

Structural Properties of Markov Modulated Revenue Management Problems

Can Özkan^a, Fikri Karaesmen^{*,a,1}, Süleyman Özekici^a

^a*Department of Industrial Engineering, Koç University, 34450 Sarıyer-Istanbul, Turkey.*

Abstract

The admission decision is one of the fundamental categories of demand-management decisions. In the dynamic model of the single-resource capacity control problem, the distribution of demand does not explicitly depend on external conditions. However, in reality, demand may depend on the current external environment which represents the prevailing economic, financial, social or other factors that affect customer behavior. We formulate a Markov Decision Process (MDP) to maximize expected revenues over a finite horizon that explicitly models the current environment. We derive some structural results of the optimal admission policy, including the existence of an environment-dependent thresholds and a comparison of threshold levels in different environments. We also present some computational results which illustrate these structural properties. Finally, we extend some of the results to a related dynamic pricing formulation.

Key words: Revenue Management, Dynamic programming, Markov Modulation

1. Introduction

Revenue management is a field that originates in the Airline Deregulation Act of 1978 (Talluri and van Ryzin (2004 a)). There have been many studies since 1978 on different aspects of revenue management. Detailed overviews can be found in Talluri and van Ryzin (2004 a) and Chiang et al. (2007). An important building block model for more complicated revenue management is single resource capacity control. It is common in airline companies to sell identical seats at different fares. The major issue is the decision process of accepting or rejecting a booking request of a certain class for a given resource. The static model in which

*Corresponding author

Email addresses: canozkan@ku.edu.tr (Can Özkan), fkaraesmen@ku.edu.tr (Fikri Karaesmen), sozekici@ku.edu.tr (Süleyman Özekici)

¹Tel: 90(212)338-1718 Fax: 90(212)338-1548

different fare classes arrive at different, nonoverlapping time stages ordered in an increasing fare class rewards, is first considered by Littlewood (1972). The dynamic programming model of this problem is analyzed by Lee and Hersh (1993), and the structure of the optimality policy is investigated by Lautenbacher and Stidham (1999). For further research on single resource capacity control, see Brumelle and McGill (1993), Talluri and van Ryzin (2004 b), Barz and Waldmann (2007), Lan et al. (2008), Birbil et al. (2009), and Aydın et al. (2009).

Many sophisticated models exist for the single resource problem in the revenue management literature. Most of these models assume that the arrival process of fare classes is independent of external factors that may be varying randomly over the planning horizon. On the other hand, there are situations where the demand rate is strongly dependent on some external process, which we call the environmental process. We model this environmental process through a Markov chain. We consider the single resource capacity control problem in revenue management in such fluctuating demand environments. We refer to the corresponding model as Markov-modulated single resource capacity control. Such a model has not been discussed widely in a revenue management context. To our knowledge, the only study that explicitly models this situation is Chapter 4 of Barz (2007). In that chapter, Barz considers an environmental process for a single-resource control problem under very general assumptions. In particular, her model considers an infinite horizon problem with possibly random planning horizons. She shows that the optimal admission policy must be of threshold type for this generic model. On the other hand, the complexity of her model prevents further structural results on the effects of time, environments and other relevant model parameters. To investigate these properties, we consider a discrete-time finite horizon problem that is less general than that of Barz but otherwise follows the standard assumptions with respect to the general literature. This model enables us to investigate time-related properties of the optimal policy and the effects of external environments. Moreover, for this model, we can also analyze the effects of varying problem parameters such as arrival rates, rewards and the transition matrix of the environmental process. This not only extends the results of Aydın et al. (2009) to a more general setting but also allows comparing optimal policies in different environments and for varying environment process parameters. Overall, this analysis presents a complete picture for this problem.

Even though the fluctuating demand environment is little studied in revenue management, there is a significant number of papers related to Markov-modulated models for inventory systems, see Song and Zipkin (1993), Özekici and Parlar (1999), Arifoğlu and Özekici (2010), Gayon et al. (2009). For example, Song and Zipkin (1993) argue that demand frequently depends on external factors which they call the current state of the world. They also

argue that this current state of the world can be described by factors based on economic, financial and other conditions. Van Ryzin (2005) emphasizes the needs for better demand modeling for revenue management. In particular, he mentions that standard demand models in revenue management treat causal variations based on external factors as noise. The environment-based framework addresses this issue. Van Ryzin (2005) points out short term market conditions as a significant factor. These include competitors' availabilities and prices. In addition, there is evidence that the aggregate demand is affected by external market forces such as currency exchange rates and energy prices. Finally, weather conditions, such as forecasted snow storms or heat waves are important short term external factors that are known to impact demand in hotel and airline revenue management. Although these external factors seem to be very different from each other, they all influence the demand. This motivates the need for modeling the effects of such factors through an environment-dependent demand model. Finally, there is reason to believe that environment-based demand may have a bigger impact in revenue management problems than in inventory replenishment problems. In inventory replenishment, the ordering decision helps absorbing some of the variability in demand. But in revenue management there is typically no replenishment opportunity and demand variability has to be addressed only by admission or pricing decisions.

The rest of the paper is organized as follows. We first provide the model notation and formulation in Section 2. Then, we identify some structural properties of the optimal admission control policy in Section 3. We analyze the effects of varying problem parameters in Section 4. In Section 5 we provide an example to illustrate our structural results and assess the benefits of using an environment based model. Next, we extend the idea of a fluctuating demand environment to a dynamic pricing problem in Section 6. Finally, in Section 7 we conclude the paper. Most of the technical details involving proofs and derivations are relegated to the Appendix.

2. Model Formulation

We formulate a discrete time, finite horizon (T periods) MDP model of the admission control problem corresponding to single-leg capacity control.

Let $E_t \in \{1, 2, \dots, M\}$ denote the randomly fluctuating external environment. $E = \{E_0, E_1, \dots, E_T\}$ is assumed to be a Markov chain with transition matrix \mathbf{P} where $p_{ij} = P\{E_{t+1} = j | E_t = i\}$. We assume that there is at most one arrival and that each arrival from a fare class can request a finite number of seats in each stage. The probability that fare class a arrives at any stage is denoted by r_{ja} when the current environment is j . The

probability of no arrival in a given environment is denoted by r_{j0} . Therefore, $\sum_{a=0}^N r_{ja} = 1$ for any j . Non-stationary demand scenarios can be handled by defining appropriate environment and transition matrices. For each fare class a , suppose there is an upper bound B_a on the number of fare products requested. Let q_{jab} denote the probability that b units of inventory is requested given that current environment is j and the requested fare class is a .

In each stage t , the firm must choose the optimal number of seats to be sold for each fare class. We assume that customers accept the scenario of a partial satisfaction of their request. Brumelle and Walczak (2003) showed that structural results on the optimal policy are not valid in case of acceptance or rejection of the whole demand when there is no environment process (Also, see Van Slyke and Young (2000) and Çil et al. (2007) for related issues). Therefore, we only analyze the case where customers accept the partial satisfaction of their requests. For each sold ticket, the reward is $c(a)$ if the fare class is a . The transition probabilities and reward function are assumed to be stationary and we suppose that the fare classes are ordered so that $c(a_1) \leq c(a_2)$ when $a_1 \leq a_2$. We let \mathbb{Z}_+ denote the set of positive integers and \mathbb{R} denote the set of real numbers.

We also use the following notations:

$v_t(x, j)$ = expected maximum revenue from period t on, given that current inventory level is x and environment is j .

$\Delta v_t(x, j) = v_t(x, j) - v_t(x - 1, j)$

$(x)^+ = \max\{x, 0\}$

$U(b, x) = \{0, 1, \dots, \min\{b, x\}\}$

The optimal expected revenue and the admission control policy for this problem can be obtained by solving the following Bellman equation

$$v_t(x, j) = \sum_{a=1}^N r_{ja} \sum_{b=1}^{B_a} q_{jab} \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - u, k) + c(a)u \right\} + r_{j0} \sum_{k=1}^M p_{jk} v_{t+1}(x, k) \quad (1)$$

with boundary conditions

$$v_t(0, j) = 0 \text{ for } j = 1, 2, \dots, M.$$

$$v_T(x, j) = 0 \text{ for any } x \in \mathbb{Z}_+ \text{ and } j = 1, 2, \dots, M.$$

For obtaining structural results, the following equivalent representation that uses the

definition of Δv_t turns out to be helpful

$$\begin{aligned}
v_t(x, j) &= \sum_{a=1}^N r_{ja} \sum_{b=1}^{B_a} q_{jab} \max_{u \in U(b, x)} \left\{ c(a)u - \sum_{k=1}^M p_{jk} \left(\sum_{z=1}^u \Delta v_{t+1}(x+1-z, k) - v_{t+1}(x, k) \right) \right\} \\
&\quad + r_{j0} \sum_{k=1}^M p_{jk} v_{t+1}(x, k) \\
&= \sum_{a=1}^N r_{ja} \sum_{b=1}^{B_a} q_{jab} \max_{u \in U(b, x)} \left\{ \sum_{z=1}^u \left(c(a) - \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x+1-z, k) \right) \right\} + \sum_{k=1}^M p_{jk} v_{t+1}(x, k)
\end{aligned} \tag{2}$$

where the sum is set to be zero when $u = 0$.

3. Structural Properties

In this section, we investigate some structural properties of the Markov-modulated single-resource capacity control problem. To begin with, it is intuitive that if we have one more inventory, then expected revenue should be larger. Similarly, expected revenue should be larger if we have more time to go. These claims can be easily proven by induction on t . Second order properties are less trivial. In the following theorem, we establish the concavity of $v_t(x, j)$ in x .

Theorem 1. $v_t(x, j)$ is a concave function in x for any environment j and time t .

We provide the proofs of this section in Appendix A. Theorem 1 establishes that $\Delta v_t(x, j)$ decreases as we increase the inventory level x . By considering (2), we can conclude that $c(a) - \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x+1-z, k)$ is decreasing in z . Therefore, in (2) we should increase $u \leq \min\{b, x\}$ until $c(a) - \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x+1-z, k)$ becomes negative or u is equal to $\min\{b, x\}$. Since $\Delta v_t(x, j)$ is decreasing in x for any j , there is a threshold level $l_t^{a,j}$ which is defined as

$$l_t^{a,j} = \min \left\{ x : c(a) \geq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k) \right\}. \tag{3}$$

Explicitly, $l_t^{a,j}$ is the maximum quantity for the inventory level such that if the current inventory level is less than $l_t^{a,j}$ it is optimal to reject any batch for fare class a in environment j . However, if the inventory on hand is greater than or equal to $l_t^{a,j}$, then demand for fare class a is satisfied until the inventory level drops to $l_t^{a,j} - 1$.

Hence the optimal decision for fare class a at stage t and environment j , when demand is b , is

$$u^* = \min \left\{ (x - l_t^{a,j} + 1)^+, b \right\}. \quad (4)$$

Theorem 1 implies that optimal admission control policies are of threshold (or booking limit) type as in standard single-resource capacity control. The difference in this case is that the thresholds now depend on the current state of the environment. Nevertheless, such policies are relatively easy to implement.

Since optimal thresholds are determined by the marginal value function via (3), we next investigate the structure of this function. First, we analyze how the marginal value function changes in time. Next proposition states a result on the marginal value of one additional inventory over time.

Proposition 1. $\Delta v_{t+1}(x, j) \leq \Delta v_t(x, j)$ for any inventory level x , environment j and time t .

Please note that Theorem 1 and Proposition 1 extend the corresponding results in Aydın et al. (2009) to a setting with multiple environments. With regard to Proposition 1, a corresponding result in Aydın et al. (2009) establishes that admission thresholds decrease as time increases when there is a single environment. When there are multiple environment states, the environment also changes over time; hence, we cannot guarantee the decrease of the thresholds over time when the environment changes. On the other hand, if the environment does not change, then we can establish the admission threshold should decrease in the next period. This result follows from comparing

$$\sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k) \leq \sum_{k=1}^M p_{jk} \Delta v_t(x, k)$$

which is obviously true by Proposition 1. Therefore, $l_t^{a,j} \geq l_{t+1}^{a,j}$ for any fare class a , time t and environment j .

Since demand varies according to the environment, optimal threshold levels change with the environment. To better understand the effects of the environment on the optimal thresholds, we must classify and order the environments. To this end, we need some assumptions on arrival probabilities and the transition matrix of the environmental process. The following classification is useful for this purpose.

Definition 1. A Markov chain is said to be IFR (Increasing Failure Rate) if the rows of its

transition probability matrix are in increasing stochastic order, i.e.,

$$f(i) = \sum_{j=k}^M p_{ij}$$

is nondecreasing in i for all $k = 1, \dots, M$. Similarly, a matrix X is said to be IFR if the rows of X are in increasing stochastic order.

In reliability theory, life distribution classifications, like IFR, play a crucial role in identifying the structure of optimal maintenance policies. This usually leads to optimal threshold policies since the IFR property implies the increasing marginal deterioration of the system. An example is the age replacement policy which states that the system is replaced as soon as its age exceeds a critical level. The reader is referred to Barlow and Proschan (1965) for basic concepts on life distribution classifications, and Keilson and Kesten (1977) for classifications of Markov chains using their transition matrices. In our context, we need to impose similar restrictions on the environmental process so that the state becomes more or less “desirable” in generating revenue.

Let \mathbf{R} be a matrix such that $R_{j,a} = r_{ja}$, and suppose \mathbf{R} is IFR. This implies that environments are ordered in terms of the arrival probability of customers from higher fare classes. For example, suppose we have 2 environments, then the second environment is said to be “better” than the first one if it is more probable to have a demand for higher reward fare-classes in the second environment.

Let $B = \max\{B_a : a = 1, 2, \dots, N\}$ and set $q_{jab} = 0$ for any $B_a < b \leq B$. Also, let \mathbf{Q} denote a 3 dimensional matrix whose (j, a, b) th component is q_{jab} as defined above. Then we define 2 dimensional submatrices of \mathbf{Q} where we fix one component of \mathbf{Q} . Let the fixed component be denoted as a superscript while the other components are denoted by subscripts. We also assume that the matrix $\mathbf{Q}_{\mathbf{j}\mathbf{b}}^{(a)}$ is IFR for a fixed a . Finally, we also assume that $\mathbf{Q}_{\mathbf{a}\mathbf{b}}^{(i)}$ is IFR for fixed i .

Last, we need a condition on the transition matrix \mathbf{P} of environment process. We assume that \mathbf{P} is IFR. This is also plausible. If the index of an environment i is higher than another environment j , then we call i a “better” environment than j by the explanation above. Since environment i is better than j , it is more likely for environment i to make a transition to an environment that is better than an arbitrary given environment. Intuitively, the probability that the current environment will transition in the future to a better environment increases as the level of the current environment increases. We now summarize all of these conditions.

Condition 2. (1) \mathbf{P} is IFR.

- (2) \mathbf{R} is IFR.
- (3) $\mathbf{Q}_{\mathbf{jb}}^{(\mathbf{a})}$ is IFR for any fare class a .
- (4) $\mathbf{Q}_{\mathbf{ab}}^{(\mathbf{j})}$ is IFR for any environment j .

The above condition imposes an order on the environments. This order is a minimal requirement for obtaining structural results as a function of the environment. When the condition holds, j is a more favorable environment than i where $i \leq j$. Let us discuss the modeling implications of Condition 2. Condition 2 (1) concerns the environment transitions. Several environment-based models in the literature have two environments in which case the condition can be easily expressed and verified. There are also special but plausible transition structures where the condition can be shown to hold (see Gayon et al. (2009) for examples in a different model). In short, the environment transitions need to have a smoothness property where better current environments are likelier to lead to better future environments which seems natural for most applications with a few environment states representing aggregate conditions. Condition 2 (2) can be viewed as a consequence of the environment classification where better current states have a more favorable demand arrival distribution. Without this condition, environment states do not necessarily have a natural order which prevents monotonicity. Condition 2 (3) states that batch sizes are likelier to be larger in better environments which also appears natural. Condition 2 (4) imposes constraints on the demand batch size as a function of the class of customers. This condition is automatically satisfied for the frequently encountered case of unit demand arrivals (see Talluri and van Ryzin (2004 a)) and for the case where the batch sizes are not class dependent.

We first investigate the expected maximum revenue from period t on for different environments at stage t under Condition 2. In particular, in the next proposition, we establish that the maximum expected revenue increases when the environment gets better .

Proposition 2. *Under Condition 2, $v_t(x, i) \leq v_t(x, j)$ for any inventory level x , environment $i \leq j$ and time t .*

From a practical perspective, Proposition 2 states that better starting environments lead to better expected revenues. Second, we consider the effect of the environment on the expected marginal value of one additional inventory. This value is important in understanding the structure of threshold values in different environments.

Proposition 3. *Under Condition 2, $\Delta v_t(x, i) \leq \Delta v_t(x, j)$ for any inventory level x , environment $i \leq j$ and time t .*

Let us discuss the implication of this proposition. Since the admission policy is determined by the structure of the difference $\Delta v_t(x, j)$, and since this difference increases in j , we can conclude that $l_t^{a,j}$ increases in j . Since the demand for a more valuable fare class will increase in probability as the environment gets better, it is optimal to protect the stock more in a better environment. For implementation purposes, this implies that the optimal admission thresholds are non-decreasing in more favorable environments. By using Propositions 1 and 3, we have the following immediate result that extends the property in Proposition 3 to different time periods.

Corollary 1. *Under Condition 2, $\Delta v_{t+1}(x, i) \leq \Delta v_t(x, j)$ for any inventory level x , environment $i \leq j$ and time t .*

By this corollary, we know that the threshold level of a fare class in a given stage will decrease in the next stage if the environment of the next stage is worse than the one in the previous stage. However, we cannot guarantee the decrease of the threshold if the environment of the next stage is better than the one in the previous stage. This is explored further in Section 5.

4. Sensitivity Analysis

In this section, we will provide results on the sensitivity of the structural properties on the model parameters. A recent paper by Çil et al. (2009) presents a general approach for this type of analysis and Aydın et al. (2009) presents corresponding results for a standard single-leg capacity control problem.

First, by setting $z_{jab} = q_{jab}r_{ja}$, we will use the following equivalent form of our problem

$$v_t(x, j) = \sum_{a=1}^N \sum_{b=1}^{B_a} z_{jab} \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - u, k) + c(a)u \right\} + r_{j0} \sum_{k=1}^M p_{jk} v_{t+1}(x, k) \quad (5)$$

with boundary conditions $v_t(0, j) = 0$ and $v_T(x, j) = 0$ for all x and t . We show the effects of changing components of arrival probabilities (\mathbf{Z}), transition matrix (\mathbf{P}), and reward function (c). We will change a component of these matrices or the reward function by a small amount and explore the effects of this change under some specific conditions.

We first provide the results on the effects of varying the arrival probabilities. Aydın et al. (2009) also considers a similar study on the effects of parameters, where only the fictitious event probability is decreased when a given arrival probability is increased. We employ a more general approach and consider decreasing any other arrival probability.

Let us increase z_{iab} by $\epsilon \geq 0$ for a given environment i , class $a \geq 1$ and batch size b . In order to have a valid probability distribution we will reduce $z_{ia_2b_2}$ by ϵ where $1 \leq a_2 \leq a$ and $b_2 \leq b$. Here, ϵ should be small enough in order to have both $z_{iab} + \epsilon$ and $z_{ia_2b_2} - \epsilon$ lie in the interval $[0, 1]$. Let $v_t^\epsilon(x, j)$ be the value function for the modified system.

Proposition 4. $v_t^\epsilon(x, j) \geq v_t(x, j)$ for any environment j , time t and inventory level x .

We provide all the proofs of this section in Appendix B.

Proposition 4 formalizes that increased demand from more valuable classes improves expected revenues. Next, we consider the effects of varying arrival probabilities on the expected marginal value of one additional inventory, since admission thresholds are determined by this value. The marginal value for the modified system is denoted by $\Delta v_t^\epsilon(x, j)$.

Proposition 5. $\Delta v_t^\epsilon(x, j) \geq \Delta v_t(x, j)$ for any environment j , time t and inventory level x .

Since expected marginal value of one additional inventory is greater in the modified system, the threshold level of the modified system for a given fare class, time and environment is greater than the one of the original model. In other words, $l_t^{a,j} \leq l_{t,\epsilon}^{a,j}$ where $l_{t,\epsilon}^{a,j}$ denotes the threshold level in the modified system. Please note that an increase in some arrival probability at a given environment i causes the admission thresholds in all environments j to increase. Propositions 4 and 5 extend the corresponding results in Aydın et al. (2009) to multiple environment states.

Second, we analyze the effects of changing a component of \mathbf{P} which is assumed to be IFR. Suppose we increase p_{ij} by $\epsilon \geq 0$. To have a valid distribution, we need to reduce another component in the i th row of \mathbf{P} with a column index smaller than j by ϵ . Again, ϵ should be small enough to make the changed components lie in $[0, 1]$ interval. These changes must preserve the IFR property of \mathbf{P} . Let the modified solution be denoted by $v_t^\epsilon(x, j)$ and the transition probability matrix by \mathbf{P}^ϵ . We have only changed the i th row of \mathbf{P} , hence the remaining rows of \mathbf{P}^ϵ are identical to \mathbf{P} . First, we compare the expected revenue of these two systems.

Proposition 6. Under Condition 2, $v_t^\epsilon(x, j) \geq v_t(x, j)$ for any environment j , time t and inventory level x .

Proposition 6 establishes that a better environment probability transition matrix leads to higher expected revenues. For practical purposes, this implies that more favorable forecasts of future demand environments results in improved expected revenues.

Next, we investigate the effects of such a change on the optimal policy. As done in the previous analysis on \mathbf{Z} , we focus on the expected marginal value of one additional inventory level. The marginal value of the modified system is denoted by $\Delta v_t^\epsilon(x, j)$. In the next proposition we show that marginal value of the modified system is greater than the original system.

Proposition 7. *Under Condition 2, $\Delta v_t^\epsilon(x, j) \geq \Delta v_t(x, j)$ for any environment j , time t and inventory level x .*

As explained before, since the expected marginal value of an additional inventory is greater in the modified system, threshold level of the modified system is greater than the one of the original system. (i.e., $l_{t,\epsilon}^{a,j} \geq l_t^{a,j}$). A more advantageous environment transition structure leads to higher admission thresholds for all environments.

Now, we investigate the sensitivity of the marginal value function in the reward of each fare class. We will increase the reward of a specific fare class and try to see its impact. We define $\Delta v_{t+1}^\epsilon(x, k)$ as the marginal value of an additional inventory. We have the following proposition about the effects of reward on the marginal value.

Proposition 8. *$\Delta v_t^\epsilon(x, j) \geq \Delta v_t(x, j)$ if $c(N)$ is increased by $\epsilon \geq 0$.*

Proposition 8 establishes increasing the reward of the highest class leads to higher admission thresholds: $l_{t,\epsilon}^{a,j} \geq l_t^{a,j}$. As before, somewhat surprisingly, a positive perturbation of $c(N)$ requires a stronger protection for class N and therefore has a non-decreasing effect for all admission thresholds. Please note a corresponding result exists in Aydın et al. (2009) for the case with a single environment state.

We have also investigated the effect of increasing the reward of any other fare class rather than the one with the highest reward. It is not always true that the marginal value of an additional inventory in a modified system is greater than the one in the original system or vice versa. We have a counter-example in the next section.

5. Numerical Illustrations

In our illustrations, we assume that an arrival customer demands only one product, this implies that $B_a = 1$ for any fare class a . First, we illustrate that the threshold level decreases as time increases and increases as the environment gets better. The transition matrix, reward

vector, and arrival probability matrix are respectively:

$$\mathbf{P} = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix} \quad c(a) = \begin{cases} 0 & \text{if } a = 0 \\ 50 & \text{if } a = 1 \\ 100 & \text{if } a = 2 \\ 200 & \text{if } a = 3 \end{cases} \quad \mathbf{R} = \begin{bmatrix} 0.7 & 0.2 & 0.1 & 0 \\ 0.1 & 0.2 & 0.2 & 0.5 \end{bmatrix} \quad (6)$$

with planning horizon $T = 500$. Note that 0 in vector c stands for the reward of the fictitious event. We only show the last 10 threshold levels in our tables. Note that \mathbf{R} has the IFR property, hence we can label the first row of \mathbf{R} as a bad environment and the second row as a good environment. Threshold levels for fare class 1 (with reward 50) and 2 (with reward 100) are given in Table 1. Recall that $l_t^{a,j}$ stands for the threshold level of fare class a at time t in environment j . As we expect, the threshold level decreases as time increases for any environment and the threshold level of a better environment is higher at any given time. Also we know that the threshold level for fare class 3 (with reward 200) is always 1 for any environment and time. Since, we always accept a request for the fare class with the highest reward. Finally, recall that Corollary 1 established that $\Delta v_{t+1}(x, i) \leq \Delta v_t(x, j)$ for $i \leq j$. It can be observed from Table 1 that the condition $i \leq j$ is crucial. In fact, we observe that $l_6^{1,1} \leq l_7^{1,2}$. Therefore, the threshold is not necessarily monotone in all cases.

Table 1: Threshold levels for fare classes 1 and 2 in both environments

Time t	1	2	3	4	5	6	7	8	9	10
$l_t^{1,1}$	4	4	3	3	3	2	2	1	1	1
$l_t^{1,2}$	8	8	7	6	5	5	4	3	2	1
$l_t^{2,1}$	2	1	1	1	1	1	1	1	1	1
$l_t^{2,2}$	6	5	5	4	4	3	3	2	2	1

Next, we investigate the effects of changing the parameters of the problem. Suppose that we decrease the arrival probability of fare class 3 from 0.5 to 0.1 and increase the arrival probability of fare class 1 from 0.2 to 0.6 in environment 2, then we compare the threshold levels for fare classes 1 and 2 in both systems. See Tables 7 and 8 in Appendix C for the comparison of threshold levels for fare classes 1 and 2. As expected, threshold levels for fare classes 1 and 2 are smaller in the modified system in any environment.

We also change the entries of \mathbf{P} while the modified \mathbf{P} matrix still has the IFR property. Suppose that we have the following modified \mathbf{P}

$$\mathbf{P}^\epsilon = \begin{bmatrix} 0.05 & 0.95 \\ 0.05 & 0.95 \end{bmatrix} \quad (7)$$

which is obtained by changing the first row of the transition matrix. The threshold levels for fare class 2 in the modified system is greater than the one in the original system as shown in Table 2 for environment 1.

Table 2: Threshold levels of fare class 2 in environment 1 (Change in transition matrix)

Time t	1	2	3	4	5	6	7	8	9	10
$l_t^{2,1}$	2	1	1	1	1	1	1	1	1	1
$l_{t,\epsilon}^{2,1}$	7	6	5	5	4	4	3	2	2	1

Further, we increase the reward of the third fare class, which is the most expensive one, from 200 to 250. Threshold levels for fare classes 1 and 2 are given in the Tables 9 and 10 in Appendix D.

The threshold levels for fare classes 1 and 2 of the modified system are greater in both environments. We also provide a counter-example for the case when the reward of any other fare class rather than the most expensive one is changed. Suppose that we change the reward of fare class 2 from 100 to 150. The threshold levels of fare class 1 in environment 1 and fare class 2 in environment 2 are given in Table 3.

Table 3: Threshold levels for fare classes 1 and 2 (Change in price of product 2)

Time t	1	2	3	4	5	6	7	8	9	10
$l_t^{1,1}$	4	4	3	3	3	2	2	1	1	1
$l_{t,\epsilon}^{1,1}$	5	4	4	3	3	2	2	2	1	1
$l_t^{2,2}$	6	5	5	4	4	3	3	2	2	1
$l_{t,\epsilon}^{2,2}$	4	4	4	3	3	3	2	2	1	1

Note that the threshold levels of fare-class one increase; however, the threshold levels of fare-class two decrease. Therefore, it is not always true that expected marginal value of an additional inventory decreases (or increases) as we increase the reward of a fare class which is not the most expensive.

Remember that when \mathbf{R} is IFR, we can order the environments. In addition to this property, if \mathbf{P} is IFR, we know that $l_t^{a,j} \leq l_t^{a,i}$ whenever $j \leq i$. However, we cannot conclude the same result when \mathbf{P} is not IFR. We have the following counter-example to show this

claim. We use the same problem parameters except the matrix

$$\mathbf{P} = \begin{bmatrix} 0.05 & 0.95 \\ 0.95 & 0.05 \end{bmatrix} \quad (8)$$

which is not IFR anymore. The threshold levels for fare class 1 (with price 50) are given in the Table 4. Even though environment 2 can be considered better than environment 1, threshold levels of fare-class 1 at times 3, 5, 7 and 9 in environment 1 are greater than those in environment 2.

Table 4: Threshold levels for fare class 1 when P is not IFR

Time t	1	2	3	4	5	6	7	8	9	10
$l_t^{1,1}$	6	5	5	4	4	3	3	2	2	1
$l_t^{1,2}$	6	5	4	4	3	3	2	2	1	1

6. The Efficiency of the Environment-Based Model

To assess the performance of our environment based model, we consider a 2-environment problem in which arriving customers demand only one product at a time. In this setting, we compare the expected revenues from our model to a simple but reasonable benchmark approach where the system manager incorrectly believes that the system will always remain in one of the environment states (i.e. the environment will not fluctuate). In this case, the manager solves a simpler standard dynamic program to find the optimal admission policy.

To implement the benchmark approach, let us define $w_t^j(x)$, the maximum expected total revenue when the environment j is the environment believed to be true by the manager. The corresponding optimal policy can be formulated by the Bellman equation

$$w_t^j(x) = \sum_{a=1}^N r_{ja} \max \{w_{t+1}^j(x-1) + c(a), w_{t+1}^j(x)\} + r_{j0}w_{t+1}^j(x) \quad (9)$$

with boundary conditions $w_T^j(x) = 0$ and $w_t^j(0) = 0$ for all x and t . \mathbf{R} and \mathbf{P} are as given in (6) and $c(0) = 0$, $c(1) = 50$, $c(3) = 200$ and we vary $c(2)$ between 65 and 185 (using a step-size of 30).

For each environment state, we compute the optimal admission policy and use this policy in our environment-based model and calculate the corresponding expected revenue for an initial inventory level of 200 and horizon length of 500 starting with environment 1. In addition, we compute the expected optimal revenue using the environment-dependent model

for the same parameters. Figure 1 reports the the percentage differences in expected revenues due to using a simpler model for different values of $c(2)$. It can be observed that the difference is consistently over 15% when the manager employs the good environment state (maybe due to optimistic expectations). On the other hand, the difference varies significantly and appears to be an increasing as a function of $c(2)$ when the manager employs the bad environment state.

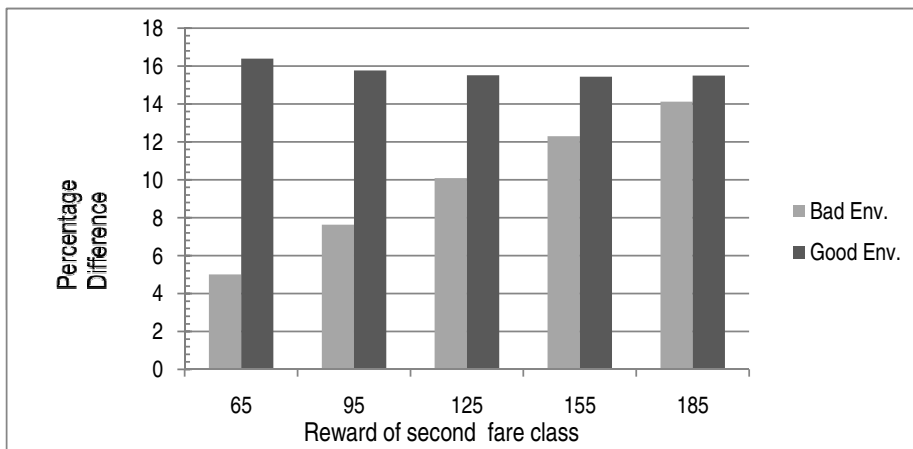


Figure 1: Percentage Differences when Single Environment Policies are Used Instead of the Environment-Based Policy

Next, we explore how the benefits of the environment-based model are affected by the demand profile similarity or dissimilarity in different environments. We use c and \mathbf{P} as given in (6) and we define $\mathbf{R}(\epsilon)$ (where $\epsilon = 0, 0.1, 0.2, 0.3, 0.4$) as follows:

$$\mathbf{R}(\epsilon) = \begin{bmatrix} 0.7 - \epsilon & 0.2 & 0.1 & 0 + \epsilon \\ 0.1 & 0.2 & 0.2 & 0.5 \end{bmatrix}.$$

Please note that increasing the value of ϵ makes the demand profiles in the two environments more similar. Therefore, when $\epsilon = 0$ the demand profiles are very different from each other and when $\epsilon = 0.4$, the demand profiles are fairly similar. For each ϵ , we repeat the same investigation as above and compare the revenues in an environment-based model with

a fixed environment model. The percentage differences as a function of the environment are reported in Table 5. As can be observed from the table, there is significant value in using the environment-based model when the demand profiles are different but as expected this value diminishes as the demand profiles of the different environments become similar.

Table 5: Percentage difference as function of R

ϵ	0	0.1	0.2	0.3	0.4
Bad Env.	8.06	4.69	1.78	0.26	0.0017
Good Env.	15.71	11.63	5.46	3.57	0.0060

Finally, we investigate the effect of total demand rate difference between the environments. The situation in mind we have is external factors that affect the aggregate demand rate in varying degrees. In particular, if the demand rate in the first environment for a given demand class a ($a = 1, 2, 3$) is r_{1a} , then the corresponding demand rate in the second environment is αr_{1a} where $0 < \alpha < 1$. For the numerical experimentation, we use c as given in (6) and the other parameters are given below.

$$\mathbf{R}(\alpha) = \begin{bmatrix} 1 - 0.9\alpha & 0.4\alpha & 0.3\alpha & 0.2\alpha \\ 0.1 & 0.4 & 0.3 & 0.2 \end{bmatrix}, \mathbf{P}_1 = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}, \mathbf{P}_2 = \begin{bmatrix} 0.7 & 0.3 \\ 0.1 & 0.9 \end{bmatrix}, \mathbf{P}_3 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

Using the above parameters, we experiment with three levels of α and repeat the earlier experimentation by comparing the revenues using the environment-based dynamic program versus revenues obtained by solving simpler single environment models. Please note that we also use three different transition matrices. The results are reported in Table 6. Once again, the benefits of using the environment-based model are important when the demand profiles (aggregate demand rates in this experiment) are different.

Table 6: Percentage difference when the total demand rate fluctuates

α	P₁		P₂		P₃	
	Good Env.	Bad Env.	Good Env.	Bad Env.	Good Env.	Bad Env.
0.25	3.56	8.15	0.20	17.34	3.62	8.49
0.5	0.62	3.18	0.07	5.13	0.51	3.10
0.75	0.07	0.06	0.01	0.13	0.06	0.06

7. Extension: Markov Modulated Dynamic Pricing

In this section, we extend our investigation to a corresponding dynamic pricing problem. A similar continuous-time problem with replenishment has been explored by Gayon et al. (2009). In dynamic pricing, customers are not segmented to different classes but they have different purchasing probabilities as a function of the offered price. The goal is to find the price to charge in a given state to maximize the expected revenue. We assume that there is only one customer in each stage and his willingness to pay is a random variable which depends on the current environment. If the current environment is j then the price he is willing to pay W_j has a distribution $F_j(v) = P\{W_j \leq v\}$. The distribution function is assumed to be differentiable and we denote the density by $f_j(p)$. We also assume that the distribution function has an inverse F_j^{-1} . Let $v_t(x, j)$ be the expected maximum revenue from period t on, given that current inventory level is x and environment is j . The manager needs to choose the price of the fare-class in each stage t with a given environment j and inventory level x . Therefore, we now have the following Bellman equation

$$v_t(x, j) = \max_{p \geq 0} \left\{ (1 - F_j(p)) \left(p + \sum_{k=1}^N p_{jk} v_{t+1}(x-1, k) \right) + F_j(p) \sum_{k=1}^N p_{jk} v_{t+1}(x, k) \right\}$$

with $v_T(x, j) = 0$ and $v_t(0, j) = 0$ as boundary conditions. Since distribution function is one-to-one, there exists a unique p such that $d = \bar{F}_j(p) = 1 - F_j(p)$ for any $0 \leq d \leq 1$. Therefore, we have the following equivalent formulation

$$v_t(x, j) = \max_{0 \leq d \leq 1} \left\{ d \left(p_j(d) + \sum_{k=1}^N p_{jk} v_{t+1}(x-1, k) \right) + (1-d) \sum_{k=1}^N p_{jk} v_{t+1}(x, k) \right\}$$

where $p_j(d) = F_j^{-1}(1-d)$. By using $\Delta v_t(x, j) = v_t(x, j) - v_t(x-1, j)$, we have

$$v_t(x, j) = \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \right\} + \sum_{k=1}^N p_{jk} v_{t+1}(x, k) \quad (10)$$

In (10), $dp_j(d)$ is the expected revenue during the current stage. Let $H_j(d)$ be the derivative of $dp_j(d)$ with respect to d , we make the following standard assumption as in Talluri and van Ryzin (2004 a).

Condition 3. For any environment j , $H_j(d)$ is a decreasing function in d , and this condition also implies that $H_j(\bar{F}_j(p)) = p - (1 - F_j(p)) / f_j(p)$ is increasing function of p .

By using this condition, we know that inner part of the maximization problem is a concave function in d , therefore; the optimal solution can be found by using

$$H_j(d^*) = \sum_{k=1}^N \Delta v_{t+1}(x, k) \quad (11)$$

For the rest of this section, we assume that $d^* \in (0, 1)$. To gain insights on the the structure of the optimal pricing policy we need to investigate the structure of Δv_t . First, we show that marginal revenue decreases as we have more inventory.

Proposition 9. $\Delta v_t(x, i)$ is a decreasing function of x for any environment i and time t .

We provide the proofs of propositions in E. Under Condition 3, Proposition 9 implies that the optimal prices are non-increasing in the inventory on hand. Next, explore the effect of time on the marginal revenue.

Proposition 10. $\Delta v_t(x, i)$ is a decreasing function of t for any environment i and inventory level x .

Proposition 10 provides further insights on the structure of the optimal pricing policy. Under Condition 3, Proposition 10 implies that the optimal prices are non-increasing in the remaining time for the same inventory level and environment. While the optimal price paths need to be non-increasing in general, they are so when the environment does not fluctuate.

8. Conclusion and Future Research

We investigated a single resource capacity control problem with a fluctuating demand environment. Modeling fluctuating demand through a Markov-modulated environment process is widely accepted in the inventory control literature. But there has not been much work on such models in capacity control problems rooted in revenue management.

We were able to provide a fairly complete set of structural results on the optimal admission policy under a Markov-modulated demand process. The structural results comprise the existence of environment-based thresholds but also extend to the effect of the time, environments and various problem parameters. Through numerical examples, we observe that the benefit from the environment-based model is significant if the conditions in different environments are distinctively different.

Some extensions of the model follow relatively easily as in the dynamic pricing case presented in Section 7. Other extensions such as consumer-choice behavior and network

revenue management merit further research. Another interesting and challenging line of extension is to consider uncertain environment transition rates or unobservable environments.

Acknowledgements: F. Karaesmen's research was partially supported by the TUBA-GEBIP programme.

References

- Arifoğlu, K., Özekici, S., 2010. Optimal policies for inventory systems with finite capacity and partially observed Markov-modulated demand and supply processes. *European Journal of Operational Research* 204, 421–483.
- Aydın, S., Akçay, Y., Karaesmen, F., 2009. On the structural properties of a discrete-time single product revenue management problem. *Operations Research Letters* 37 (4), 273–279.
- Barlow, R. E., Proschan, F., 1965. *Mathematical Theory of Reliability*. John Wiley and Sons, New York.
- Barz, C., 2007. Risk-averse capacity control in revenue management. No. 597. Springer-Verlag.
- Barz, C., Waldmann, K., 2007. Risk-sensitive capacity control in revenue management. *Mathematical Methods of Operations Research* 65 (3), 565–579.
- Birbil, Ş. İ., Frenk, J., Gromicho, J., Zhang, S., 2009. The role of robust optimization in single-leg airline revenue management. *Management Science* 55 (1), 148–163.
- Brumelle, S., McGill, J., 1993. Airline seat allocation with multiple nested fare classes. *Operations Research* 41 (1), 127–137.
- Brumelle, S., Walczak, D., 2003. Dynamic Airline Revenue Management with Multiple Semi-Markov Demand. *Operations Research* 51 (1), 137–148.
- Chiang, W., Chen, J., Xu, X., 2007. An overview of research on revenue management: current issues and future research. *International Journal of Revenue Management* 1 (1), 97–128.
- Çil, E., Örmeci, E., Karaesmen, F., 2007. Structural results on a batch acceptance problem for capacitated queues. *Mathematical Methods of Operations Research* 66 (2), 263–274.

- Çil, E., Örmeci, E., Karaesmen, F., 2009. Effects of system parameters on the optimal policy structure in a class of queueing control problems. *Queueing Systems* 61, 273–304.
- Gayon, J., Talay-Değirmenci, I., Karaesmen, F., Örmeci, E., 2009. Optimal pricing and production policies of a make-to-stock system with fluctuating demand. *Probability in the Engineering and Informational Sciences* 23 (2), 205–230.
- Keilson, J., Kesten, A., 1977. *Monotone Matrices and Monotone Markov Processes*. *Stochastic processes and their Applications* 5, 231–241.
- Lan, Y., Gao, H., Ball, M., Karaesmen, I., 2008. Revenue management with limited demand information. *Management Science* 54 (9), 1594–1609.
- Lautenbacher, C., Stidham, S., 1999. The Underlying Markov Decision Process in the Single-Leg Airline Yield-Management Problem. *Transportation Science* 33 (2), 136–146.
- Lee, T., Hersh, M., 1993. A Model for Dynamic Airline Seat Inventory Control with Multiple Seat Bookings. *Transportation Science* 27, 252–265.
- Littlewood, K., 1972. Forecasting and control of passengers. In *Proceedings of the 12th Annual AGIFORS Symposium*, 95–128.
- Özekici, M., Parlar, M., 1999. Inventory models with unreliable suppliers in a random environment. *Annals of Operations Research* 91, 123–136.
- Song, J., Zipkin, P., 1993. Inventory control in a fluctuating demand environment. *Operations Research* 41, 351–370.
- Talluri, K., van Ryzin, G., 2004 a. *The Theory and Practice of Revenue Management*. Springer.
- Talluri, K., van Ryzin, G., 2004 b. Revenue Management Under a General Discrete Choice Model of Consumer Behavior. *Management Science* 50 (1), 15–33.
- van Ryzin, G.J., 2005. Future of revenue management: Models of demand. *Journal of Revenue & Pricing Management* 4, 204–210.
- Van Slyke, R., Young, Y., 2000. Finite horizon stochastic knapsacks with applications to yield management. *Operations Research* 48, 155–172.

On-line Appendix to Structural Properties of Markov Modulated Revenue Management Problems

A. Monotonicity Proofs (Section 3)

In this appendix, we present the proofs of our results in Section 3 using induction. Before proving the concavity of $v_t(x, j)$ in x , we first state a lemma by Lautenbacher and Stidham (1999).

Lemma 1. *Suppose $g : \mathbb{Z}_+ \rightarrow \mathbb{R}$ is concave. Let $f : \mathbb{Z}_+ \rightarrow \mathbb{R}$ be defined by*

$$f(x) = \max_{\beta=0,1,\dots,m} \{\beta p + g(x - \beta)\}$$

for any given $p \geq 0$, and nonnegative integer $m \leq x$. Then, f is concave in $x \geq 0$.

Using Lemma 1, we next establish the concavity of the value function in x for each environment.

Proof of Theorem 1. Since $v_T(x, j)$ is zero for any x and j we have the concavity of $v_T(x, j)$ in x for any j . Suppose that $v_{t+1}(x, j)$ is a concave function of x for any environment j . We can use lemma 1 by taking g as $\sum_{k=1}^M p_{jk} v_{t+1}(x, k)$, p as $c(a)$, and m as $\min\{b, x\}$. Therefore,

$$\max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - u, k) + c(a)u \right\} \quad (12)$$

is concave in x for any product a , batch size $0 \leq b \leq B_a$. Since equation (1) is positive linear combination of (12) and $v_{t+1}(x, k)$, we have the concavity of $v_t(x, j)$ in x for any environment j . \square

Proof of Proposition 1. Since $v_t(x, j)$ is increasing in x , $\Delta v_t(x, j) \geq 0$. Also $\Delta v_T(x, j) = 0$, which implies that $\Delta v_T(x, j) \leq \Delta v_{T-1}(x, j)$. Suppose $\Delta v_{t+2}(x, j) \leq \Delta v_{t+1}(x, j)$ for any environment j and inventory level x . Consider the following inequality,

$$\begin{aligned} & \max_{u_1 \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+2}(x - u_1, k) + c(a) u_1 \right\} - \max_{u_2 \in U(b, x-1)} \left\{ \sum_{k=1}^M p_{jk} v_{t+2}(x - 1 - u_2, k) + c(a) u_2 \right\} \\ & \leq \\ & \max_{u_3 \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - u_3, k) + c(a) u_3 \right\} - \max_{u_4 \in U(b, x-1)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - 1 - u_4, k) + c(a) u_4 \right\} \end{aligned} \quad (13)$$

for any a , and batch size $0 \leq b \leq B_a$. It is sufficient to show that this inequality holds, in order to conclude that $\Delta v_{t+1}(x, j) \leq \Delta v_t(x, j)$, since the remaining terms in $\Delta v_t(x, j) - \Delta v_{t+1}(x, j)$ are clearly positive by using the induction hypothesis.

Let u_i^* be the optimal value of u_i in (13). We should note that $l_{t+1}^{a,j} \leq l_t^{a,j}$ for any product a and environment j . This can be easily seen by considering the induction hypothesis and (3). As a result, we have $u_3^* \leq u_1^*$. Also, we know that $u_1^* - u_2^*$ is either 1 or zero. Same reasoning is valid for $u_3^* - u_4^*$. If they are equal, then this is possible only either $u_1^* = u_2^* = 0$ or $u_1^* = u_2^* = b$.

Therefore, there are six cases we need to consider for the possible values of $u_1^*, u_2^*, u_3^*, u_4^*$.

Case	$(u_1^*, u_2^*, u_3^*, u_4^*)$	Inequality (13) simplifies to
1	$(0, 0, 0, 0)$	$\sum_{k=1}^M p_{jk} \Delta v_{t+2}(x, k) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k)$
2	$(y_2 + 1, y_2, 0, 0)$	$c(a) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k)$
3	$(b, b, 0, 0)$	$\sum_{k=1}^M p_{jk} \Delta v_{t+2}(x - b, k) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k)$
4	$(y_2 + 1, y_2, y_1 + 1, y_1)$	$c(a) \leq c(a)$
5	$(b, b, y_1 + 1, y_1)$	$\sum_{k=1}^M p_{jk} \Delta v_{t+2}(x - b, k) \leq c(a)$
6	(b, b, b, b)	$\sum_{k=1}^M p_{jk} \Delta v_{t+2}(x - b, k) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x - b, k)$

Here, y_1 and y_2 are integers such that $0 \leq y_1 \leq y_2 \leq b - 1$. Case 1 and 6 are true due to the induction hypothesis. Also case 4 is automatically true. In case 2, suppose that $\sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k) < c(a)$, then we should accept at least one customer when current inventory level is x at stage t but $u_3^* = 0$. Therefore, inequality in case 2 is true. In case 5, suppose that $\sum_{k=1}^M p_{jk} \Delta v_{t+2}(x - b, k) > c(a)$. Then, at time $t + 1$, accepted batch size is less than $b - 1$ when current inventory level is x . However $u_2^* = b$, which means that the inequality in case 5 is also true. In case 3, we have $c(a) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k)$ since $u_3^* = 0$. Also we have $\sum_{k=1}^M p_{jk} \Delta v_{t+2}(x - b, k) \leq c(a)$ since $u_2^* = b$. Note that, these inequalities can be shown by using the methodology used in case 2 and 5. Hence, we have the inequality of case 3. Consequently, Δv_t decreases in t . \square

Proof of Proposition 2. For $t = T$ we have the result trivially since $v_T(x, i) = 0$ for any inventory level x and environment i . Suppose $v_{t+1}(x, i) \leq v_{t+1}(x, j)$ for any $i \leq j$. We provide some definitions to make the proof clearer. Let

$$W(a, b, j) = \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - u, k) + c(a)u \right\}$$

and

$$S(a, j) = \sum_{b=1}^B q_{jab} W(a, b, j)$$

$$S(0, j) = \sum_{k=1}^M p_{jk} v_{t+1}(x, k)$$

for $a = 1, 2, \dots, N$. Then, we need to show the following

$$\sum_{a=0}^N r_{ia} S(a, i) \leq \sum_{a=0}^N r_{ja} S(a, j)$$

First of all, it is clear that $W(a, b, j)$ is nondecreasing in b , and a . Also, since \mathbf{P} is IFR, by the induction hypothesis we know that $W(a, b, j)$ is nondecreasing in j . Hence $S(a, j)$ is nondecreasing in j , because $\mathbf{Q}_{\mathbf{j}\mathbf{b}}^{(\mathbf{a})}$ is IFR. Also, by induction hypothesis we know $S(0, j)$ is nondecreasing in j . We also need to show that $S(a, j)$ is a nondecreasing function in a . Take $a \in \{1, 2, \dots, N\}$, since $W(a, b, j)$ is nondecreasing in a and $\mathbf{Q}_{\mathbf{a}\mathbf{b}}^{(\mathbf{j})}$ is IFR, we know that $S(a, j)$ is nondecreasing in the domain $\{1, 2, \dots, N\}$. It is also easy to show that $S(0, j) \leq S(1, j)$ hence $S(a, j)$ is nondecreasing in a . Since $S(a, i) \leq S(a, j)$ and $S(a, j)$ is a nondecreasing function in a , we have $v_t(x, i) \leq v_t(x, j)$ by using the IFR property of \mathbf{R} . \square

Proof of Proposition 3. Clearly, we have $\Delta v_T(x, i) = \Delta v_T(x, j) = 0$. Suppose $\Delta v_{t+1}(x, i) \leq \Delta v_{t+1}(x, j)$ for any $i \leq j$. Let

$$W(a, b, j) = \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - u, k) + c(a)u \right\} - \max_{u \in U(b, x-1)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - 1 - u, k) + c(a)u \right\}$$

for $a = 1, 2, \dots, N$. Also define,

$$S(a, j) = \sum_{b=1}^B q_{jab} W(a, b, j)$$

$$S(0, j) = \sum_{k=1}^M p_{jk} (v_{t+1}(x, k) - v_{t+1}(x - 1, k))$$

for $a = 1, 2, \dots, N$. After making these definitions, we need to show

$$\sum_{a=0}^M r_{ia} S(a, i) \leq \sum_{a=0}^M r_{ja} S(a, j)$$

for any environment $i \leq j$. First of all we will show $W(a, b, i) \leq W(a, b, j)$ for any $a \in \{1, 2, \dots, N\}$ and b . Let u_1^* be the optimal decision when current inventory is x and environment

is i , and let u_2^* be the optimal decision when the current inventory is $x - 1$ and environment is i (u_3^* and u_4^* are also defined in a similar fashion for environment j). Using the same reasoning that we used in the proof of the Proposition 1, we have the following relations and results by noting that $l_t^{a,j} \geq l_t^{a,i}$.

C	$(u_1^*, u_2^*, u_3^*, u_4^*)$	Inequality $W(a, b, i) \leq W(a, b, j)$ simplifies to
1	$(0, 0, 0, 0)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k)$
2	$(y_2 + 1, y_2, 0, 0)$	$c(a) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k)$
3	$(b, b, 0, 0)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x, k)$
4	$(y_2 + 1, y_2, y_1 + 1, y_1)$	$c(a) \leq c(a)$
5	$(b, b, y_1 + 1, y_1)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq c(a)$
6	(b, b, b, b)	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq \sum_{k=1}^M p_{jk} \Delta v_{t+1}(x - b, k)$

Here, y_1 and y_2 are integers such that $0 \leq y_1 \leq y_2 \leq b - 1$. Note that case 6 and case 1 are obviously true due to the induction hypothesis and the IFR property of \mathbf{P} . The remaining cases are true as explained in the proof of Proposition 1.

Secondly, we will show that $W(a, b, i)$ is nondecreasing in ordered quantity b for any i and $a \in \{1, 2, \dots, N\}$. Take $1 \leq b < B$. Let u_1^* be the optimal decision when current inventory is x and ordered quantity is b , and let u_2^* be the optimal decision when the current inventory is $x - 1$ and ordered quantity is b (u_3^* and u_4^* are also defined in a similar fashion for ordered quantity $b + 1$). We have 4 cases,

C	$(u_3^* u_4^*)$	Results	$W(a, b, i) < W(a, b + 1, i)$ reduces to
1	$(0, 0)$	$(u_1^*, u_2^*) = (0, 0)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) \leq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k)$
2	$(b + 1, b + 1)$	$x - l_t^{a,j} \geq b + 1$ $(u_1^*, u_2^*) = (b, b)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b - 1, k)$
3	$(b + 1, b)$	$x - l_t^{a,j} = b$ $(u_1^*, u_2^*) = (b, b)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq c(a)$
4	$(y, y - 1)$	$(u_1^*, u_2^*) = (y, y - 1)$	$c(a) \leq c(a)$

where $1 \leq y < b + 1$. Case 1 and 4 are obviously true since right-hand side and left-hand side are equal in both cases. Case 2 is also true since $\Delta v_{t+1}(x)$ is nonincreasing in x . In case 3, suppose $\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) > c(a)$. Then this fact contradicts with $l_t^{a,j} = x - b$.

Clearly, $S(0, i) \leq S(0, j)$ since \mathbf{P} is IFR. Since $W(a, b, i) \leq W(a, b, j)$ and $W(a, b, i)$ is nondecreasing in ordered quantity b for any i and $a \in \{1, 2, \dots, N\}$, by using the IFR property of $\mathbf{Q}_{\mathbf{jb}}^{(\mathbf{a})}$ we have

$$S(a, i) \leq S(a, j)$$

for any $a \in \{0, 1, \dots, N\}$. Now, it is sufficient to show that $S(a, j)$ is nondecreasing in a to show $\Delta v_t(x, i) \leq \Delta v_t(x, j)$ because we can use the IFR property of \mathbf{R} to conclude our result. Take $a_1, a_2 \in \{1, 2, \dots, N\}$ with $a_1 \leq a_2$. Since $c(a_1) \leq c(a_2)$, we know that $l_t^{a_1, j} \geq l_t^{a_2, j}$. Also, we have already shown that $W(a, b, j)$ is nondecreasing in b . It is sufficient to prove $W(a_1, b, j) \leq W(a_2, b, j)$, then we can use the IFR property of $\mathbf{Q}_{\mathbf{ab}}^{(\mathbf{j})}$ and $W(a, b, j)$'s being nondecreasing in b to conclude that $S(a_1, j) \leq S(a_2, j)$. Let u_1^* be the optimal decision when current inventory is x and product type is a_1 , and let u_2^* be the optimal decision when the current inventory is $x - 1$ and product type is a_1 (u_3^* and u_4^* are also defined in a similar fashion for product type a_2). Then, there are six cases for the values of $(u_1^*, u_2^*, u_3^*, u_4^*)$ as before.

C	$(u_1^*, u_2^*, u_3^*, u_4^*)$	Inequality $W(a_1, b, i) \leq W(a_2, b, i)$ simplifies to
1	$(0, 0, 0, 0)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) \leq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k)$
2	$(0, 0, y_1 + 1, y_1)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) \leq c(a_2)$
3	$(0, 0, b, b)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) \leq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k)$
4	$(y_2 + 1, y_2, y_1 + 1, y_1)$	$c(a_1) \leq c(a_2)$
5	$(y_2 + 1, y_2, b, b)$	$c(a_1) \leq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k)$
6	(b, b, b, b)	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k)$

Here, y_1 and y_2 are integers such that $0 \leq y_2 \leq y_1 \leq b - 1$. Note that in case 1 and 6, right hand sides and left hand sides are identical. Case 4 is true since $c(a_1) \leq c(a_2)$. In case 2, suppose $\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) > c(a_2)$ then we should not sell any product of type a_2 when current inventory level is x at time t , but $u_3^* = y_1 + 1 \geq 1$. Case 3 is also true since $\Delta v_{t+1}(x)$ is nonincreasing in x . In case 5, suppose $c(a_1) > \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k)$. Since $y_2 = u_2^* \leq b - 1$, we have $x - b + 1 \leq l_t^{a_1, j}$ and this result contradicts with our assumption. Therefore $S(a_1, i) \leq S(a_2, i)$ for $a_1, a_2 \in \{1, 2, \dots, N\}$. Also, we need to show

$S(0, i) \leq S(1, i)$. We have the following inequality since $W(1, b, i)$ is nondecreasing in b .

$$\begin{aligned} W(1, 1, i) &= \max_{u_1 \in \{0,1\}} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}(x - u, k) + c(1)u \right\} - \max_{u_2 \in \{0,1\}} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}(x - 1 - u, k) + c(1)u \right\} \\ &\leq S(1, i) = \sum_{b=1}^B q_{j1b} W(1, b, i) \end{aligned}$$

It is sufficient to show $W(1, 1, i) \geq S(0, i)$. Note that we have $u_1^* \geq u_2^*$. Therefore, we have 3 cases,

C	(u_1^*, u_2^*)	Inequality $W(1, 1, i) \geq S(0, i)$ reduces to
1	$(1, 1)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - 1, k) \geq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k)$
2	$(0, 0)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) \geq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k)$
3	$(1, 0)$	$c(1) \geq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k)$

Case 2 is obviously true, also case 1 is true since $\Delta v_{t+1}(x)$ is nonincreasing in x . In case 3 suppose $\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) > c(a)$ but this contradicts with $u_1^* = 1$.

Hence, $S(a_1, i) \leq S(a_2, i)$ for $a_1, a_2 \in \{0, 1, \dots, N\}$ and $S(a, i) \leq S(a, j)$ for $i \leq j$. Since \mathbf{R} is IFR, we have

$$\Delta v_t(x, i) = \sum_{a=0}^M r_{ia} S(a, i) \leq \sum_{a=0}^M r_{ja} S(a, j) = \Delta v_t(x, j)$$

□

B. Sensitivity Analysis Proofs (Section 4)

Proof of Proposition 4. Clearly $v_T^\epsilon(x, j) = v_T(x, j) = 0$, suppose $v_{t+1}^\epsilon(x, j) \geq v_{t+1}(x, j)$ for any environment j and inventory level x . For a given product a , and amount $b \in \{1, 2, \dots, B_a\}$, by using the induction hypothesis we know

$$\sum_{k=1}^M p_{jk} v_{t+1}(x - u, k) + c(a)u \leq \sum_{k=1}^M p_{jk} v_{t+1}^\epsilon(x - u, k) + c(a)u.$$

for any $0 \leq u \leq \min \{b, x\}$. Hence we have

$$\max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x - u, k) + c(a)u \right\} \leq \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}^\epsilon(x - u, k) + c(a)u \right\} \quad (14)$$

Consider any environment $j \neq i$. Then $v_t^\epsilon(x, j) \geq v_t(x, j)$ which is clear from inequality (14).

When we consider i as an environment, it is sufficient to show the following

$$\max_{u \in U(b_2, x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - u, k) + c(a_2)u \right\} \leq \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - u, k) + c(a)u \right\}$$

Note that, since $b_2 \leq b$ and $a_2 \leq a$ we know

$$\max_{u \in U(b_2, x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - u, k) + c(a_2)u \right\} \leq \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - u, k) + c(a_2)u \right\}$$

and

$$\max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}^\epsilon(x - u, k) + c(a_2)u \right\} \leq \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}^\epsilon(x - u, k) + c(a)u \right\}$$

Hence we have the result. \square

Proof of Proposition 5. Clearly, $\Delta v_T^\epsilon(x, j) = \Delta v_T(x, j) = 0$. Suppose $\Delta v_{t+1}^\epsilon(x, j) \geq \Delta v_{t+1}(x, j)$ for any environment j and inventory level x . Consider any environment $j \neq i$. Then as done in proposition 1, it can be shown that $\Delta v_t^\epsilon(x, j) \geq \Delta v_t(x, j)$. When we consider i as an environment, it is sufficient to show the following inequality,

$$\begin{aligned} & \max_{u \in U(b_2, x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - u, k) + c(a_2)u \right\} - \max_{u \in U(b_2, x-1)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - 1 - u, k) + c(a_2)u \right\} \\ & \leq \\ & \max_{u \in U(b, x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - u, k) + c(a)u \right\} - \max_{u \in U(b, x-1)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - 1 - u, k) + c(a)u \right\} \end{aligned}$$

Note that the right and left hand sides are similar to the definition of $W(a, b, j)$ in the proof of proposition 3, and W is nondecreasing in b . (None of the IFR properties are used

to show this, hence the same proof is also valid in here.) Therefore, it is sufficient to show,

$$\begin{aligned} \max_{u \in U(b,x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x-u, k) + c(a_2)u \right\} - \max_{u \in U(b,x-1)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x-1-u, k) + c(a_2)u \right\} \\ \leq \\ \max_{u \in U(b,x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x-u, k) + c(a)u \right\} - \max_{u \in U(b,x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x-1-u, k) + c(a)u \right\} \end{aligned}$$

Also we know that W is nondecreasing in a as done in the proof of proposition 3. (Again the IFR properties are not used to show this.) Hence, $\Delta v_t^\epsilon(x, j) \geq \Delta v_t(x, j)$ for any environment j and inventory level x . \square

Proof of Proposition 6. Clearly, $v_T^\epsilon(x, j) = v_T(x, j) = 0$. Suppose $v_{t+1}^\epsilon(x, j) \geq v_{t+1}(x, j)$. It is easy to verify $v_t^\epsilon(x, j) \geq v_t(x, j)$ when $j \neq i$ since components of \mathbf{P} remain same except the i^{th} row. Now, suppose that the current environment is i . Take any $a \in \{1, 2, \dots, N\}$ and $1 \leq b \leq B_a$. It is sufficient to show

$$\max_{u \in U(b,x)} \left\{ \sum_{k=1}^M p_{jk} v_{t+1}(x-u, k) + c(a)u \right\} \leq \max_{u \in U(b,x)} \left\{ \sum_{k=1}^M p_{jk}^\epsilon v_{t+1}^\epsilon(x-u, k) + c(a)u \right\}$$

By proposition 2, we know that $v_{t+1}(x-u, k)$ and also $v_{t+1}^\epsilon(x-u, k)$ are nondecreasing function in k . Therefore,

$$\sum_{k=1}^M p_{jk} v_{t+1}(x-u, k) \leq \sum_{k=1}^M p_{jk}^\epsilon v_{t+1}^\epsilon(x-u, k)$$

for any $u \in \{0, 1, \dots, \min\{b, x\}\}$ \square

Proof of Proposition 7. Clearly, $\Delta v_T^\epsilon(x, j) = \Delta v_T(x, j) = 0$. Suppose $\Delta v_{t+1}^\epsilon(x, j) \geq \Delta v_{t+1}(x, j)$ for any environment j and inventory level x . Consider any environment $j \neq i$, then it is easy to verify $\Delta v_t^\epsilon(x, j) \geq \Delta v_t(x, j)$ as done in the proof of proposition 1, because

j^{th} row of \mathbf{P} and \mathbf{P}^ϵ are identical. When the environment is i , it is sufficient to show

$$\begin{aligned} & \max_{u_1 \in U(b,x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}(x - u_1, k) + c(a) u_1 \right\} - \max_{u_2 \in U(b,x-1)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}(x - 1 - u_2, k) + c(a) u_2 \right\} \\ & \leq \\ & \max_{u_3 \in U(b,x)} \left\{ \sum_{k=1}^M p_{ik}^\epsilon v_{t+1}^\epsilon(x - u_3, k) + c(a) u_3 \right\} - \max_{u_4 \in U(b,x-1)} \left\{ \sum_{k=1}^M p_{ik}^\epsilon v_{t+1}^\epsilon(x - 1 - u_4, k) + c(a) u_4 \right\} \end{aligned}$$

for any $1 \leq a \leq N$, $1 \leq b \leq B_a$. Since $\Delta v_{t+1}^\epsilon(x, j) \geq \Delta v_{t+1}(x, j)$ and $\Delta v_t^\epsilon(x, k)$ is nondecreasing in k ,

$$\sum_{k=1}^M p_{ik}^\epsilon \Delta v_{t+1}^\epsilon(x, k) \geq \sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k)$$

Then $l_{t,\epsilon}^{a,j} \geq l_t^{a,j}$. Let u_i^* be the optimal value of u_i in the inequality above. As a result, we have $u_3^* \leq u_1^*$. Also, we know that $u_1^* - u_2^*$ is either 1 or zero. The same reasoning is valid for $u_3^* - u_4^*$. If they are equal, then this is possible only either $u_1^* = u_2^* = 0$ or $u_1^* = u_2^* = b$.

Therefore, there are six cases we need to consider for the possible values of $u_1^*, u_2^*, u_3^*, u_4^*$.

Case	$(u_1^*, u_2^*, u_3^*, u_4^*)$	Inequality simplifies to
1	$(0, 0, 0, 0)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) \leq \sum_{k=1}^M p_{ik}^\epsilon \Delta v_{t+1}^\epsilon(x, k)$
2	$(y_2 + 1, y_2, 0, 0)$	$c(a) \leq \sum_{k=1}^M p_{ik}^\epsilon \Delta v_{t+1}^\epsilon(x, k)$
3	$(b, b, 0, 0)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq \sum_{k=1}^M p_{ik}^\epsilon \Delta v_{t+1}^\epsilon(x, k)$
4	$(y_2 + 1, y_2, y_1 + 1, y_1)$	$c(a) \leq c(a)$
5	$(b, b, y_1 + 1, y_1)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq c(a)$
6	(b, b, b, b)	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq \sum_{k=1}^M p_{ik}^\epsilon \Delta v_{t+1}^\epsilon(x - b, k)$

Here, y_1 and y_2 are integers such that $0 \leq y_1 \leq y_2 \leq b - 1$. Case 1 is obviously true as shown above. In case 5, suppose $\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) > c(a)$, then accepted batch size should be less than b at time t when current inventory level is x , but this result contradicts with $u_2^* = b$. Similarly, in case 2, suppose $c(a) > \sum_{k=1}^M p_{ik}^\epsilon \Delta v_{t+1}^\epsilon(x, k)$. Then, we need to satisfy at least one of the requested amount at time t in modified system when current inventory

level is x , but we have $u_3^* = 0$. In case 3, we have $c(a) \leq \sum_{k=1}^M p_{ik}^\epsilon \Delta v_{t+1}^\epsilon(x, k)$ since $u_3^* = 0$.

Also we have $\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq c(a)$ since $u_2^* = b$. Note that, these inequalities can be shown by using the methodology used in case 2 and 5. Hence we have the inequality of case 3. Case 6 is also true by the induction assumption and the fact that $\Delta v_{t+1}^\epsilon(x, k)$ is nondecreasing function of k . Therefore, we have $\Delta v_t^\epsilon(x, j) \geq \Delta v_{t+1}^\epsilon(x, j)$ for any environment j , time t and inventory level x . \square

Proof of Proposition 8. We denote the modified function of price as $c_\epsilon(a)$ where $c(a)$ and $c_\epsilon(a)$ are identical expect $a = N$. At the terminal stage we trivially have $\Delta v_T^\epsilon(x, j) = \Delta v_T(x, j) = 0$ or any inventory level x and environment j . Suppose $\Delta v_{t+1}^\epsilon(x, j) \geq \Delta v_{t+1}(x, j)$ for $\forall x, j$. It is sufficient to show,

$$\begin{aligned} & \max_{u_1 \in U(b, x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}(x - u_1, k) + c(a) u_1 \right\} - \max_{u_2 \in U(b, x-1)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}(x - 1 - u_2, k) + c(a) u_2 \right\} \\ & \leq \\ & \max_{u_3 \in U(b, x)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - u_3, k) + c_\epsilon(a) u_3 \right\} - \max_{u_4 \in U(b, x-1)} \left\{ \sum_{k=1}^M p_{ik} v_{t+1}^\epsilon(x - 1 - u_4, k) + c_\epsilon(a) u_4 \right\} \end{aligned} \quad (15)$$

for any $1 \leq a \leq N$, $1 \leq b \leq B_a$. When $a \neq N$, then this inequality is true by a similar proof to that of proposition 1. In case of $a = N$, we know that threshold level is always 0, therefore, optimal quantity is $u_3^* = \min\{b, x\}$ when inventory level is x . Similarly $u_1^* = \min\{b, x\}$, $u_2^* = \min\{b, x - 1\} = u_4^*$.

Case	$(u_1^*, u_2^*, u_3^*, u_4^*)$	Inequality (15) simplifies to
1	(b, b, b, b)	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x - b, k) \leq \sum_{k=1}^M p_{ik} \Delta v_{t+1}^\epsilon(x - b, k)$
2	$(0, 0, 0, 0)$	$\sum_{k=1}^M p_{ik} \Delta v_{t+1}(x, k) \leq \sum_{k=1}^M p_{ik} \Delta v_{t+1}^\epsilon(x, k)$
3	$(x, x - 1, x, x - 1)$	$c(N) \leq c_\epsilon(N)$

Case 3 is true due to the increase in $c(N)$. Case 1 and 2 are also true by the induction hypothesis. \square

C. Sensitivity to Arrival Probability

In this section, we present results on the sensitivity of the optimal thresholds with respect to the arrival probability of fare classes 1 and 3. In particular, we vary the arrival probability of fare class 3 from 0.5 to 0.1, and the arrival probability of fare class 1 from 0.2 to 0.6 in environment 2. Threshold levels for fare classes 1 and 2 are given in Tables 7 and 8.

Table 7: Threshold levels of fare class 1

Time t	1	2	3	4	5	6	7	8	9	10
$l_t^{1,1}$	4	4	3	3	3	2	2	1	1	1
$l_{t,\epsilon}^{1,1}$	3	3	3	2	2	2	1	1	1	1
$l_t^{1,2}$	8	8	7	6	5	5	4	3	2	1
$l_{t,\epsilon}^{1,2}$	5	5	5	4	4	3	3	2	2	1

Table 8: Threshold levels of fare class 2

Time t	1	2	3	4	5	6	7	8	9	10
$l_t^{2,1}$	2	1	1	1	1	1	1	1	1	1
$l_{t,\epsilon}^{2,1}$	1	1	1	1	1	1	1	1	1	1
$l_t^{2,2}$	6	5	5	4	4	3	3	2	2	1
$l_{t,\epsilon}^{2,2}$	2	2	2	2	2	2	1	1	1	1

D. Sensitivity to Price

In this section, we present results on the sensitivity of the optimal thresholds with respect to the revenue of the third fare class, which is the most expensive one. In particular, we vary the revenue from 200 to 250. Threshold levels for fare classes 1 and 2 are given in Tables 9 and 10.

Table 9: Threshold levels of fare class 1

Time t	1	2	3	4	5	6	7	8	9	10
$l_t^{1,1}$	4	4	3	3	3	2	2	1	1	1
$l_{t,\epsilon}^{1,1}$	5	4	4	3	3	2	2	1	1	1
$l_t^{1,2}$	8	8	7	6	5	5	4	3	2	1
$l_{t,\epsilon}^{1,2}$	9	8	7	6	6	5	4	3	2	1

Table 10: Threshold levels of fare class 2

Time t	1	2	3	4	5	6	7	8	9	10
$l_t^{2,1}$	2	1	1	1	1	1	1	1	1	1
$l_{t,\epsilon}^{2,1}$	3	2	2	2	1	1	1	1	1	1
$l_t^{2,2}$	6	5	5	4	4	3	3	2	2	1
$l_{t,\epsilon}^{2,2}$	6	6	5	5	4	4	3	3	2	1

E. Proofs for the Pricing Model (Section 6)

Proof of Proposition 9. We know that $v_T(x, j) = 0$, therefore $\Delta v_T(x, j) = 0$. Assume that Δv_{t+1} is increasing function of x . Then

$$\begin{aligned}
\Delta v_t(x-1, j) - \Delta v_t(x, j) &= \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-1, k) - \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \\
&+ \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-1, k) \right\} \\
&- \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-2, k) \right\} \\
&- \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \right\} \\
&+ \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-1, k) \right\}
\end{aligned}$$

Let d_1 be optimal solution for $\max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_t(x, k) \right\}$ and d_2 be optimal solution for $\max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_t(x-2, k) \right\}$, then we have

$$\begin{aligned}
\Delta v_t(x-1, j) - \Delta v_t(x, j) &\geq \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-1, k) - \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \\
&\quad + d_1 p_j (d_1) - d_1 \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-1, k) \\
&\quad - d_2 p_j (d_2) + d_2 \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-2, k) \\
&\quad - d_1 p_j (d_1) + d_1 \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \\
&\quad + d_2 p_j (d_2) - d_2 \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-1, k)
\end{aligned}$$

After cancellations and rearranging the terms, we have

$$\begin{aligned}
\Delta v_t(x-1, j) - \Delta v_t(x, j) &\geq (1-d_1) \sum_{k=1}^N p_{jk} (\Delta v_{t+1}(x-1, k) - \Delta v_{t+1}(x, k)) \\
&\quad + d_2 \sum_{k=1}^N p_{jk} (\Delta v_{t+1}(x-2, k) - \Delta v_{t+1}(x-1, k))
\end{aligned}$$

Since $0 \leq d_1, d_2 \leq 1$, right hand side of the last inequality is greater than 0 by using the induction hypothesis. Hence we have the result. \square

Proof of Proposition 10. We know that $\Delta v_T(x, j) = 0$. Also, one can easily show that $v_t(x, j)$ is an increasing function of x by using induction. Therefore, $\Delta v_T(x, j) \leq \Delta v_{T-1}(x, j)$. Assume that $\Delta v_{t+1}(x, j) \leq \Delta v_t(x, j)$ for any inventory level x and environ-

ment j . Then

$$\begin{aligned}
\Delta v_{t-1}(x, j) - \Delta v_t(x, j) &= \sum_{k=1}^N p_{jk} \Delta v_t(x, k) - \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \\
&+ \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_t(x, k) \right\} \\
&- \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_t(x-1, k) \right\} \\
&- \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \right\} \\
&+ \max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-1, k) \right\}
\end{aligned}$$

Let d_2 be optimal solution for $\max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_t(x-1, k) \right\}$ and d_3 be the optimal solution for $\max_{0 \leq d \leq 1} \left\{ dp_j(d) - d \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \right\}$. Then

$$\begin{aligned}
\Delta v_{t-1}(x, j) - \Delta v_t(x, j) &\geq \sum_{k=1}^N p_{jk} (\Delta v_t(x, k) - \Delta v_{t+1}(x, k)) \\
&+ d_3 p_j(d_3) - d_3 \sum_{k=1}^N p_{jk} \Delta v_t(x, k) \\
&- d_2 p_j(d_2) + d_2 \sum_{k=1}^N p_{jk} \Delta v_t(x-1, k) \\
&- d_3 p_j(d_3) + d_3 \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x, k) \\
&+ d_2 p_j(d_2) - d_2 \sum_{k=1}^N p_{jk} \Delta v_{t+1}(x-1, k)
\end{aligned}$$

After cancellations and rearranging the terms we have

$$\begin{aligned}
\Delta v_{t-1}(x, j) - \Delta v_t(x, j) &\geq (1 - d_3) \sum_{k=1}^N p_{jk} (\Delta v_t(x, k) - \Delta v_{t+1}(x, k)) \\
&+ d_2 \sum_{k=1}^N p_{jk} (\Delta v_t(x-1, k) - \Delta v_{t+1}(x-1, k))
\end{aligned}$$

By using the induction hypothesis and $0 \leq d_2, d_3 \leq 1$, we have the result.

□