \right) \\ &\geq v\Delta V_k(x) + y\theta_A \left(\max \left(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x+1)-c)}, 0 \right) \right. \\ &\quad \left. - \max \left(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x)-c)}, 0 \right) \right) \\ &\geq v\Delta V_k(x) - y\theta_A \max \left(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x)-c)}, 0 \right). \end{aligned}$$

If $\Delta V_k(x) \leq c$, then $\max(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x)-c)}, 0) = 0$ and hence $\Delta V_{k+1}(x) \geq v\Delta V_k(x) \geq 0$.

Else, if $\Delta V_k(x) \geq c$, then

$$\begin{aligned} \Delta V_{k+1}(x) &\geq v\Delta V_k(x) - y\theta_A \left(1 - e^{-\frac{1}{\theta_A}(\Delta V_k(x)-c)} \right) \\ &\geq v\Delta V_k(x) - y\theta_A \frac{\Delta V_k(x) - c}{\theta_A} \geq (v - y)\Delta V_k(x) \geq 0. \end{aligned}$$

Here we have used the fact that $1 - e^{-s} \leq s$ when $s \geq 0$. Also note that $v \geq y$ as $v \geq y\kappa^*(x)$ for every admissible value of $\kappa^*(x)$ and $\kappa^*(x) = 1$ is admissible as $\Delta V(x) = c$ is possible.

(c)

$$\begin{aligned} &\Delta V_{k+1}(x) - \Delta V_{k+1}(x + 1) \\ &= z\theta_D \left(e^{-\frac{f(x-1)}{\theta_D}} - 2e^{-\frac{f(x)}{\theta_D}} + e^{-\frac{f(x+1)}{\theta_D}} \right) + v(\Delta V_k(x) - \Delta V_k(x + 1)) \\ &\quad + y\theta_A \left(2\max \left(1 - e^{-\frac{g(x+1)}{\theta_A}}, 0 \right) - \max \left(1 - e^{-\frac{g(x)}{\theta_A}}, 0 \right) \right. \\ &\quad \left. - \max \left(1 - e^{-\frac{g(x+2)}{\theta_A}}, 0 \right) \right), \end{aligned}$$

where $f(x) = p - \Delta V(x)$ and $g(x) = \Delta V(x) - c$. Note that, as $0 \leq f(x) \leq p$ and $f(x)$ is increasing in x , we have

$$e^{-\frac{f(x-1)}{\theta_D}} - e^{-\frac{f(x)}{\theta_D}} \geq 0.$$

Also, as $g(x)$ is decreasing in x ,

$$\max \left(1 - e^{-\frac{g(x+1)}{\theta_A}}, 0 \right) - \max \left(1 - e^{-\frac{g(x+2)}{\theta_A}}, 0 \right) \geq 0.$$

Therefore

$$\begin{aligned} & \Delta V_{k+1}(x) - \Delta V_{k+1}(x + 1) \\ & \geq z\theta_D \left(e^{-\frac{f(x+1)}{\theta_D}} - e^{-\frac{f(x)}{\theta_D}} \right) + v(\Delta V_k(x) - \Delta V_k(x + 1)) \\ & \quad + y\theta_A \left(\max\left(1 - e^{-\frac{g(x+1)}{\theta_A}}, 0\right) - \max\left(1 - e^{-\frac{g(x)}{\theta_A}}, 0\right) \right) \\ & = z\theta_D \left(e^{-\frac{f(x+1)}{\theta_D}} - e^{-\frac{f(x)}{\theta_D}} \right) + z((p - \Delta V_k(x + 1)) - (p - \Delta V_k(x))) \\ & \quad + y\theta_A \left(\max\left(1 - e^{-\frac{g(x+1)}{\theta_A}}, 0\right) - \max\left(1 - e^{-\frac{g(x)}{\theta_A}}, 0\right) \right) \\ & \quad + y((\Delta V_k(x) - c) - (\Delta V_k(x + 1) - c)) \\ & \quad + (v - y - z)(\Delta V_k(x) - \Delta V_k(x + 1)) \\ & \geq z\theta_D \left(\left(e^{-\frac{f(x+1)}{\theta_D}} + \frac{f(x + 1)}{\theta_D} \right) - \left(e^{-\frac{f(x)}{\theta_D}} + \frac{f(x)}{\theta_D} \right) \right) \\ & \quad + y\theta_A \left(\left(\max\left(1 - e^{-\frac{g(x+1)}{\theta_A}}, 0\right) - \frac{g(x + 1)}{\theta_A} \right) \right. \\ & \quad \left. - \left(\max\left(1 - e^{-\frac{g(x)}{\theta_A}}, 0\right) - \frac{g(x)}{\theta_A} \right) \right). \end{aligned}$$

$e^{-s} + s$ is an increasing function of s when $s \geq 0$, so

$$\left(e^{-\frac{f(x+1)}{\theta_D}} + \frac{f(x + 1)}{\theta_D} \right) - \left(e^{-\frac{f(x)}{\theta_D}} + \frac{f(x)}{\theta_D} \right) \geq 0.$$

Also $\max(1 - e^{-s}, 0) - s$ is decreasing in s and, as $g(x)$ is a decreasing function of x ,

$$\left(\max\left(1 - e^{-\frac{g(x+1)}{\theta_A}}, 0\right) - \frac{g(x + 1)}{\theta_A} \right) - \left(\max\left(1 - e^{-\frac{g(x)}{\theta_A}}, 0\right) - \frac{g(x)}{\theta_A} \right) \geq 0.$$

Therefore,

$$\Delta V_{k+1}(x) - \Delta V_{k+1}(x + 1) \geq 0. \quad \square$$

Hence we have proved here that if we restrict ourselves to the class of Markovian policies and nature is also restricted to choose a Markovian policy to hurt the decision maker then a threshold policy is optimum. Specifically we proved that there exists a threshold $b \in [0, \infty]$ such that

$$\begin{cases} c \leq V(x + 1) - V(x) \leq p & \text{for } x \leq b \\ V(x + 1) - V(x) < c & \text{for } x > b. \end{cases} \tag{33}$$

Coupled with (24) we have the following policy:

$$\lambda^*(x) = \begin{cases} y & \text{if } x \leq b \\ 0 & \text{if } x > b. \end{cases} \tag{34}$$

Next we will show that the policy remains optimal even if nature is free to choose any non-Markovian policy. Specifically we prove that if we choose the threshold policy and nature is free to choose anything, nature would choose the Markovian policy to hurt most.

Theorem 3 *Suppose we choose the input and output intensities according to (34) and (25). Suppose we allow “nature”, acting as the adversary, to choose any arbitrary \mathcal{F}_t -predictable processes κ_t and η_t to hurt the decision maker so that the expected profit is minimized. Then nature would choose the Markovian policy as given by (22), where the value function in the equation is the optimal one when both nature and decision maker are allowed to choose only Markovian policies.*

Proof For any given arbitrary processes $(\kappa_t, \eta_t), t \geq 0$, suppose we consider a situation where nature follows (κ_t, η_t) up to time t and then follows the Markovian policy given by (22) after that. The value function associated with this (denoted by \tilde{V}_t) can be expressed as follows:

$$\begin{aligned} \tilde{V}_t(x) = \mathbb{E}_x \int_0^t e^{-\delta s} & (p\mu^*(X_s)\eta_s - c\lambda^*(X_s)\kappa_s + \theta_A\lambda^*(X_s)(1 - \kappa_s + \kappa_s \ln \kappa_s) \\ & + \theta_A\mu^*(X_s)(1 - \eta_s + \eta_s \ln \eta_s) ds) + \mathbb{E}_x [e^{-\delta t} V(X_t)]. \end{aligned} \tag{35}$$

To derive the second expectation in the above equation consider

$$\int_0^t e^{-\delta s} dV(X_s) = e^{-\delta t} V(X_t) - V(X(0)) + \delta \int_0^t e^{-\delta s} V(X_s) ds. \tag{36}$$

Taking expectation on both sides of the equality, we have

$$\mathbb{E}_x \int_0^t e^{-\delta s} dV(X_s) = \mathbb{E}_x [e^{-\delta t} V(X_t)] - V(x) + \mathbb{E}_x \left[\delta \int_0^t e^{-\delta s} V(X_s) ds \right]. \tag{37}$$

We can calculate the left most term in the above expression as:

$$\begin{aligned} \mathbb{E}_x \int_0^t e^{\delta s} dV(X_s) & = \mathbb{E}_x \int_0^t e^{-\delta s} ([\Delta V(X_s)]dA_s - [\Delta V(X_s - 1)]dD_s) \\ & = \mathbb{E}_x \int_0^t e^{-\delta s} (\kappa_s\lambda^*(X_s)\Delta V(X_s) - \eta_s\mu^*(X_s)\Delta V(X_s - 1)) ds. \end{aligned} \tag{38}$$

The first equality follows from the fact that there are only two possible transitions, upward and downward, and the second equality follows from (6). From (36) and (38) we get:

$$\mathbb{E}_x[e^{-\delta t} V(X_t)] = V(x) + \mathbb{E}_x \int_0^t e^{-\delta s} (\kappa_s \lambda^*(X_s) \Delta V(X_s) - \eta_s \mu^*(X_s) \Delta V(X_s - 1) - \delta V(X_s)) ds. \tag{39}$$

From (23) we have

$$\delta V(X_s) = \theta_A \lambda^*(x) \left(1 - \exp\left(-\frac{1}{\theta_A} (\Delta V(X_s) - c)\right) \right) + \theta_D \mu^*(x) \left(1 - \exp\left(-\frac{1}{\theta_D} (p - \Delta V(X_s - 1))\right) \right). \tag{40}$$

From (39) and (40) we get

$$\begin{aligned} \mathbb{E}_x[e^{-\delta t} V(X_t)] &= V(x) + \mathbb{E}_x \int_0^t e^{-\delta s} (\kappa_s \lambda^*(X_s) \Delta V(X_s) - \eta_s \mu^*(X_s) \Delta V(X_s - 1)) ds \\ &\quad - \mathbb{E}_x \int_0^t e^{-\delta s} \theta_A \lambda^*(x) \left(1 - \exp\left(-\frac{1}{\theta_A} (\Delta V(X_s) - c)\right) \right) ds \\ &\quad + \mathbb{E}_x \int_0^t e^{-\delta s} \theta_D \mu^*(x) \left(1 - \exp\left(-\frac{1}{\theta_D} (p - \Delta V(X_s - 1))\right) \right) ds. \end{aligned} \tag{41}$$

Substituting $\mathbb{E}_x[e^{-\delta t} V(X(t))]$ from (41) to (35) we obtain

$$\begin{aligned} \tilde{V}_t(x) &= \mathbb{E}_x \int_0^t e^{-\delta s} \left(\lambda^*(X_s) \left[-c\kappa_s + \theta_A \left(\kappa_s (\ln \kappa_s - 1) + e^{-\frac{1}{\theta_A} (\Delta V(X_s) - c)} \right) \right. \right. \\ &\quad \left. \left. + \kappa_s \Delta V(X_s) \right] \right) ds \\ &\quad + \mathbb{E}_x \int_0^t e^{-\delta s} \left(\mu^*(X_s) \left[p\eta_s + \theta_D \left(\eta_s (\ln \eta_s - 1) + e^{-\frac{1}{\theta_D} (p - \Delta V(X_s - 1))} \right) \right. \right. \\ &\quad \left. \left. - \eta_s \Delta V(X_s - 1) \right] \right) ds + V(x). \end{aligned} \tag{42}$$

We now prove that the integrands in the expression are non-negative, i.e.,

$$-c\kappa_s + \theta_A \left(\kappa_s (\ln \kappa_s - 1) + e^{-\frac{1}{\theta_A} (\Delta V(X_s) - c)} \right) + \kappa_s \Delta V(X_s) \geq 0 \tag{43}$$

and

$$p\eta_s + \theta_D \left(\eta_s (\ln \eta_s - 1) + e^{-\frac{1}{\theta_D} (p - \Delta V(X_s - 1))} \right) - \eta_s \Delta V(X_s - 1) \geq 0. \tag{44}$$

But this is straightforward as expressions (43) and (44) are convex in κ and η respectively and, from the first order conditions, the values of κ_s and η_s that minimize

the integrands are:

$$\begin{aligned} \kappa_s &= e^{-\frac{1}{\theta_A}(\Delta V(X_s)-c)} \\ \eta_s &= e^{-\frac{1}{\theta_D}(p-\Delta V(X_s-1))}. \end{aligned} \tag{45}$$

Substituting the minimizing value of κ_s in (43) and η_s in (44) we get zeros. Hence we have proved that

$$\tilde{V}_t(x) \geq V(x). \tag{46}$$

□

A similar analysis would prove that if nature chose the Markovian policy as defined in (22) and we are free to choose any policy, we will again choose the threshold policy. So even if we are free to choose anything and nature is restricted to Markovian policies, we will choose the threshold policy. Giving more freedom to nature will only worsen the performance. So it makes sense for us to choose the threshold policy. Hence a threshold policy is optimum even if we are free to choose any \mathcal{F}_t -predictable intensities and in that case nature would also choose a Markovian policy to hurt us most.

4 Effect of ambiguity parameter on threshold control

In this section we will study the effect of change in ambiguity levels on threshold control.

We define the optimal value function explicitly as a function of $\phi := (\theta_A, \theta_D)$ as

$$\begin{aligned} V^\phi(x) := \max_u \min_{\kappa, \eta} \mathbb{E}_x \left[\int_0^\infty e^{-\delta t} (p\mu_s - c\kappa_s\lambda_s + \theta_A\lambda_s(1 - \kappa_s + \kappa_s \ln \kappa_s) \right. \\ \left. + \theta_D\mu_s(1 - \eta_s + \eta_s \ln \eta_s)) ds \right]. \end{aligned} \tag{47}$$

We also define a partial order on ϕ , i.e., $\phi_1 \geq \phi_2$ if $\theta_{1A} \geq \theta_{2A}$ and $\theta_{1D} \geq \theta_{2D}$.

The following property of the value function is obvious from its definition.

Proposition 1 *If $\phi_1 \leq \phi_2$ then $V^{\phi_1}(x) \leq V^{\phi_2}(x)$ for all $x \in \mathcal{N}$.*

Let $b(\phi)$ be the value of the optimal threshold control corresponding to the parameter ϕ . We now show that the threshold remains bounded.

Proposition 2 *$b(\phi) < \infty$ for all $\phi \in [0, \infty] \times [0, \infty]$.*

Proof If $b(\phi) = \infty$ for some ϕ then $\lim_{x \rightarrow \infty} V^\phi(x) = \infty$ as $\Delta V^\phi(x) > c$ for all x . But the function $V^\phi(\cdot)$ is uniformly (in x) less than the value function for the unambiguous problem. The value function of the unambiguous problem can be uniformly

bounded by setting $\alpha_t = z$ and $\beta_t = 0$ in (8). Hence $\lim_{x \rightarrow \infty} V^\phi(x) = \infty$ is not possible. \square

We can now prove that the optimal threshold control is monotone in θ_A for fixed θ_D .

Proposition 3 *Let $\phi_1 = (\theta_{1A}, \theta_D)$ and $\phi_2 = (\theta_{2A}, \theta_D)$. If $\theta_{1A} < \theta_{2A}$ then $b(\phi_1) \geq b(\phi_2)$.*

Proof If $x > b(\phi)$ for some $\phi = (\theta_A, \theta_D)$, then $\Delta V^\phi(x) < c$ and hence

$$V^\phi(x) = z\theta_D \left(1 - e^{-\frac{1}{\theta_D}(p - \Delta V^\phi(x-1))} \right) + vV^\phi(x) \tag{48}$$

which implies

$$\delta V^\phi(x) + z\theta_D \left(e^{-\frac{1}{\theta_D}(p - \Delta V^\phi(x-1))} \right) = z\theta_D. \tag{49}$$

Subtracting $\delta V^\phi(x - 1)$ from both sides we get

$$\delta(\Delta V^\phi(x - 1)) + z\theta_D \left(e^{-\frac{1}{\theta_D}(p - \Delta V^\phi(x-1))} \right) = z\theta_D - \delta V^\phi(x - 1). \tag{50}$$

Suppose on the contrary $b(\phi_1) < b(\phi_2) < \infty$. By definition of $b(\phi_2)$, $\Delta V^{\phi_2}(b(\phi_2)) \geq c$ but $\Delta V^{\phi_2}(b(\phi_2) + 1) < c$. Therefore substituting $x = b(\phi_2) + 1$ we get the following from (50):

$$\delta(\Delta V^{\phi_2}(b(\phi_2))) + z\theta_D \left(e^{-\frac{1}{\theta_D}(p - \Delta V^{\phi_2}(b(\phi_2)))} \right) = z\theta_D - \delta V^{\phi_2}(b(\phi_2)). \tag{51}$$

Also as $b(\phi_1) < b(\phi_2)$

$$\delta(\Delta V^{\phi_1}(b(\phi_2))) + z\theta_D \left(e^{-\frac{1}{\theta_D}(p - \Delta V^{\phi_1}(b(\phi_2)))} \right) = z\theta_D - \delta V^{\phi_1}(b(\phi_2)). \tag{52}$$

As $\phi_2 \geq \phi_1$, $V^{\phi_2}(b(\phi_2)) \geq V^{\phi_1}(b(\phi_2))$. Hence

$$\begin{aligned} &\delta(\Delta V^{\phi_2}(b(\phi_2))) + z\theta_D \left(e^{-\frac{1}{\theta_D}(p - \Delta V^{\phi_2}(b(\phi_2)))} \right) \\ &\leq \delta(\Delta V^{\phi_1}(b(\phi_2))) + z\theta_D \left(e^{-\frac{1}{\theta_D}(p - \Delta V^{\phi_1}(b(\phi_2)))} \right). \end{aligned} \tag{53}$$

The function $\delta s + z\theta_D e^{-\frac{1}{\theta_D}(p-s)}$ is an increasing function of s for $s \geq 0$. So the only way (53) can be true is if

$$\Delta V^{\phi_2}(b(\phi_2)) \leq \Delta V^{\phi_1}(b(\phi_2)).$$

As $\Delta V^{\phi_2}(b(\phi_2)) \geq c$, so $\Delta V^{\phi_1}(b(\phi_2)) \geq c$. This contradicts $b(\phi_1) < b(\phi_2)$. \square

Threshold Control Variation with θ_A and θ_D . Parameters are $z=0.5, y=0.4, c=1, p=5, \delta=0.05$

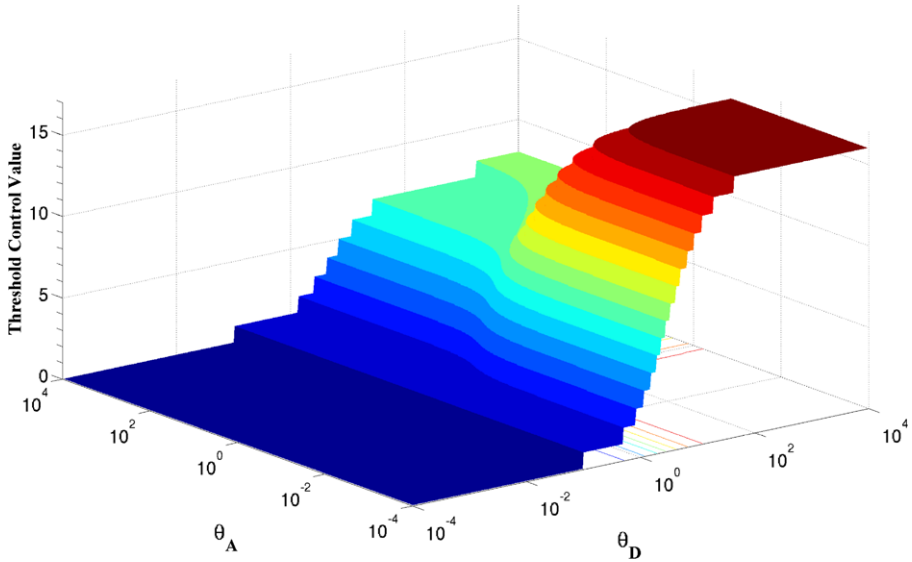


Fig. 2 Threshold control variation with ambiguity levels

Threshold Control Variation with $\theta_A = \theta_D$. Parameters are $z=0.5, y=0.4, c=1, p=5, \delta=0.05$

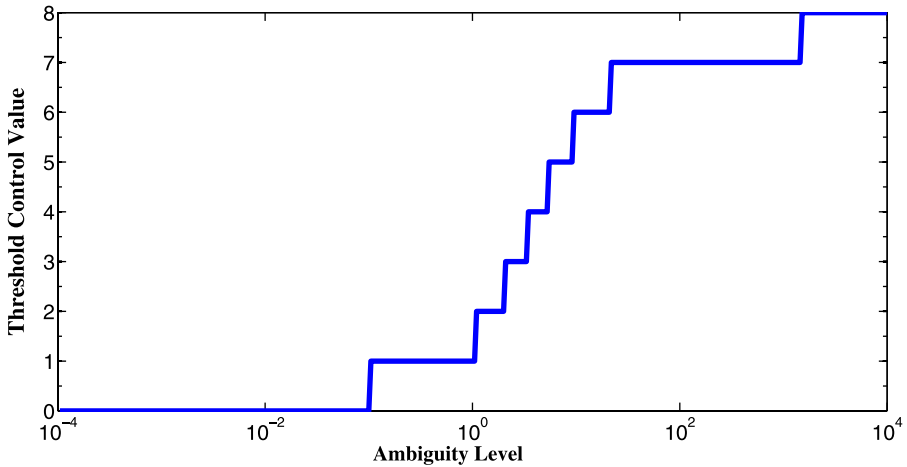


Fig. 3 Threshold is increasing if $\theta_A = \theta_D$

Numerical experiments also indicate (for various choices of parameters) that the threshold value is an increasing function of θ_D . Thus the ambiguity in arrival and ambiguity in departure appear to act in opposite directions (see for example Fig. 2). It is therefore important to consider the case when the two ambiguity levels are the

same. In our numerical experiments for $\theta_A = \theta_D$, the threshold control is increasing in the common ambiguity level (Fig. 3).

5 Conclusion

In this paper, we considered the problem of worst-case robust intensity control of the arrival and departure processes of a single-state queueing system, where model ambiguity is represented using the notion of relative entropy. A novel feature of our model is that we consider different levels of uncertainty for the arrival and departure processes. We prove that the optimal robust control for our model is of threshold type.

The paper can be extended in several directions. One possibility is to consider state-dependent capacity limits. Another extension is to consider multistage networks. Finally it would be interesting to consider a decentralized version of our problem where arrivals and departures are controlled by separate entities.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Bertsekas, D.P.: Dynamic Programming and Optimal Control, vol. II. Athena Scientific, Belmont (1995)
2. Bremaud, P.: Point Processes and Queues: Martingale Dynamics. Springer, New York (1981)
3. Chen, H., Yao, D.: Optimal intensity control of a queueing system with state-dependent capacity limit. *IEEE Trans. Autom. Control*, **35**(4) (1990)
4. Durrett, R.: Probability: Theory and Examples. Duxbury, N. Scituate (2003)
5. Hansen, L.P., Sargent, T.J., Turmuhambetova, G.A., Williams, N.: Robust control and model misspecification. *J. Econ. Theory* **128**(1), 45–90 (2006)
6. Jacod, J.: Calcul stochastique et problèmes de martingales. In: *Lecture Notes in Math.*, vol. 714. Springer, Berlin (1979)
7. Li, L.: A stochastic theory of the firm. *Math. Oper. Res.* **13**, 447–466 (1988)
8. Lim, A.E.B., Shanthikumar, J.G.: Relative entropy, exponential utility, and robust dynamic pricing. *Oper. Res.* **55**(2), 198–214 (2007)
9. Lim, A.E.B., Shanthikumar, J.G., Watwai, T.: Robust multi-product pricing and a new class of risk-sensitive control problems. Working Paper (2009)
10. Petersen, I.R., James, M.R., Dupuis, P.: Minimax optimal control of stochastic uncertain systems with relative entropy constraints. *IEEE Trans. Autom. Control* **45**, 398–412 (2000)
11. Pra, P. Dai, Meneghini, L., Runggaldier, W.J.: Connections between stochastic control and dynamic games. *Math. Control Signals Syst.* **9**, 303–326 (1996)
12. Serfozo, R.: Optimal control of random walks, birth and death processes, and queues. *Adv. Appl. Probab.* **13**, 61–83 (1981)
13. Shapley, L.S.: Stochastic games. *Proc. Natl. Acad. Sci.* **39**, 1095–1100 (1953)
14. Stidham, S.: Optimal control of admission to a queueing system. *IEEE Trans. Autom. Control* **30**, 705–713 (1985)