Appendices to "Dynamic Capacity Management with General Upgrading"

The appendices contain the supplemental materials to the main paper. Appendices A and B are devoted to the proofs of the main results in this paper. Appendix C provides the analysis of the multi-horizon model with replenishment in Section 6. A detailed description of the numerical studies can be found in Appendix D. Finally, Appendix E offers more details of the proofs that have been omitted in Appendix B for concision.

Appendix A: Preliminary

A.1. Notations

The following notations are used in this appendix to simplify our exposition. Consider a vector $\mathbf{Z}=(z_1,\cdots,z_N)\in\Re^N$, for $1\le i<j\le N$, we define

$$(\mathbf{Z})_i = z_i$$

$$(\mathbf{Z})_{i,\cdots,j} = (z_i, z_{i+1},\cdots,z_j)$$

$$\mathbf{Z}_{ij} = (z_1,\cdots,z_{i-1},z_i+1,z_{i+1},\cdots,z_{j-1},z_j-1,z_{j+1},\cdots,z_N).$$

Notice that the above notations are still valid for $\mathbf{Z}=(z_r,\cdots,z_k)$ $(1<r\le i\le j\le k<N)$ if we artificially set $\mathbf{Z}=(0,\cdots,0,z_r,\cdots,z_k,0,\cdots,0)\in\Re^N$.

For state vector $\mathbf{N}^t$, recall the effective state $\hat{\mathbf{N}}_r^t$ of classes $(1,\cdots,r)$ defined in Definition 2. If $r=N$, we use $\hat{\mathbf{N}}^t$ instead of $\hat{\mathbf{N}}_N^t$ to simplify our notation.

Moreover, for class i $(1\le i\le N)$ in period t $(1\le t\le T)$, we define

$$\partial_i^-\Theta^t(\mathbf{Z})=\frac{\partial}{\partial z_i^-}\Theta^t(\mathbf{Z}),\qquad \partial_i^+\Theta^t(\mathbf{Z})=\frac{\partial}{\partial z_i^+}\Theta^t(\mathbf{Z}).$$

Recall Δ_{ij}^{-+} and Δ_{ij}^{+-} $(1\le i<j\le N)$, we have

$$\Delta_{ij}^{-+}\Theta^t(\mathbf{Z})=\frac{\partial}{\partial z_i^-}\Theta^t(\mathbf{Z})-\frac{\partial}{\partial z_j^+}\Theta^t(\mathbf{Z}),\qquad \Delta_{ij}^{+-}\Theta^t(\mathbf{Z})=\frac{\partial}{\partial z_i^+}\Theta^t(\mathbf{Z})-\frac{\partial}{\partial z_j^-}\Theta^t(\mathbf{Z}).$$

Using the notations above, the protection level $p_{ij}=p$ in period t if and only if $\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t)\le\alpha_{ij}\le\Delta_{ij}^{-+}\Theta^t(\mathbf{N}^t)$ from (8), where $\mathbf{N}^t=(n_1^t,\cdots,n_{i-1}^t,p,0,\cdots,0,-p,n_{j+1}^t,\cdots,n_N^t)$.

In the essence of Δ_{ij}^{-+} and Δ_{ij}^{+-}, we define the marginal perturbation of class i and j (referred to as MP_{ij} hereafter) as $\Theta^t(\mathbf{Z}+\epsilon(\mathbf{e}_i-\mathbf{e}_j))-\Theta^t(\mathbf{Z})$, where $\epsilon\in\Re$ is a small number and $\mathbf{e}_s$ $(s=i,j)$ is the unit vector with 1 in position s.

A.2. Independence Property

Consider a state vector $\mathbf{N}^t=(n_1^t,\cdots,n_N^t)$ and its effective state $\hat{\mathbf{N}}_{i-1}^t=(\hat{n}_1^t,\cdots,\hat{n}_{i-1}^t,n_i^t,\cdots,n_N^t)$ in period t. In Lemma EC.4 and EC.5, we will show Θ^t has the following independence property if $(n_{i+1}^t,\cdots,n_{j-1}^t)\le 0$ and $n_j^t<0$:

1. In period t $(1 \le t \le T-1)$,

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t), \qquad \Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{-+}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t).$$

2. $\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t)$ and $\Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t)$ are independent of the values of $(n_j^t, \cdots, n_N^t)$.

Given the independence property of Θ^{t+1}, the protection levels in period t have a similar property. Specifically, consider two different state vectors $\mathbf{N} = (n_1, \cdots, n_N)^\intercal$ and $\mathbf{N}' = (n_1', \cdots, n_N')^\intercal$ with the same effective state for the first $i-1$ classes. If $(n_{i+1}, \cdots, n_{j-1}) = (n_{i+1}', \cdots, n_{j-1}') \le 0$ and $n_j = n_j' < 0$, then the protection level p_{ij} under state $\mathbf{N}$ is the same as that under $\mathbf{N}'$. Furthermore, the protection level p_{ij} under state $\mathbf{N}$ is independent of the values of $(n_j, \cdots, n_N)$. Hereafter, when speaking of the independence property, we do not distinguish between Θ^{t+1} and the protection levels in period t, since the proper interpretation is usually clear from the context.

Remark EC.1. Note that the independence property holds under the conditions $(n_{i+1}^t, \cdots, n_{j-1}^t) \le 0$ and $n_j^t < 0$. However, in the proofs of Lemma EC.4 and EC.5, we only need $n_j^t \le 0$ to prove the results of $\Delta_{ij}^{+-}\Theta^{t+1}$.

A.3. Foundation Results

Lemma 1 gives the condition of splitting the N-class general upgrading problem into subproblems, which reduces the complexity of the analysis.

Lemma 1 *Consider an N-class general upgrading problem with state $\mathbf{N}^t = (n_1^t, n_2^t, \cdots, n_N^t)^\intercal$ in period t. If $\sum_{s=k}^{i} n_s^t \le 0$ for all class $k \le i$, then the problem can be separated into two independent subproblems: an upper part consisting of classes $(1, \cdots, i)$, and a lower part consisting of classes $(i+1, \cdots, N)$.*

Proof. This result holds if none of the optimal policies would upgrade demand j by capacity k when there remains unmet demand i $(k < i < j)$ in the same period. For simplicity, we only prove the latter claim in the integer case. For any demand sample path $(\mathbf{D}^t, \cdots, \mathbf{D}^T)$, let $(\mathbf{Y}^t, \cdots, \mathbf{Y}^T)$ be the optimal decisions. We assume without loss of generality that $y_{i-1,j}^t = (\mathbf{Y}^t)_{i-1,j} \ge 1$ $(i < j)$ while there remains unmet demand i after period t.

We construct decisions $\bar{\mathbf{Y}}^s$ $(s = t, \cdots, T)$ that yield higher profit than the optimal decisions, which will be a contradiction. Let $\bar{\mathbf{Y}}^t$ be the same as $\mathbf{Y}^t$ except that $\bar{y}_{i-1,i}^t = y_{i-1,i}^t + 1$ and $\bar{y}_{i-1,j}^t = y_{i-1,j}^t - 1$. In the remaining periods s $(s = t+1, \cdots, T)$, we apply allocation decision $\bar{\mathbf{Y}}^s = \mathbf{Y}^s$ whenever $\mathbf{Y}^s$ is feasible. If the optimal decisions are feasible in periods $t+1$ to T, the profit increase by using $\bar{\mathbf{Y}}^s$ $(t \le s \le T)$ instead of the optimal decisions is $\alpha_{i-1,i} - \alpha_{i-1,j} + (T-t+1)(g_i - g_j) > 0$, which is a contradiction.

Otherwise, let l $(t+1 \le l \le T)$ be the first period that $\mathbf{Y}^l$ is not feasible. From our construction, it is clear that there exists $y_{ki}^l \ge 1$ $(k < i)$ in $\mathbf{Y}^l$ that is not feasible after applying $\bar{\mathbf{Y}}^s$ $(s = t, \cdots, l-1)$.

Let $\bar{\mathbf{Y}}^l$ be the same as $\mathbf{Y}^l$ except that $\bar{y}^l_{ki} = y^l_{ki} - 1$ and $\bar{y}^l_{kj} = y^l_{kj} + 1$. Since the states after applying $\bar{\mathbf{Y}}^s$ $(s = t, \cdots, l)$ are the same as that for $\mathbf{Y}^s$ $(s = t, \cdots, l)$, $\bar{\mathbf{Y}}^s = \mathbf{Y}^s$ $(s = l+1, \cdots, T)$ are feasible in the remaining periods. Thus, the profit increase by using $\bar{\mathbf{Y}}^s$ $(t \le s \le T)$ instead of the optimal ones is $(l-t)(g_i - g_j) > 0$, which contradicts the optimality assumption.

This concludes our proof. □

Lemmas EC.1 and EC.2 illustrate the bounds of the profit differences under different states.

Lemma EC.1. *Consider a state vector* $\mathbf{N} = (n_1, \cdots, n_N)$ *with* $n_i \ge 0$ *and* $n_j \ge 0$ $(1 \le i < j \le N)$. *Then,*

$$\partial^+_i \Theta^t(\mathbf{N}) - \partial^+_j \Theta^t(\mathbf{N}) \ge u_j - u_i \tag{EC.1}$$

and

$$\partial^-_i \Theta^t(\mathbf{N}) - \partial^-_j \Theta^t(\mathbf{N}) \ge u_j - u_i \qquad \text{if } n_i > 0 \text{ and } n_j > 0. \tag{EC.2}$$

Proof. We use the sample path argument to prove (EC.1). For each demand sample path, it is sufficient to prove

$$\Theta^t(\mathbf{N} + \epsilon\mathbf{e}_i) - \Theta^t(\mathbf{N} + \epsilon\mathbf{e}_j) \ge \epsilon(u_j - u_i), \tag{EC.3}$$

where $\epsilon > 0$, $\mathbf{e}_s$ $(s = i, j)$ is the unit vector with 1 in position s. The same argument can be applied to (EC.2).

Given a demand sample path $(\mathbf{D}^t, \cdots, \mathbf{D}^T)$, let $(\mathbf{Y}^t, \cdots, \mathbf{Y}^T)$ be the corresponding optimal solutions in period t to T under initial state $\mathbf{N} + \epsilon\mathbf{e}_j$ in period t. For initial state $\mathbf{N} + \epsilon\mathbf{e}_i$, we sequentially construct solutions $(\bar{\mathbf{Y}}^t, \cdots, \bar{\mathbf{Y}}^T)$ based on $(\mathbf{Y}^t, \cdots, \mathbf{Y}^T)$ from period t to T. Specifically, $\bar{\mathbf{Y}}^l = \mathbf{Y}^l$ in period l $(t \le l \le T)$ if $\mathbf{Y}^l$ is feasible, and we write $\epsilon_l = 0$. Otherwise, if $\mathbf{Y}^l$ is not feasible, from the assumption of the initial states, the total demands which are satisfied by capacity j in $\mathbf{Y}^l$ is greater than the existing capacity j with initial state $\mathbf{N} + \epsilon\mathbf{e}_i$, and we denote the difference as ϵ_l $(0 < \epsilon_l \le \epsilon_1)$. To construct a feasible solution $\bar{\mathbf{Y}}^l$, we use capacity i to satisfy demands which cannot be fulfilled by capacity j. By applying such $(\bar{\mathbf{Y}}^t, \cdots, \bar{\mathbf{Y}}^T)$, the unmet demands in periods t to T are the same for both initial states, and $\sum_{l=t}^T \epsilon_l \le \epsilon$.

Note that $\alpha_{si} - \alpha_{sj} = u_j - u_i < 0$ for any class s $(s \ge j)$, and unmet demand vectors in period t to T are the same for both initial states. Since $(\bar{\mathbf{Y}}^t, \cdots, \bar{\mathbf{Y}}^T)$ are feasible solutions to the general upgrading problem with initial state $\mathbf{N} + \epsilon\mathbf{e}_i$, we have

$$\Theta^t(\mathbf{N} + \epsilon\mathbf{e}_i) - \Theta^t(\mathbf{N} + \epsilon\mathbf{e}_j) \ge (u_j - u_i)\sum_{l=t}^T \epsilon_l \ge \epsilon(u_j - u_i),$$

which completes the proof. □

Lemma EC.2. *Consider a state vector* $\mathbf{N} = (n_1, \cdots, n_N)$ *with* $n_i \le 0$ *and* $n_j \le 0$ $(1 \le i < j \le N)$. *Then,*

$$\partial_i^+ \Theta^t(\mathbf{N}) - \partial_j^+ \Theta^t(\mathbf{N}) \ge r_j - r_i \qquad \text{if } n_i < 0 \text{ and } n_j < 0$$

and

$$\partial_i^- \Theta^t(\mathbf{N}) - \partial_j^- \Theta^t(\mathbf{N}) \ge r_j - r_i.$$

Proof. It is similar to the proof of Lemma EC.1. □

Appendix B: Proofs of the Main Results

This section presents the proofs of the main results in the paper. The proofs of some intermediate results are lengthy and therefore presented in Appendix E, including Lemmas EC.6 to EC.10 and Propositions EC.1 to EC.3.

In §B.1, we prove the desired properties in period T. §B.2 considers a general period t by following the similar logic for period T. §B.3 completes the optimality proof. §B.4 proves two properties of the protection levels.

B.1. Final Period T

Lemma EC.3. *The PSR policy solves the general upgrading problem* (2) *in period* T *with all protection levels being* 0.

Proof. Note that $\Theta^{T+1} \equiv 0$ and the solution $\mathbf{Y}^T$ generated by the PSR is a Monge sequence which solves the general upgrading problem in period T (see Bassok et al. 1999). □

We follow the notations in the main paper. Recall the state vector $\mathbf{N}^t = (n_1^t, \cdots, , n_N^t)$ in period t, and $\hat{\mathbf{N}}_{i-1}^t = (\hat{n}_1^t, \cdots, \hat{n}_{i-1}^t, n_i^t, \cdots, n_N^t)$, where $(\hat{n}_1^t, \cdots, \hat{n}_{i-1}^t)$ is the effective state of $(n_1^t, \cdots, n_{i-1}^t)$. Then, Lemma EC.4 shows the independence property of Θ^T.

Lemma EC.4. *Consider an* N*-class general upgrading problem in period* $T-1$ *with state vector* $\mathbf{N}^{T-1}$, *where* $(n_{i+1}^{T-1}, \cdots, n_{j-1}^{T-1}) \le 0$ *and* $n_j^{T-1} < 0$. *Then,*

$$\Delta_{ij}^{+-} \Theta^T(\mathbf{N}^{T-1}) = \Delta_{ij}^{+-} \Theta^T(\hat{\mathbf{N}}_{i-1}^{T-1}), \qquad \Delta_{ij}^{-+} \Theta^T(\mathbf{N}^{T-1}) = \Delta_{ij}^{-+} \Theta^T(\hat{\mathbf{N}}_{i-1}^{T-1}). \tag{EC.4}$$

In addition, they are independent of the values of $(n_j^{T-1}, \cdots, n_N^{T-1})$.

Proof. For any $t = 1, \cdots, T$, given $\mathbf{D}^t = (d_1, \cdots, d_N)$ as realized demand in period t, we have

$$\Delta_{ij}^{+-} \Theta^t(\mathbf{N}^{t-1}) = \Delta_{ij}^{+-} \mathbb{E}\left\{\Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)\right\} = \mathbb{E}\left\{\Delta_{ij}^{+-} \Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)\right\} \tag{EC.5}$$

and

$$\Delta_{ij}^{-+} \Theta^t(\mathbf{N}^{t-1}) = \Delta_{ij}^{-+} \mathbb{E}\left\{\Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)\right\} = \mathbb{E}\left\{\Delta_{ij}^{-+} \Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)\right\}. \tag{EC.6}$$

Both the continuity of $\Theta^t(\mathbf{N}^{t-1}|\mathbf{D}^t)$ and the existence of its left and right derivatives (see Rockafellar 1996) assure the last equality in (EC.5-EC.6) (see Zorich 2004, P.409).

We focus on Δ_{ij}^{+-} in (EC.4) since the same method applies to Δ_{ij}^{-+}. For any demand realization $\mathbf{D}^T = (d_1, \cdots, d_N)$ in period T, we next show

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}^{T-1}|\mathbf{D}^T) = \Delta_{ij}^{+-}\Theta^T(\hat{\mathbf{N}}_{i-1}^{T-1}|\mathbf{D}^T), \tag{EC.7}$$

and it is independent of the values of $(n_j^{T-1}, \cdots, n_N^{T-1})$.

For any $\mathbf{D}^T$, without loss of generality, we assume classes $(1, \cdots, N)$ can not be separated based on $\mathbf{N}^{T-1} - \mathbf{D}^T$. Otherwise, from Lemma 1, we can consider independent subproblems instead. With this assumption, classes $(1, \cdots, N)$ are also not separable based on $\hat{\mathbf{N}}_{i-1}^{T-1} - \mathbf{D}^T$ by Proposition EC.1 in Appendix E.

To solve the N-class general upgrading problem in period T, we first solve subproblems $(1, \cdots, i-1)$ with initial state $(\mathbf{N}^{T-1})_{1,\cdots,i-1}$ and $(\hat{\mathbf{N}}_{i-1}^{T-1})_{1,\cdots,i-1}$ by the PSR. Then, we use the PSR to solve the subproblem $(1, \cdots, N)$, where the initial states of classes $(1, \cdots, i-1)$ are the states after solving the subproblem $(1, \cdots, i-1)$ by the PSR.

Since the upgrading problem in period T is a transportation problem, given the special cost structure, the optimal allocation decisions in subproblem $(1, \cdots, i-1)$ are independent from classes $(i, \cdots, N)$. Particularly, the optimal decisions within classes $(1, \cdots, i-1)$ remain unchanged with respect to MP_{ij}. Moreover, from Proposition EC.2, the result of applying the PSR to subproblem $(1, \cdots, i-1)$ with initial state $(\mathbf{N}^{T-1})_{1,\cdots,i-1}$ is the same as that with initial state $(\hat{\mathbf{N}}_{i-1}^{T-1})_{1,\cdots,i-1}$. In other words, the initial states in subproblem $(1, \cdots, N)$ are the same for both initial states $(\mathbf{N}^{T-1}, \mathbf{D}^T)$ and $(\hat{\mathbf{N}}_{i-1}^{T-1}, \mathbf{D}^T)$. Thus, (EC.7) is true. In addition, $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}^{T-1}|\mathbf{D}^T)$ is independent of the values of $(n_j^{T-1}, \cdots, n_N^{T-1})$ from Lemma EC.7. This completes the proof. □

B.2. Earlier Periods

Lemma EC.5 proves the independence property of Θ^{t+1} by backward induction.

LEMMA EC.5. *Consider an N-class general upgrading problem in period t with state vector $\mathbf{N}^t$, where $(n_{i+1}^t, \cdots, n_{j-1}^t) \le 0$ and $n_j^t < 0$. If the PSR policy solves the general upgrading problem in period $t+1$ and the independence property holds for Θ^{t+2}, then,*

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t), \qquad \Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{-+}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t). \tag{EC.8}$$

In addition, $\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t)$ and $\Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t)$ are independent of the values of $(n_j^t, \cdots, n_N^t)$.

Proof. As discussed in the proof of Lemma EC.4, Δ_{ij}^{+-} and Δ_{ij}^{-+} in (EC.8) are well-defined. We prove the equality regarding Δ_{ij}^{+-} in (EC.8) and the corresponding independence property for any demand realization $\mathbf{D}^{t+1} = (d_1, \cdots, d_N)$ in period $t+1$. From Lemma 1, we can assume classes $(1, \cdots, N)$ are not separable under $\mathbf{N}^t - \mathbf{D}^{t+1}$, which is also true under $\hat{\mathbf{N}}^t_{i-1} - \mathbf{D}^{t+1}$ by Proposition EC.1.

Splitting the N-class general upgrading problem into subproblems: $(1, \cdots, i-1)$, $(1, \cdots, j)$ and $(1, \cdots, N)$, we start with the subproblem $(1, \cdots, i-1)$.

1. Because the protection levels within classes $(1, \cdots, i-1)$ in period $t+1$ are defined by Θ^{t+2}, which satisfies the independence property by assumption, the allocation decisions within classes $(1, \cdots, i-1)$ in period $t+1$ remain unchanged with respect to MP_{ij}. Let $\mathbf{N}'_{i-1}$ be the outcome of applying the PSR policy to subproblem $(1, \cdots, i-1)$ with states $((\mathbf{N}^t)_{1,\cdots,i-1}, (\mathbf{D}^{t+1})_{1,\cdots,i-1})$. Denote k $(1 \le k \le i-1)$ as the highest class such that $(\mathbf{N}'_{i-1})_{k,\cdots,i-1} \ge 0$ and $(\mathbf{N}'_{i-1})_k > 0$. Since the PSR is optimal in period $t+1$ by assumption, we only need to consider upgrading decisions among classes $(k, \cdots, N)$ in the rest of the subproblems. Similarly, we can define $\hat{\mathbf{N}}'_{i-1}$ and $\hat{k}$ for subproblem $(1, \cdots, i-1)$ with states $((\hat{\mathbf{N}}^t_{i-1})_{1,\cdots,i-1}, (\mathbf{D}^{t+1})_{1,\cdots,i-1})$. From Proposition EC.3, we know that $\hat{k} = k$ and $(\hat{\mathbf{N}}'_{i-1})_{k,\cdots,i-1} = (\mathbf{N}'_{i-1})_{k,\cdots,i-1}$. In other words, after solving subproblem $(1, \cdots, i-1)$, the initial state of classes $(k, \cdots, N)$ are the same for both $\mathbf{N}^t$ and $\hat{\mathbf{N}}^t_{i-1}$. Notice that we assume both k and $\hat{k}$ exist; otherwise, both k and $\hat{k}$ do not exist from Proposition EC.3, which means that considering upgrading decisions in classes $(i, \cdots, N)$ is sufficient, which is a simpler case.
2. From the definition of the protection levels, although there is no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$, the states of classes $(1, \cdots, k-1)$ can still affect the protection levels within classes $(k, \cdots, N)$ in period $t+1$. Fortunately, the effective state of $(\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1}$ is the same as that of $(\mathbf{N}'_{i-1})_{1,\cdots,k-1}$ by Proposition EC.2. From the independence property assumption of Θ^{t+2}, the protection levels within classes $(k, \cdots, N)$ are the same for both initial states.

To summarize, for initial states $\mathbf{N}^t$ and $\hat{\mathbf{N}}^t_{i-1}$, the capacities of classes $(k, \cdots, i-1)$ after solving subproblem $(1, \cdots, i-1)$, which can upgrade the demands in classes $(i, \cdots, N)$, are the same. Moreover, the protection levels within classes $(k, \cdots, N)$ are also the same. Therefore, we only analyze the allocation decisions within classes $(k, \cdots, N)$ under initial state $\mathbf{N}^t$, which can again be split into subproblems $(k, \cdots, j)$ and $(k, \cdots, N)$.

Apply the PSR to subproblem $(k, \cdots, j)$ with state $(\mathbf{N}_j, (0, \cdots, 0, (\mathbf{D}^{t+1})_{i,\cdots,j}))$, where $\mathbf{N}_j = ((\mathbf{N}'_{i-1})_{k,\cdots,i-1}, (\mathbf{N}^t)_{i,\cdots,j})$, and let $\mathbf{N}'_j$ be the resulting states of classes $(k, \cdots, j)$ after applying $\mathbf{Y}_j$, which are the optimal allocation decisions within classes $(k, \cdots, j)$. Since $(\mathbf{N}^t)_{i+1,\cdots,j} \le 0$, the protection levels used in subproblem $(k, \cdots, j)$, which determine the upgrades from classes $(k, \cdots, i)$, only depend on $(\mathbf{N}'_{i-1})_{k,\cdots,i-1}$ by the independence property assumption of Θ^{t+2}. We consider two cases based on whether there is unmet demand j in $\mathbf{N}'_j$:

1. $(\mathbf{N}'_j)_j = 0$: Define h $(k \le h \le i)$ as the class which satisfies the last unit of demand j when the PSR solves subproblem $(k, \cdots, j)$. In fact,

$$h = \begin{cases} r, & \text{if } r < i \text{ and } \sum_{s=r+1}^{i-1} (\mathbf{N}'_{i-1})_s \le -\sum_{s=i}^{j} ((\mathbf{N}^t)_s - d_s) < \sum_{s=r}^{i-1} (\mathbf{N}'_{i-1})_s \\ i, & \text{if } \sum_{s=i}^{j} ((\mathbf{N}^t)_s - d_s) > 0. \end{cases}$$

In this case, $\mathbf{N}'_j$ is the same as the result of applying the greedy upgrading to subproblem $(k, \cdots, j)$, i.e., $\mathbf{N}'_j = \hat{\mathbf{N}}_j$, where $\hat{\mathbf{N}}_j$ is the effective state of $\mathbf{N}_j$. Specifically,

$$(\hat{\mathbf{N}}_j)_l = \begin{cases} (\mathbf{N}'_{i-1})_l, & \text{if } k \le l < h \\ \sum_{s=h}^{i-1} (\mathbf{N}'_{i-1})_s + \sum_{s=i}^{j} ((\mathbf{N}^t)_s - d_s), & \text{if } l = h < i \\ \sum_{s=i}^{j} ((\mathbf{N}^t)_s - d_s), & \text{if } l = h = i \\ 0, & \text{otherwise,} \end{cases} \tag{EC.9}$$

for class l $(k \le l \le j)$. Note that class h $(k \le h \le i)$ must exist since classes $(1, \cdots, N)$ are not separable, and h and $\hat{\mathbf{N}}_j$ remain the same with respect to MP_{ij}. Furthermore, from the discussion of $\mathbf{N}'_j$, we can see that $\mathbf{Y}_j$ is the same as optimal allocation decisions given initial state $(\mathbf{N}_j, (0, \cdots, 0, (\mathbf{D}^{t+1})_{i,\cdots,j}))$ in period T where the protection levels are zero. Hence,

$$\begin{aligned} \Theta^{t+1}(\mathbf{N}^t | \mathbf{D}^{t+1}) =& \Theta^T \left((\mathbf{N}^t)_{1,\cdots,i-1} - \mathbf{N}'_{i-1} | (\mathbf{D}^{t+1})_{1,\cdots,i-1} \right) + \Theta^T \left(\mathbf{N}_j | (0, \cdots, 0, (\mathbf{D}^{t+1})_{i,\cdots,j}) \right) \\ &+ \Theta^{t+1} \Big(((\mathbf{N}'_{i-1})_{1,\cdots,k-1}, \hat{\mathbf{N}}_j, (\mathbf{N}^t)_{j+1,\cdots,N}) | (0, \cdots, 0, (\mathbf{D}^{t+1})_{j+1,\cdots,N}) \Big), \end{aligned} \tag{EC.10}$$

where the first two terms are the corresponding revenues of subproblems $(1, \cdots, i-1)$ and $(k, \cdots, j)$, and the last term is the sum of the current revenue of subproblem $(k, \cdots, j)$ and the expected value in the remaining periods. Thus,

$$\Delta_{ij}^{+-} \Theta^{t+1}(\mathbf{N}^t | \mathbf{D}^{t+1}) = \Delta_{ij}^{+-} \Theta^T \left(\mathbf{N}_j | (0, \cdots, 0, (\mathbf{D}^{t+1})_{i,\cdots,j}) \right), \tag{EC.11}$$

which is clearly independent of $(n_{j+1}^t, \cdots, n_N^t)$. Also, (EC.11) is independent of n_j^t by Lemma EC.7. Note that the first term in (EC.10) has been omitted from (EC.11) since the allocation decisions in subproblem $(1, \cdots, i-1)$ remain unchanged with respect to MP_{ij}. Moreover, the last term in (EC.10) has also been dropped from (EC.11) because its initial states remain the same with respect to MP_{ij}.

Similarly, for initial state $\hat{\mathbf{N}}_{i-1}^t$, we have

$$\Delta_{ij}^{+-} \Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t | \mathbf{D}^{t+1}) = \Delta_{ij}^{+-} \Theta^T \left(\mathbf{N}_j | (0, \cdots, 0, (\mathbf{D}^{t+1})_{i,\cdots,j}) \right)$$

since the allocation decisions in subproblem $(k, \cdots, j)$ are the same for both initial state $\mathbf{N}^t$ and $\hat{\mathbf{N}}_{i-1}^t$. Therefore, we have

$$\Delta_{ij}^{+-} \Theta^{t+1}(\mathbf{N}^t | \mathbf{D}^{t+1}) = \Delta_{ij}^{+-} \Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t | \mathbf{D}^{t+1}),$$

which is independent of the values of $(n_j^t, \cdots, n_N^t)$ by (EC.11);

2. $(\mathbf{N}'_j)_j < 0$: Since the PSR is optimal in period $t+1$, there is no upgrade between classes $(k,\cdots,j)$ and $(j+1,\cdots,N)$. By the definition of the effective state, $\hat{\mathbf{N}}_j$ in (EC.9), which remains unchanged with respect to MP_{ij}, is the effective state of $\mathbf{N}'_j$. Thus, the allocation decisions within classes $(j+1,\cdots,N)$ stay the same with respect to MP_{ij} by the independence property assumption of Θ^{t+2}, and we denote $\mathbf{N}'_{j+}$ as the result of applying the PSR to classes $(j+1,\cdots,N)$. Therefore, we have

$$\begin{aligned}
&\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1})\\
=&\Theta^T\left((\mathbf{N}^t)_{1,\cdots,i-1}-\mathbf{N}'_{i-1}|(\mathbf{D}^{t+1})_{1,\cdots,i-1}\right)+\Theta^T\left(\mathbf{N}_j-\mathbf{N}'_j|(0,\cdots,0,(\mathbf{D}^{t+1})_{k,\cdots,j})\right)\\
&+\Theta^T\left((\mathbf{N}^t)_{j+1,\cdots,N}-\mathbf{N}'_{j+}|(0,\cdots,0,(\mathbf{D}^{t+1})_{j+1,\cdots,N})\right)+\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right),
\end{aligned} \tag{EC.12}$$

where the first three terms are the corresponding revenues of subproblems $(1,\cdots,i-1)$, $(k,\cdots,j)$, and $(j,\cdots,N)$, and the last term is the expected revenue-to-go function. As we discussed earlier, we have

$$\begin{aligned}
&\Delta^{+-}_{ij}\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1})\\
=&\Delta^{+-}_{ij}\Theta^T\left(\mathbf{N}_j-\mathbf{N}'_j|(0,\cdots,0,(\mathbf{D}^{t+1})_{k,\cdots,j})\right)+\frac{\partial}{\partial n_i^+}\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)\\
&-\frac{\partial}{\partial n_j^-}\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right),
\end{aligned} \tag{EC.13}$$

where the first term is independent of $(n^t_{j+1},\cdots,n^t_N)$ by construction. Moreover, recall that the protection levels used in subproblem $(k,\cdots,j)$ only depend on $(\mathbf{N}'_{i-1})_{k,\cdots,i-1}$, and demand j is not fully satisfied in this case, thus the allocation decisions $\mathbf{Y}_j$ as well as $\mathbf{N}_j-\mathbf{N}'_j$, which is the capacity used in subproblem $(k,\cdots,j)$, do not depend on n^t_j. Hence, the first term in (EC.13) is also independent of n^t_j. Similarly, for initial state $\hat{\mathbf{N}}^t_{i-1}$, we have

$$\begin{aligned}
&\Delta^{+-}_{ij}\Theta^{t+1}(\hat{\mathbf{N}}^t_{i-1}|\mathbf{D}^{t+1})\\
=&\Delta^{+-}_{ij}\Theta^T\left(\mathbf{N}_j-\mathbf{N}'_j|(0,\cdots,0,(\mathbf{D}^{t+1})_{k,\cdots,j})\right)+\frac{\partial}{\partial n_i^+}\Theta^{t+2}\left((\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)\\
&-\frac{\partial}{\partial n_j^-}\Theta^{t+2}\left((\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right).
\end{aligned} \tag{EC.14}$$

To complete the proof, from (EC.13) and (EC.14), we use the induction assumption of Θ^{t+2} to show

$$\begin{aligned}
&\frac{\partial}{\partial n_i^+}\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)-\frac{\partial}{\partial n_j^-}\Theta^{t+2}\left((\mathbf{N}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)\\
=&\frac{\partial}{\partial n_i^+}\Theta^{t+2}\left((\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right)-\frac{\partial}{\partial n_j^-}\Theta^{t+2}\left((\hat{\mathbf{N}}'_{i-1})_{1,\cdots,k-1},\mathbf{N}'_j,\mathbf{N}'_{j+}\right),
\end{aligned} \tag{EC.15}$$

which is independent of $(n^t_j,\cdots,n^t_N)$. First of all, since there is no upgrade between classes $(1,\cdots,k-1)$ and $(k,\cdots,N)$ in period $t+1$, and the PSR sequentially satisfies demands in each class, the marginal change of n_i^t only affects the state of a single class in $\mathbf{N}_j'$, which is the same for both initial states $\mathbf{N}^t$ and $\hat{\mathbf{N}}_{i-1}^t$. Denote such a class as r, then $k \le r \le j$. Given $(\mathbf{N}_{i-1}')_{1,\cdots,k-1}$ and $(\hat{\mathbf{N}}_{i-1}')_{1,\cdots,k-1}$ have the same effective state from the previous argument, to apply the induction assumption, we only need to show $(\mathbf{N}_j')_{r+1,\cdots,j} \le 0$ where $(\mathbf{N}_j')_j < 0$ by assumption. Suppose to the contrary that $(\mathbf{N}_j')_l > 0$ for class l $(r < l < j)$. Note that initial states $(\mathbf{N}^t)_{i+1,\cdots,j} \le 0$, thus class $l \le i$. Since the demands in classes $(i,\cdots,j)$ should be satisfied by class l prior to class r by the PSR, given $(\mathbf{N}_j')_l > 0$, there is no upgrade between classes $(k,\cdots,l-1)$ and $(l,\cdots,j)$, i.e., the marginal change of n_i^t should not affect the state of class r, which is a contradiction. Hence, by applying the induction assumption to (EC.15), we have

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t|\mathbf{D}^{t+1}) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t|\mathbf{D}^{t+1}),$$

which is independent of the values of $(n_j^t,\cdots,n_N^t)$. This concludes the proof.

□

B.3. Optimality

Proposition 2 *Consider an N-class general upgrading problem in period t $(1 \le t \le T)$ with state vector $\mathbf{N}^t$, where $(n_{i+1}^t,\cdots,n_{j-1}^t) \le 0$ and $n_j^t < 0$.*

1. *We have*

$$\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{+-}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t), \qquad \Delta_{ij}^{-+}\Theta^{t+1}(\mathbf{N}^t) = \Delta_{ij}^{-+}\Theta^{t+1}(\hat{\mathbf{N}}_{i-1}^t),$$

both of which are independent of the values of $(n_j^t,\cdots,n_N^t)$.

2. *The PSR policy solves the general upgrading problem in period t.*

Proof. In the proof, we show the two properties in Proposition 2 can be preserved under backward induction. The proof of period T is given in the end of this proof.

Suppose they are true for Θ^{t+1}, we verify the two properties for Θ^t.

2. Optimality of the PSR policy

Consider initial state $\mathbf{X}^t = (x_1,\cdots,x_N)$ and $\tilde{\mathbf{D}}^t = (\tilde{d}_1,\cdots,\tilde{d}_N)$, for any demand realization $\mathbf{D}^t = (d_1,\cdots,d_N)$ in period t, we next verify $\mathbf{Y}^t = (y_{ij})_{N\times N}$ derived by the PSR are optimal in period t. First, from the discussion in the main paper, $y_{kk} = \min(d_k + \tilde{d}_k, x_k)$ $(1 \le k \le N)$ in the PSR is optimal.

For upgrading decisions y_{ij} $(i > j)$ in $\mathbf{Y}^t$, we consider an equivalent representation of the general upgrading problem in (2). Let $\mathbf{Z} = (z_1, \cdots, z_N)^\top = \mathbf{Y}^t\mathbf{1}$, the optimal solution $\mathbf{W} = (w_{ij})_{N\times N}$ in the following linear program is the same as $\mathbf{Y}^t = (y_{ij})_{N\times N}$ in period t:

$$\begin{aligned}\max_{\mathbf{W}\geq 0} \quad & \sum_{1\leq i\leq j\leq N} \alpha_{ij} w_{ij} \\ \text{s.t.} \quad & \sum_j w_{ij} \leq z_i, \quad i = 1, 2, \cdots, N, \\ & \sum_i w_{ij} \leq d_j + \tilde{d}_j, \quad j = 1, 2, \cdots, N.\end{aligned} \tag{EC.16}$$

Since the parallel allocation is optimal, $z_i = x_i$ $(1 \leq i \leq N)$ in $\mathbf{Z}$ is optimal if $x_i \leq d_i + \tilde{d}_i$. Furthermore, we need to show the optimality of z_i for all classes i's with $x_i > d_i + \tilde{d}_i$, i.e., the classes with surplus capacities after the parallel allocation. Since the general upgrading problem is concave, we only need to examine $\frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N})$ and $\frac{\partial}{\partial z_i^-}\Theta^{t+1}(\mathbf{N})$, where

$$\mathbf{N} = \mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t - \mathbf{Y}^t\mathbf{1} + (\mathbf{Y}^t)^\intercal\mathbf{1} = \mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t - \mathbf{Z} + (\mathbf{W})^\intercal\mathbf{1}$$

is the state at the beginning of period $t+1$.

Without loss of generality, we assume class 1 is the highest class with $x_1 > d_1 + \tilde{d}_1$ and analyze the optimality of z_1 by cases.

1. $z_1 = d_1 + \tilde{d}_1$: We only need to prove that increasing z_1 is suboptimal since $z_1 \geq y_{11} = d_1 + \tilde{d}_1$. Let k $(k > 1)$ be the highest class with $(\mathbf{N})_k < 0$. Note that z_1 is clearly optimal if class k does not exist, i.e., there is no backlogged demand in classes $(1, \cdots, N)$ in $\mathbf{N}$.
    (a) $(\mathbf{N})_{2,\cdots,k-1} = 0$: When solving the protection level p_{1k} and the allocation decision y_{1k} by the PSR, $(\mathbf{N})_{1,\cdots,k}$ are the states of classes $(1, \cdots, k)$. Meanwhile, the upgrading decisions within classes $(k+1, \cdots, N)$ have not been considered, whose states are the states after the parallel allocation, i.e., $(\mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t)_{k+1,\cdots,N}$. Thus,

    $$\begin{aligned}0 \geq & \alpha_{1k} - \Delta_{1k}^{-+}\Theta^{t+1}\left((\mathbf{N})_{1,\cdots,k}, (\mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t)_{k+1,\cdots,N}\right) \\ = & \alpha_{1k} - \Delta_{1k}^{-+}\Theta^{t+1}(\mathbf{N}) = \alpha_{1k} + \frac{\partial}{\partial z_1^+}\Theta^{t+1}(\mathbf{N}),\end{aligned} \tag{EC.17}$$

    where the first equality is from the independence property assumption of Θ^{t+1}, and the second equality follows from the fact that $\mathbf{N}$ changes to $\mathbf{N} + \epsilon(-\mathbf{e}_1 + \mathbf{e}_k)$ when z_1 marginally changes to $z_1 + \epsilon$, where $\epsilon > 0$. Hence, increasing z_1 is suboptimal.

    (b) There exists class i $(1 < i < k)$ with $(\mathbf{N})_i > 0$: Without loss of generality, we assume that i is the lowest class in $(2, \cdots, k-1)$ with $(\mathbf{N})_i > 0$. In this case, the PSR considers protection level p_{ik} and ignores the potential upgrade from class 1 to k, and we will show it is indeed optimal to do so. Since $(\mathbf{N})_{1,\cdots,k}$ are the states of classes $(1, \cdots, k)$ when considering the protection level p_{ik} by the PSR, and $\mathbf{N}$ changes to $\mathbf{N}+\epsilon(-\mathbf{e}_i+\mathbf{e}_k)$ when z_i marginally changes to $z_i+\epsilon$, where $\epsilon>0$. We have

$$0\geq \alpha_{ik}-\Delta_{ik}^{-+}\Theta^{t+1}\left((\mathbf{N})_{1,\cdots,k},(\mathbf{X}^t-\tilde{\mathbf{D}}^t-\mathbf{D}^t)_{k+1,\cdots,N}\right)=\alpha_{ik}+\frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N}).$$

Moreover, because $(\mathbf{N})_1>0$ and $(\mathbf{N})_i>0$,

$$\partial_1^-\Theta^{t+1}(\mathbf{N})-\partial_i^-\Theta^{t+1}(\mathbf{N})\geq u_i-u_1$$

from Lemma EC.1.

Notice that $\mathbf{N}$ changes to $\mathbf{N}+\epsilon(-\mathbf{e}_1+\mathbf{e}_k)$ when z_1 marginally changes to $z_1+\epsilon$, then

$$\frac{\partial}{\partial z_1^+}\Theta^{t+1}(\mathbf{N})-\frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N})=\partial_i^-\Theta^{t+1}(\mathbf{N})-\partial_1^-\Theta^{t+1}(\mathbf{N}). \tag{EC.18}$$

Thus, from $\alpha_{ik}-\alpha_{1k}=u_1-u_i$, we have

$$\frac{\partial}{\partial z_1^+}\Theta^{t+1}(\mathbf{N})+\alpha_{1k}\leq\frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N})+\alpha_{1k}+u_1-u_i=\frac{\partial}{\partial z_i^+}\Theta^{t+1}(\mathbf{N})+\alpha_{ik}\leq 0, \tag{EC.19}$$

which means increasing z_1 is not optimal.

2. $z_1>d_1+\tilde{d}_1$: Let j $(j>1)$ be the lowest class with $y_{1j}>0$ in $\mathbf{Y}^t$. Similar to the previous case, from the PSR, $(\mathbf{N})_{1,\cdots,j}$ are the states after performing the last unit of upgrade y_{1j}. In this case, $\mathbf{N}$ changes to $\mathbf{N}+\epsilon(\mathbf{e}_1-\mathbf{e}_j)$ when z_1 marginally changes to $z_1-\epsilon$, where $\epsilon>0$, then

$$0\leq\alpha_{1j}-\Delta_{1j}^{+-}\Theta^{t+1}\left((\mathbf{N})_{1,\cdots,j},(\mathbf{X}^t-\tilde{\mathbf{D}}^t-\mathbf{D}^t)_{j+1,\cdots,N}\right)=\alpha_{1j}+\frac{\partial}{\partial z_1^-}\Theta^{t+1}(\mathbf{N}). \tag{EC.20}$$

Thus, decreasing current z_1 is costly.

Furthermore, for all class i $(1<i<j)$ with $x_i>d_i+\tilde{d}_i$, $z_i=x_i$ by the PSR policy. First, we only need to show decreasing these z_i's is not optimal.

When z_i marginally changes to $z_i-\epsilon$, there is a *chain reaction*. From (EC.16), decreasing z_i by ϵ is equivalent to reducing the upgrade y_{ik_i} by ϵ, where k_i is the lowest class upgraded by capacity i. Then, unmet demand k_i increases by ϵ unit, and demand k_i will be upgraded by capacity s $(1\leq s<i)$, which is the lowest class with $x_s>\tilde{d}_s+d_s$. Meanwhile, k_s, the lowest class of demands upgraded by capacity s prior to changing z_i, has an additional ϵ unit unmet demand, which can be similarly analyzed as class k_i. The *chain reaction* continues, and $\mathbf{N}$ changes to $\mathbf{N}+\epsilon(\mathbf{e}_i-\mathbf{e}_j)$, i.e., the unmet demand j is increased by ϵ unit.

When z_i marginally changes to $z_i-\epsilon$, only unmet demand j and capacity i changed in the aforementioned *chain reaction*, then the objective function in (EC.16) decreases by $\epsilon\alpha_{ij}$. Meanwhile, given $(\mathbf{N})_1\geq 0$ and $(\mathbf{N})_i=0$, similar to (EC.18), we have

$$\frac{\partial}{\partial z_1^-}\Theta^{t+1}(\mathbf{N})-\frac{\partial}{\partial z_i^-}\Theta^{t+1}(\mathbf{N})\leq u_1-u_i$$

by Lemma EC.1. Thus, from $\alpha_{ij} - \alpha_{1j} = u_1 - u_i$,

$$\frac{\partial}{\partial z_i^-}\Theta^{t+1}(\mathbf{N}) + \alpha_{ij} \geq \frac{\partial}{\partial z_1^-}\Theta^{t+1}(\mathbf{N}) + u_i - u_1 + \alpha_{ij} = \frac{\partial}{\partial z_1^-}\Theta^{t+1}(\mathbf{N}) + \alpha_{1j} \geq 0.$$

Hence, $z_i = x_i$ is optimal for all class i $(1 < i < j)$ with $x_i > d_i + \tilde{d}_i$.

Next, we have to prove that increasing z_1 itself is also suboptimal.

(a) $(\mathbf{N})_j < 0$: In this case, the protection level p_{1j} is binding in the PSR, i.e., the upgrade between classes 1 and j stops when the quantity of capacity 1 reaches p_{1j}. From the definition of p_{1j}, and the fact that $\mathbf{N}$ changes to $\mathbf{N} + \epsilon(-\mathbf{e}_1 + \mathbf{e}_j)$ when z_1 marginally changes to $z_1 + \epsilon$, we have

$$\begin{aligned} 0 \geq & \alpha_{1j} - \Delta_{1j}^{-+}\Theta^{t+1}\left((\mathbf{N})_{1,\cdots,j}, (\mathbf{X}^t - \tilde{\mathbf{D}}^t - \mathbf{D}^t)_{j+1,\cdots,N}\right) \\ = & \alpha_{1j} + \frac{\partial}{\partial z_1^+}\Theta^{t+1}(\mathbf{N}). \end{aligned} \tag{EC.21}$$

From (EC.20) and (EC.21), we know the optimality of z_1.

(b) $(\mathbf{N})_j = 0$: The upgrading decision y_{1j} is bounded because there is no unmet demand j remaining, and we do not have (EC.21) directly from solving p_{1j}. However, similar to the case when $z_1 = d_1 + \tilde{d}_1$, increasing z_1 is still suboptimal. Particularly, if there exists k $(k > j)$ as the highest class with $(\mathbf{N})_k < 0$, and $(\mathbf{N})_s = 0$ for all class s $(j < s < k)$, then we have (EC.17) that affirms the optimality of z_1. On the other hand, if there exists class i $(j < i < k)$ with $(\mathbf{N})_i > 0$, then (EC.19) is valid, which also proves the optimality of z_1.

To summarize, we have proved that z_1 is optimal. In addition, if $z_1 > d_1 + \tilde{d}_1$ and class j is the lowest class with $y_{1j} > 0$ in $\mathbf{Y}^t$, we have also shown the optimality of z_i $(1 < i < j)$ with $x_i > d_i + \tilde{d}_i$. The same argument can be sequentially applied to the rest of z_s's since $(\mathbf{N})_{1,\cdots,s-1}$ are the states of classes $(1, \cdots, s-1)$ when solving the protection levels within classes $(s, \cdots, N)$ in the PSR policy.

Therefore, the PSR policy solves the general upgrading problem in period t.

1. Independence property of Θ^t

As the PSR solves the general upgrading problem in period t, and the independence property of Θ^{t+1} holds by Property 1 of the induction assumption, all requirements of Lemma EC.5 are satisfied, thus the independence property of Θ^t also holds.

To conclude the proof, we now consider period T. The PSR solves the general upgrading problem in period T by Lemma EC.3. And Lemma EC.4 asserts the independence property of Θ^T. Therefore, we can use the backward induction and complete the proof. □

B.4. Properties of the Protection Levels

PROPOSITION 3 *If initial capacity $\mathbf{X}^1$ and demand $\mathbf{D}^1, \cdots, \mathbf{D}^T$ are integer valued, there exists an integer valued optimal policy $\mathbf{Y}^1, \cdots, \mathbf{Y}^T$ derived by the PSR policy.*

Proof. The proof is similar to the proof of Proposition 3 in Shumsky and Zhang (2009). ☐

PROPOSITION 4 *For the same* $(n_1^t, \cdots, n_{i-1}^t)$ *in period* t $(1 \le t \le T)$, $p_{ij} \le p_{i,j+1}$ *when* $i < j$.

Proof. Suppose to the contrary that $p_{ij} > p_{i,j+1}$ in period t. Let $\bar{p} = \frac{p_{ij}+p_{i,j+1}}{2}$, and denote $\mathbf{N} = (n_1^t, \cdots, n_{i-1}^t, \bar{p}, 0, \cdots, 0, n_{j+2}^t, \cdots, n_N^t)$. From the concavity in Proposition 1 and the independence property in Proposition 2, we have $\Delta_{i,j+1}^{+-}\Theta^{t+1}(\mathbf{N}) \le \alpha_{i,j+1}$ given $\bar{p} > p_{i,j+1}$. Similarly, we have $\Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}) \ge \alpha_{ij}$ since $\bar{p} < p_{ij}$.

However, since $\alpha_{ij} - \alpha_{i,j+1} = r_j + g_j - r_{j+1} - g_{j+1}$ and $g_j > g_{j+1}$,

$$
\begin{aligned}
&\Delta_{i,j+1}^{+-}\Theta^{t+1}(\mathbf{N}) = \partial_i^{+}\Theta^{t+1}(\mathbf{N}) - \partial_{j+1}^{-}\Theta^{t+1}(\mathbf{N})\\
\ge &\partial_i^{+}\Theta^{t+1}(\mathbf{N}) - \partial_j^{-}\Theta^{t+1}(\mathbf{N}) + r_{j+1} - r_j = \Delta_{ij}^{+-}\Theta^{t+1}(\mathbf{N}) + r_{j+1} - r_j\\
\ge &\alpha_{ij} + r_{j+1} - r_j > \alpha_{i,j+1},
\end{aligned}
$$

where the first inequality follows from Lemma EC.2. This is a contradiction. Hence, $p_{ij} \le p_{i,j+1}$ when $i < j$. ☐

Appendix C: Multi-Horizon Model with Replenishment

PROPOSITION 6 *Suppose the firm starts with an initial capacity* $\mathbf{X} \le \mathbf{X}^*$ *in* (11)*. The firm's optimal replenishment policy in the multi-horizon model is a base stock policy with the optimal base stock level* $\mathbf{X}^*$*. Furthermore, the PSR policy solves the optimal allocation decisions within each horizon.*

Proof. We prove this proposition by induction. Let $V_k(\mathbf{X}, \tilde{\mathbf{D}})$ $(1 \le k \le K)$ denote the expected revenue-to-go function at the beginning of the k-th horizon with capacity $\mathbf{X}$ and backlogged demand $\tilde{\mathbf{D}}$. Where possible, the index of periods in each horizon is denoted by superscripts while subscripts denote the index of horizons. We inductively prove the following three properties.

1. The PSR policy optimally solves the capacity allocation decisions in the k-th horizon;
2. The optimal replenishment policy in the k-th horizon is a base stock policy with the optimal base stock level $\mathbf{X}^*$;
3. If $\mathbf{X} \le \mathbf{X}^*$, $V_k(\mathbf{X}, \tilde{\mathbf{D}})$ is affine in $\mathbf{X}$ with slope $\mathbf{C}$ and $\tilde{\mathbf{D}}$ with slope $\boldsymbol{\alpha} - \mathbf{C}$.

Suppose all properties hold in the $(k+1)$-th horizon. It suffices to consider capacity $\mathbf{X} \le \mathbf{X}^*$. Since $V_{k+1}(\mathbf{X}, \tilde{\mathbf{D}})$ is affine in $(\mathbf{X}, \tilde{\mathbf{D}})$, in horizon k we have

$$\begin{aligned}
\Theta_k^T(\mathbf{X}, \tilde{\mathbf{D}}) &= \mathop{\mathbb{E}}_{\mathbf{D}_k^T} \left\{ \max_{\mathbf{Y}_k^T} \left[H(\mathbf{Y}_k^T | \tilde{\mathbf{D}}; \mathbf{D}_k^T) - \mathbf{h}\mathbf{X}_k^{T+1} + \gamma V_{k+1}(\mathbf{X}_k^{T+1}, \tilde{\mathbf{D}}_k^{T+1}) \right] \right\} \\
&= \mathop{\mathbb{E}}_{\mathbf{D}_k^T} \left\{ \max_{\mathbf{Y}_k^T} \left[H(\mathbf{Y}_k^T | \tilde{\mathbf{D}}; \mathbf{D}_k^T) - \mathbf{h}\mathbf{X}_k^{T+1} + \gamma V_{k+1}(\mathbf{X}^*, 0) + \gamma \mathbf{C}(\mathbf{X}_k^{T+1} - \mathbf{X}^*) + \gamma(\boldsymbol{\alpha} - \mathbf{C})\tilde{\mathbf{D}}_k^{T+1} \right] \right\} \\
&= \mathop{\mathbb{E}}_{\mathbf{D}_k^T} \left\{ \max_{\mathbf{Y}_k^T} \left[H(\mathbf{Y}_k^T | \tilde{\mathbf{D}}; \mathbf{D}_k^T) + (\gamma\mathbf{C} - \mathbf{h})\mathbf{X}_k^{T+1} + \gamma(\boldsymbol{\alpha} - \mathbf{C})\tilde{\mathbf{D}}_k^{T+1} + \gamma\left(V_{k+1}(\mathbf{X}^*, 0) - \mathbf{C}\mathbf{X}^*\right) \right] \right\} \\
&= \mathop{\mathbb{E}}_{\mathbf{D}_k^T} \left\{ \max_{\mathbf{Y}_k^T} \left[H(\mathbf{Y}_k^T | \tilde{\mathbf{D}}; \mathbf{D}_k^T) + (\gamma\mathbf{C} - \mathbf{h})\mathbf{X}_k^{T+1} + \gamma(\boldsymbol{\alpha} - \mathbf{C})\tilde{\mathbf{D}}_k^{T+1} \right] \right\} + \gamma\left(V_{k+1}(\mathbf{X}^*, 0) - \mathbf{C}\mathbf{X}^*\right),
\end{aligned} \tag{EC.22}$$

where $\gamma\left(V_{k+1}(\mathbf{X}^*, 0) - \mathbf{C}\mathbf{X}^*\right)$ is a constant. From the proof of Proposition 2, the PSR policy optimally solves the T-th period allocation decisions in the k-th horizon, where the protection levels are based on $\Theta^{T+1}(\mathbf{X}^{T+1}, \tilde{\mathbf{D}}^{T+1})$ in (10). Moreover, it is clear that $\Theta_k^T(\mathbf{X}, \tilde{\mathbf{D}})$ is also concave in $(\mathbf{X}, \tilde{\mathbf{D}})$ from the proof of Proposition 1. Inductively, for $t = T-1, \cdots, 1$, we know

$$\Theta_k^t(\mathbf{X}, \tilde{\mathbf{D}}) = \mathop{\mathbb{E}}_{\mathbf{D}_k^t} \left\{ \max_{\mathbf{Y}_k^t} \left[H(\mathbf{Y}_k^t | \tilde{\mathbf{D}}; \mathbf{D}_k^t) + \Theta_k^{t+1}(\mathbf{X}_k^{t+1}, \tilde{\mathbf{D}}_k^{t+1}) \right] \right\}$$

is concave in $(\mathbf{X}, \tilde{\mathbf{D}})$, and we can show that the PSR policy solves the capacity allocation decisions for horizon k.

From the Bellman equation, we have

$$V_k(\mathbf{X}, \tilde{\mathbf{D}}) = \max_{\mathbf{Z} \ge \mathbf{X}} G(\mathbf{Z}),$$

where $G(\mathbf{Z}) = \Theta_k^1(\mathbf{Z}, \mathbf{0}) + \boldsymbol{\alpha}\tilde{\mathbf{D}} - \mathbf{C}(\mathbf{Z} - \mathbf{X} + \tilde{\mathbf{D}})$. From (EC.22), we have

$$G(\mathbf{Z}) = \Pi\left(\mathbf{Z}; \gamma\mathbf{C} - \mathbf{h}, \gamma(\boldsymbol{\alpha} - \mathbf{C})\right) + (\boldsymbol{\alpha} - \mathbf{C})\tilde{\mathbf{D}} + \mathbf{C}(\mathbf{X} - \gamma\mathbf{X}^*) + \gamma V_{k+1}(\mathbf{X}^*, 0).$$

By the definition of $\mathbf{X}^*$, the optimal replenishment policy in the k-th horizon is a base stock policy with optimal base stock level $\mathbf{X}^*$. Furthermore, for $\mathbf{X} \le \mathbf{X}^*$,

$$V_k(\mathbf{X}, \tilde{\mathbf{D}}) = \Pi\left(\mathbf{X}^*; \gamma\mathbf{C} - \mathbf{h}, \gamma(\boldsymbol{\alpha} - \mathbf{C})\right) + (\boldsymbol{\alpha} - \mathbf{C})\tilde{\mathbf{D}} + \mathbf{C}(\mathbf{X} - \gamma\mathbf{X}^*) + \gamma V_{k+1}(\mathbf{X}^*, 0)$$

is affine in $\mathbf{X}$ with slope $\mathbf{C}$ and $\tilde{\mathbf{D}}$ with slope $\boldsymbol{\alpha} - \mathbf{C}$.

To conclude the proof, we consider the last horizon profit, $V_K(\mathbf{X}, \tilde{\mathbf{D}})$. Since

$$\Theta_K^T(\mathbf{X}, \tilde{\mathbf{D}}) = \mathop{\mathbb{E}}_{\mathbf{D}_K^T} \left\{ \max_{\mathbf{Y}_K^T} \left[H(\mathbf{Y}_K^T | \tilde{\mathbf{D}}; \mathbf{D}_K^T) - \mathbf{h}\mathbf{X}_K^{T+1} + \gamma\mathbf{C}\mathbf{X}_K^{T+1} + \gamma(\boldsymbol{\alpha} - \mathbf{C})\tilde{\mathbf{D}}_K^{T+1} \right] \right\}$$

by definition, the optimality of the PSR policy can be similarly proved. Meanwhile, if $\mathbf{X} \le \mathbf{X}^*$,

$$V_K(\mathbf{X}, \tilde{\mathbf{D}}) = \max_{\mathbf{Z} \ge \mathbf{X}} \left[\Pi\left(\mathbf{Z}; \gamma \mathbf{C} - \mathbf{h}, \gamma(\boldsymbol{\alpha} - \mathbf{C})\right) + (\boldsymbol{\alpha} - \mathbf{C})\tilde{\mathbf{D}} + \mathbf{C}\mathbf{X} \right],$$

so the base stock policy is optimal and $V_K(\mathbf{X}, \tilde{\mathbf{D}})$ is affine in $\mathbf{X}$ with slope $\mathbf{C}$ and $\tilde{\mathbf{D}}$ with slope $\boldsymbol{\alpha} - \mathbf{C}$. Therefore, all properties hold for the K-th horizon, which completes the proof. □

Appendix D: Numerical Studies

D.1. Numerical study with $N = 4$ and $T = 3$

In Table 1, we consider problems with size $N = 4$ and $T = 3$. For such a problem size, we can use backward induction to calculate the firm's optimal revenue, which serves as a benchmark to evaluate the performance of the RCEC heuristics. Below we describe the design of the numerical study in detail. The description consists of three parts: demand patterns, economic parameters, and initial capacity.

D.1.1. Demand patterns To cover a wide range of demand scenarios, we consider 13 evolution patterns for product demand means in Table EC.1. For each evolution pattern, we define vectors $\boldsymbol{\mu}^t = (\mu_1^t, \cdots, \mu_N^t)^\intercal$ $(t = 1, \cdots, T)$, where μ_i^t is the demand mean of product i in period t. The demand mean patterns in Table EC.1 cover some typical realistic scenarios. For instance, in pattern 4, the expected demand for high-quality products are higher than that for low-quality products when the period is close to the end of horizon, which corresponds to revenue management situations.

Given an evolution pattern $\boldsymbol{\mu}^t$ $(t = 1, \cdots, T)$ for the demand means, we generate a sample of random demands for each product in each period. Specifically, given the demand mean $\boldsymbol{\mu}^t$ in period t, we generate demand $\mathbf{D}^t$ by using either Poisson distribution or multivariate normal distribution with covariance matrix

$$\begin{pmatrix} 0.5 & 0.15 & 0.075 & 0.0375 \\ 0.15 & 0.5 & 0.15 & 0.075 \\ 0.075 & 0.15 & 0.5 & 0.15 \\ 0.0375 & 0.075 & 0.15 & 0.5 \end{pmatrix} * \boldsymbol{\mu}^t,$$

and

$$\begin{pmatrix} 0.5 & -0.15 & -0.075 & -0.0375 \\ -0.15 & 0.5 & -0.15 & -0.075 \\ -0.075 & -0.15 & 0.5 & -0.15 \\ -0.0375 & -0.075 & -0.15 & 0.5 \end{pmatrix} * \boldsymbol{\mu}^t.$$

The first covariance matrix represents positive correlation between the products, while the second represents negative correlation between the products. For normal distribution, we truncate the demand realizations at zero and round them to the nearest integer values. By the above construction, there are total $39 = 3 * 13$ demand scenarios.

Pattern	Description	Example ($T = 3$)
Linear	1. Product 1 demand increases, product 2 demand is flat, product 3 and 4 demands decrease with the same rate.	$\begin{pmatrix} 2 & 4 & 6 & 6 \\ 4 & 4 & 4 & 4 \\ 6 & 4 & 2 & 2 \end{pmatrix}$
	2. Product 1 demand increases, product 2 demand is flat, product 3 demand decreases, product 4 demand decreases in half of the rate of product 3.	$\begin{pmatrix} 2 & 4 & 6 & 4 \\ 4 & 4 & 4 & 3 \\ 6 & 4 & 2 & 2 \end{pmatrix}$
	3. Product 1, 2, 3 and 4 demands are flat.	$\begin{pmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{pmatrix}$
	4. Product 1 and 2 demands increase with the same rate, product 3 and 4 demands decrease with the same rate.	$\begin{pmatrix} 2 & 2 & 6 & 6 \\ 4 & 4 & 4 & 4 \\ 6 & 6 & 2 & 2 \end{pmatrix}$
	5. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3 demand decreases, product 4 demand decreases in half of the rate of product 3.	$\begin{pmatrix} 2 & 2 & 6 & 4 \\ 4 & 3 & 4 & 3 \\ 6 & 4 & 2 & 2 \end{pmatrix}$
	6. Product 1 and 2 demands increase with the same rate, product 3 demand is flat, product 4 demand decreases.	$\begin{pmatrix} 2 & 2 & 4 & 6 \\ 4 & 4 & 4 & 4 \\ 6 & 6 & 4 & 2 \end{pmatrix}$
	7. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3 demand is flat, product 4 demand decreases.	$\begin{pmatrix} 2 & 2 & 4 & 6 \\ 4 & 3 & 4 & 4 \\ 6 & 4 & 4 & 2 \end{pmatrix}$
	8. Product 1 and 2 demands increase with the same rate, product 3 and 4 demands are flat.	$\begin{pmatrix} 2 & 2 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 6 & 6 & 4 & 4 \end{pmatrix}$
	9. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3 and 4 demands are flat.	$\begin{pmatrix} 2 & 2 & 4 & 4 \\ 4 & 3 & 4 & 4 \\ 6 & 4 & 4 & 4 \end{pmatrix}$
Alternating	10. Products 1 and 3 start with positive demand, while products 2 and 4 start with zero demand.	$\begin{pmatrix} 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{pmatrix}$
	11. Products 1 and 3 start with positive demand, where demand 3 is smaller than demand 1 in each period, and products 2 and 4 start with zero demand, where demand 4 is smaller than demand 2 in each period.	$\begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 6 & 0 & 2 \\ 6 & 0 & 2 & 0 \end{pmatrix}$
	12. Products 2 and 4 start with positive demand, while products 1 and 3 start with zero demand.	$\begin{pmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \end{pmatrix}$
	13. Products 2 and 4 start with positive demand, where demand 4 is smaller than demand 2 in each period, and products 1 and 3 start with zero demand, where demand 3 is smaller than demand 1 in each period.	$\begin{pmatrix} 0 & 6 & 0 & 2 \\ 6 & 0 & 2 & 0 \\ 0 & 6 & 0 & 2 \end{pmatrix}$

Table EC.1 Demand patterns with 4 products.

D.1.2. Economic parameters We also consider a wide variety of values for the economic parameters while using the same backorder cost $(g_1, g_2, g_3, g_4) = (1.0, 0.9, 0.8, 0.7)$. Recall the upgrading revenue is given by $\alpha_{ij} = r_j + g_j - u_i$; instead of specifying r_j and u_i, we choose to specify α_{ij}, which is sufficient for the numerical study. Four different matrices of $\boldsymbol{\alpha} = (\alpha_{ij})_{4\times 4}$ have been considered in the numerical study. The capacity costs (c_1, c_2, c_3, c_4) are decided by $c_i = 0.3\alpha_{ii}$ $(i = 1, \cdots, 4)$ for each matrix.

1. Matrix 1: The parallel revenue decreases in product class, and the upgrading revenues are close to the parallel revenue.
$$\begin{pmatrix} 16 & 14 & 12 & 10 \\ 0 & 15 & 13 & 11 \\ 0 & 0 & 14 & 12 \\ 0 & 0 & 0 & 13 \end{pmatrix}$$
2. Matrix 2: The parallel revenue decreases in product class, and the upgrading revenues are decreasing in the number of levels of upgrading (e.g., 1-step upgrading revenue is 11 and 2-step upgrading is either 7 or 8).
$$\begin{pmatrix} 16 & 11 & 7 & 4 \\ 0 & 15 & 11 & 8 \\ 0 & 0 & 14 & 11 \\ 0 & 0 & 0 & 13 \end{pmatrix}$$
3. Matrix 3: The parallel revenue decreases in product class, and α_{12} and α_{34} are higher than α_{23}.
$$\begin{pmatrix} 16 & 14 & 5 & 3 \\ 0 & 15 & 6 & 4 \\ 0 & 0 & 14 & 12 \\ 0 & 0 & 0 & 13 \end{pmatrix}$$
4. Matrix 4: The parallel revenue is constant across products, the 2-step upgrading revenue is constant, and α_{23} is higher than the other 1-step upgrading revenue.
$$\begin{pmatrix} 16 & 10 & 9 & 3 \\ 0 & 16 & 15 & 9 \\ 0 & 0 & 16 & 10 \\ 0 & 0 & 0 & 16 \end{pmatrix}$$

D.1.3. Initial capacity When choosing the initial capacity, we start with optimal capacity level $\mathbf{X}_{\text{RCEC}}$ using the RCEC heuristic. To ensure the robustness of the results, we also consider a number of variants of $\mathbf{X}_{\text{RCEC}}$, among which some are extreme capacity scenarios. In particular, we use $\mathbf{X}_{\text{RCEC}} = (x_1^{\text{RCEC}}, x_2^{\text{RCEC}}, x_3^{\text{RCEC}}, x_4^{\text{RCEC}})$ to construct the following five patterns of initial capacity:

1. $\mathbf{X} = \lambda \mathbf{X}_{\text{RCEC}}$
2. For each $i \in \{1,2,3\}$: $(\mathbf{X})_i = \lambda(x_i^{\text{RCEC}} + x_{i+1}^{\text{RCEC}}),\ (\mathbf{X})_{i+1} = 0,\ (\mathbf{X})_s = \lambda x_s^{\text{RCEC}},\ \forall\ s \in \{1,2,3,4\} \setminus \{i, i+1\}$
3. For each $i \in \{2,3,4\}$: $(\mathbf{X})_i = 0,\ (\mathbf{X})_s = \lambda x_s^{\text{RCEC}},\ \forall\ s \in \{1,2,3,4\} \setminus \{i\}$
4. $\mathbf{X} = \lambda(x_1^{\text{RCEC}} + x_2^{\text{RCEC}}, 0, x_3^{\text{RCEC}} + x_4^{\text{RCEC}}, 0)$

5. $\mathbf{X} = \lambda(x_1^{\text{RCEC}} + x_3^{\text{RCEC}}, x_2^{\text{RCEC}} + x_4^{\text{RCEC}}, 0, 0)$,

where $\lambda = \{0.75,\ 1,\ 1.25\}$. Each pattern corresponds to a realistic or extreme scenario. For instance, in Pattern 2, a certain product has extremely low inventory level while the adjacent high-quality product is abundant; in Pattern 5, the last two products have extremely low investment while there are plenty of higher level products. Note that in some of the patterns (e.g., Patterns 2-5), upgrading would be frequently performed. The parameter λ is used to adjust the capacity-demand ratio (e.g., $\lambda = 0.75$ implies that the aggregate capacity level is relatively low). For each λ, there are 9 initial capacity scenarios; so there are totally 27 capacity scenarios.

To summarize, we test $39 * 4 * 27 = 4212$ problem instances by the above construction. They cover a wide range of possible situations that may arise in practice.

D.2. Numerical study with $N = 5$ and $T \in \{3, 15, 30, 60\}$

This is the major numerical study in this paper; it serves several purposes. First, we test the performance of the RCEC heuristic for problems with larger sizes in Tables 2 and 3; second, we examine the value of multi-step upgrading in Tables 4 and 5; third, we investigate the importance of the allocation mechanism and the capacity sizing decision in Table 6. To make the results comparable across different T values, we make a couple of assumptions: (1) For each product i, the expected total demand throughout the sales horizon is the same for different T values; that is, the sum $\sum_{t=1}^{T} \mu_i^t$ in each demand evolution pattern $\boldsymbol{\mu}^t = (\mu_1^t, \cdots, \mu_N^t)$ $(t = 1, \cdots, N)$ is the same for different T values, which is set to be 60 for each i. (2) For each parameter combination, the capacity cost is the same for different T's. (3) The per-period goodwill cost is linearly decreasing in T (using a constant goodwill cost independent of T will not affect the performance result about RCEC). Below we describe the design of the numerical study in detail. Again the description consists of three parts: demand patterns, economic parameters, and initial capacity.

D.2.1. Demand patterns Similar to the first numerical study in Section D.1, we consider 13 demand evolution patterns in Table EC.2.

Again, given an evolution pattern $\boldsymbol{\mu}^t$ $(t = 1, \cdots, T)$ for the demand means, we generate a sample of random demands for each product in each period. Specifically, given the demand mean $\boldsymbol{\mu}^t$ in period t, we generate demand $\mathbf{D}^t$ by using either Poisson distribution or multivariate normal distribution with a positive covariance matrix

$$\begin{pmatrix} 0.5 & 0.15 & 0.12 & 0.09 & 0.06 \\ 0.15 & 0.5 & 0.15 & 0.12 & 0.09 \\ 0.12 & 0.15 & 0.5 & 0.15 & 0.12 \\ 0.09 & 0.12 & 0.15 & 0.5 & 0.15 \\ 0.06 & 0.09 & 0.12 & 0.15 & 0.5 \end{pmatrix} * \boldsymbol{\mu}^t,$$

Pattern	Description	Example ($T=3$)
Linear	1. Product 1 demand increases, product 2 and 3 demands are flat, product 4 and 5 demands decrease with the same rate.	$\begin{pmatrix} 4 & 8 & 8 & 12 & 12 \\ 8 & 8 & 8 & 8 & 8 \\ 12 & 8 & 8 & 4 & 4 \end{pmatrix}$
	2. Product 1 demand increases, product 2 and 3 demands are flat, product 4 demand decreases, product 5 demand decreases in half of the rate of product 4.	$\begin{pmatrix} 4 & 8 & 8 & 12 & 8 \\ 8 & 8 & 8 & 8 & 6 \\ 12 & 8 & 8 & 4 & 4 \end{pmatrix}$
	3. Product 1, 2, 3, 4 and 5 demands are flat.	$\begin{pmatrix} 8 & 8 & 8 & 8 & 8 \\ 8 & 8 & 8 & 8 & 8 \\ 8 & 8 & 8 & 8 & 8 \end{pmatrix}$
	4. Product 1, 2 and 3 demands increase with the same rate, product 4 and 5 demands decrease with the same rate.	$\begin{pmatrix} 4 & 4 & 4 & 12 & 12 \\ 8 & 8 & 8 & 8 & 8 \\ 12 & 12 & 12 & 4 & 4 \end{pmatrix}$
	5. Product 1 demand increases, product i ($i=2,3$) demand increases in half of the rate of product $i-1$, product 4 demand decreases, product 5 demand decreases in half of the rate of product 4.	$\begin{pmatrix} 4 & 4 & 4 & 12 & 8 \\ 8 & 6 & 5 & 8 & 6 \\ 12 & 8 & 6 & 4 & 4 \end{pmatrix}$
	6. Product 1 and 2 demands increase with the same rate, product 3 demand is flat, product 4 and 5 demands decrease with the same rate.	$\begin{pmatrix} 4 & 4 & 8 & 12 & 12 \\ 8 & 8 & 8 & 8 & 8 \\ 12 & 12 & 8 & 4 & 4 \end{pmatrix}$
	7. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3 demand is flat, product 4 demand decreases, product 5 demand decreases in half of the rate of product 4.	$\begin{pmatrix} 4 & 4 & 8 & 12 & 8 \\ 8 & 6 & 8 & 8 & 6 \\ 12 & 8 & 8 & 4 & 4 \end{pmatrix}$
	8. Product 1 and 2 demands increase with the same rate, product 3, 4 and 5 demands are flat.	$\begin{pmatrix} 4 & 4 & 8 & 8 & 8 \\ 8 & 8 & 8 & 8 & 8 \\ 12 & 12 & 8 & 8 & 8 \end{pmatrix}$
	9. Product 1 demand increases, product 2 demand increases in half of the rate of product 1, product 3, 4 and 5 demands are flat.	$\begin{pmatrix} 4 & 4 & 8 & 8 & 8 \\ 8 & 6 & 8 & 8 & 8 \\ 12 & 8 & 8 & 8 & 8 \end{pmatrix}$
Alternating	10. Products 1, 3 and 5 start with positive demand, while products 2 and 4 start with zero demand.	$\begin{pmatrix} 8 & 0 & 8 & 0 & 8 \\ 0 & 8 & 0 & 8 & 0 \\ 8 & 0 & 8 & 0 & 8 \end{pmatrix}$
	11. Products 1, 3 and 5 start with positive demand, where demand i ($i=3,5$) is smaller than demand $i-2$ in each period, and products 2 and 4 start with zero demand, where demand 4 is smaller than demand 2 in each period.	$\begin{pmatrix} 16 & 0 & 8 & 0 & 4 \\ 0 & 16 & 0 & 8 & 0 \\ 16 & 0 & 8 & 0 & 4 \end{pmatrix}$
	12. Products 2 and 4 start with positive demand, while products 1, 3 and 5 start with zero demand.	$\begin{pmatrix} 0 & 8 & 0 & 8 & 0 \\ 8 & 0 & 8 & 0 & 8 \\ 0 & 8 & 0 & 8 & 0 \end{pmatrix}$
	13. Products 2 and 4 start with positive demand, where demand 4 is smaller than demand 2 in each period, and products 1, 3 and 5 start with zero demand, where demand i ($i=3,5$) is smaller than demand $i-2$ in each period.	$\begin{pmatrix} 0 & 16 & 0 & 8 & 0 \\ 16 & 0 & 8 & 0 & 4 \\ 0 & 16 & 0 & 8 & 0 \end{pmatrix}$

Table EC.2 Demand patterns with 5 products.

and a negative covariance matrix

$$\begin{pmatrix} 0.5 & -0.15 & -0.12 & -0.09 & -0.06 \\ -0.15 & 0.5 & -0.15 & -0.12 & -0.09 \\ -0.12 & -0.15 & 0.5 & -0.15 & -0.12 \\ -0.09 & -0.12 & -0.15 & 0.5 & -0.15 \\ -0.06 & -0.09 & -0.12 & -0.15 & 0.5 \end{pmatrix} * \boldsymbol{\mu}^t.$$

The rest of the details are the same as in the first numerical study and therefore omitted. There are totally $3 * 13 = 39$ demand scenarios.

D.2.2. Economic parameters For all problem instances, we use the same backorder cost vector $T * (g_1, \cdots, g_5) = (6.0, 5.7, 5.4, 5.1, 4.8)$. Then we use 4 sets of parameters for the revenue $(r_1, \cdots, r_5)$, usage cost $(u_1, \cdots, u_5)$ and capacity cost $(c_1, \cdots, c_5)$, which is explained next. Since the vector $(g_1, \cdots, g_5)$ decreases in T, each element of $\boldsymbol{\alpha} = (\alpha_{ij})_{N \times N}$ decreases in T as well, where α_{15} is the smallest entry with $T = 15$. For illustration, we provide $\boldsymbol{\alpha}$ for $T = 15$ in each of the parameter set for revenue $(r_1, \cdots, r_5)$, usage cost $(u_1, \cdots, u_5)$ and capacity cost $(c_1, \cdots, c_5)$. In particular, given α_{15} in $T = 15$, the capacity cost $(c_1, \cdots, c_5)$ are determined by $c_1 = 0.7 * \alpha_{15}$, and $c_{i+1} = c_i - 0.05$ $(i > 1)$ for parameter set 1 and $c_{i+1} = c_i - 0.02$ $(i > 1)$ for the other sets.

1. Parameter set 1: Upgrading revenue is close to the parallel revenue.

$$\begin{aligned} (r_1, \cdots, r_5) &= (20, 18, 16, 14, 12) \\ (u_1, \cdots, u_5) &= (5, 4, 3, 2, 1) \\ (c_1, \cdots, c_5) &= (5.124, 5.074, 5.024, 4.974, 4.924). \end{aligned}$$

For $T = 15$, there is

$$\boldsymbol{\alpha} = \begin{pmatrix} 15.4 & 13.38 & 11.36 & 9.34 & 7.32 \\ 0 & 14.38 & 12.36 & 10.34 & 8.32 \\ 0 & 0 & 13.36 & 11.34 & 9.32 \\ 0 & 0 & 0 & 12.34 & 10.32 \\ 0 & 0 & 0 & 0 & 11.32 \end{pmatrix}$$

2. Parameter set 2: Revenues of 1-step upgrading are almost identical for different classes.

$$\begin{aligned} (r_1, \cdots, r_5) &= (26, 21, 17, 14, 12) \\ (u_1, \cdots, u_5) &= (11, 7, 4, 2, 1) \\ (c_1, \cdots, c_5) &= (0.924, 0.904, 0.884, 0.864, 0.844). \end{aligned}$$

For $T = 15$, there is

$$\boldsymbol{\alpha} = \begin{pmatrix} 15.4 & 10.38 & 6.36 & 3.34 & 1.32 \\ 0 & 14.38 & 10.36 & 7.34 & 5.32 \\ 0 & 0 & 13.36 & 10.34 & 8.32 \\ 0 & 0 & 0 & 12.34 & 10.32 \\ 0 & 0 & 0 & 0 & 11.32 \end{pmatrix}$$

3. Parameter set 3: α_{12} is much smaller than parallel revenue α_{11}. However, α_{23}, α_{34} and α_{45} are close to α_{22}, α_{33} and α_{44}, respectively.

$$
\begin{aligned}
(r_1,\cdots,r_5) &= (26,21,17,14,12)\\
(u_1,\cdots,u_5) &= (11,7,4,2,1)\\
(c_1,\cdots,c_5) &= (2.324,2.304,2.284,2.264,2.244).
\end{aligned}
$$

For $T=15$, there is

$$
\boldsymbol{\alpha} = \begin{pmatrix} 15.4 & 9.38 & 7.36 & 5.34 & 3.32\\ 0 & 14.38 & 12.36 & 10.34 & 8.32\\ 0 & 0 & 13.36 & 11.34 & 9.32\\ 0 & 0 & 0 & 12.34 & 10.32\\ 0 & 0 & 0 & 0 & 12.32 \end{pmatrix}
$$

4. Parameter set 4: α_{12} and α_{45} are close to parallel revenues α_{11} and α_{44}, respectively. However, α_{23} and α_{34} are much smaller than α_{22} and α_{33}.

$$
\begin{aligned}
(r_1,\cdots,r_5) &= (26,21,17,14,12)\\
(u_1,\cdots,u_5) &= (11,7,4,2,1)\\
(c_1,\cdots,c_5) &= (1.624,1.604,1.584,1.564,1.544).
\end{aligned}
$$

For $T=15$,

$$
\boldsymbol{\alpha} = \begin{pmatrix} 15.4 & 13.38 & 8.36 & 4.34 & 2.32\\ 0 & 14.38 & 9.36 & 5.34 & 3.32\\ 0 & 0 & 13.36 & 9.34 & 7.32\\ 0 & 0 & 0 & 12.34 & 10.32\\ 0 & 0 & 0 & 0 & 11.32 \end{pmatrix}
$$

There are totally 4 parameter combinations.

D.2.3. Initial capacity Similar to the first numerical study, we use $\mathbf{X}_{\text{RCEC}}$ to construct the following five patterns of initial capacity.

1. $\mathbf{X} = \lambda \mathbf{X}_{\text{RCEC}}$
2. For each $i,j \in \{1,2,3,4,5\}$ with $i<j$: $(\mathbf{X})_i = \lambda\left((\mathbf{X}_{\text{RCEC}})_i + (\mathbf{X}_{\text{RCEC}})_j\right)$, $(\mathbf{X})_j = 0$, $(\mathbf{X})_s = \lambda(\mathbf{X}_{\text{RCEC}})_s$, $\forall\ s \in \{1,2,3,4,5\}\setminus\{i,j\}$
3. For each $i \in \{2,3,4,5\}$: $(\mathbf{X})_i = 0$, $(\mathbf{X})_s = \lambda(\mathbf{X}_{\text{RCEC}})_s$, $\forall\ s \in \{1,2,3,4,5\}\setminus\{i\}$
4. $\mathbf{X} = \lambda\left((\mathbf{X}_{\text{RCEC}})_1 + (\mathbf{X}_{\text{RCEC}})_2, 0, (\mathbf{X}_{\text{RCEC}})_3 + (\mathbf{X}_{\text{RCEC}})_4, 0, (\mathbf{X}_{\text{RCEC}})_5\right)$
5. $\mathbf{X} = \lambda\left((\mathbf{X}_{\text{RCEC}})_1, (\mathbf{X}_{\text{RCEC}})_2 + (\mathbf{X}_{\text{RCEC}})_4, 0, (\mathbf{X}_{\text{RCEC}})_3 + (\mathbf{X}_{\text{RCEC}})_5, 0\right)$,

where $\lambda = \{0.7, 0.9, 1.0, 1.1, 1.3\}$. Again the parameter λ is used to adjust the capacity-demand ratio (e.g., $\lambda = 0.7$ implies that the aggregate capacity level is relatively low). For each λ, there are 17 initial capacity scenarios; so there are totally 85 capacity scenarios.

To summarize, we test $39 * 4 * 85 = 13260$ problem instances by the above construction. They cover a wide range of possible situations that may arise in practice.

D.3. Impact of Allocation Mechanism: Suboptimal k-Step Upgrading

We examine the profit loss of adopting suboptimal, k-step upgrading ($k = 0, \cdots, N-2$). Given $\mathbf{X}_{\text{RCEC}}$ as the optimal initial capacity under full-step upgrading, define the profit loss of using only k-step upgrading as

$$\Delta_{k-step} = \left| \frac{\Pi_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}}) - \Pi^k_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}})}{\Pi_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}})} \right| * 100\%, \qquad k = 0, 1, 2, 3,$$

where $\Pi^k_{\text{RCEC}}(\mathbf{X}_{\text{RCEC}})$ is the revenue from the k-step upgrading. The statistics are presented in Table EC.3. We can see that the magnitudes of profit losses are still generally much larger than those for $\Delta_{\mathbf{X}_{\text{CB}}}$ and $\Delta_{\mathbf{X}_{\text{NV}}}$ (given in Table 6).

	Mean	Std.	Median	90%-percentile	Max.
Δ_{0-step}	4.29	4.28	2.85	9.36	31.80
Δ_{1-step}	1.08	1.39	0.45	3.37	7.72
Δ_{2-step}	0.33	0.56	0.091	1.04	3.54
Δ_{3-step}	0.10	0.25	0.01	0.29	1.98

Table EC.3 Profit loss of suboptimal allocation with k-step upgrading.

Appendix E: Additional Details

The following lemma shows the relation between $\mathbf{N}$ and its effective state $\hat{\mathbf{N}}$.

Lemma EC.6. *Suppose* $\hat{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_N)$ *is the effective state of* $\mathbf{N} = (n_1, \cdots, n_N)$, *then* $\sum_{s=i}^{j} \hat{n}_s \le \sum_{s=i}^{j} n_s$ *if* $\hat{n}_i > 0$, *and* $\sum_{s=i}^{j} \hat{n}_s \ge \sum_{s=i}^{j} n_s$ *if* $\hat{n}_{j+1} > 0$. *Especially,* $\sum_{s=i}^{N} \hat{n}_s \ge \sum_{s=i}^{N} n_s$.

Proof. The proof follows from the definition of the effective state. For any class k ($1 \le k \le N$), when applying the greedy upgrading to $\mathbf{N}$, there is no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ if $\hat{n}_k > 0$, and such an upgrade may exist if $\hat{n}_k \le 0$. Hence, $\sum_{s=i}^{j} \hat{n}_s \le \sum_{s=i}^{j} n_s$ if $\hat{n}_i > 0$, and $\sum_{s=i}^{j} \hat{n}_s \ge \sum_{s=i}^{j} n_s$ if $\hat{n}_{j+1} > 0$. The same argument shows that $\sum_{s=i}^{N} \hat{n}_s \ge \sum_{s=i}^{N} n_s$. ☐

The next proposition shows that separation can be preserved under the effective state operation.

Proposition EC.1. *Suppose* $\hat{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_N)$ *is the effective state of* $\mathbf{N} = (n_1, \cdots, n_N)$. *For any demand realization* $\mathbf{D}$, *class* i ($i < N$) *is the lowest class which is separable in* $\mathbf{N} - \mathbf{D}$ *if and only if* i *is the lowest class which is separable in* $\hat{\mathbf{N}} - \mathbf{D}$.

Proof. Suppose class i is the lowest separable class in $\mathbf{N} - \mathbf{D}$ but is not separable in $\hat{\mathbf{N}} - \mathbf{D}$. Then, there exists class a $k \le i$ such that $\sum_{s=k}^{i} (\hat{n}_s - d_s) > 0$. First, we must have $k < i$; otherwise, $n_i \ge \hat{n}_i > d_i \ge 0$, which means class i is not separable in $\mathbf{N} - \mathbf{D}$ and is a contradiction. Given $k < i$, without loss of generality, we assume k is the lowest class with $\sum_{s=k}^{i} (\hat{n}_s - d_s) > 0$, which also implies $\hat{n}_k > d_k \ge 0$ since $\sum_{s=k+1}^{i} (\hat{n}_s - d_s) \le 0$. Thus, $\sum_{s=k}^{i} n_s \ge \sum_{s=k}^{i} \hat{n}_s$ by Lemma EC.6, and

$\sum_{s=k}^{i}(n_s-d_s)\geq\sum_{s=k}^{i}(\hat{n}_s-d_s)>0$ which contradicts the assumption of class i being separable in $\mathbf{N}-\mathbf{D}$. Hence, class i must be separable in $\hat{\mathbf{N}}-\mathbf{D}$ as well.

Next, we prove that i is the lowest separable class in $\hat{\mathbf{N}}-\mathbf{D}$. Suppose to the contrary that class $i'>i$ is the lowest separable class in $\hat{\mathbf{N}}-\mathbf{D}$, i.e., $\sum_{s=k}^{i'}(\hat{n}_s-d_s)\leq 0$ for all classes k $(1\leq k\leq i')$. Then, $\hat{n}_{i'+1}-d_{i'+1}>0$; otherwise, $i'+1$ will be the lowest separable class. Because class i is the lowest separable class in $\mathbf{N}-\mathbf{D}$ and $i'>i$, there exists class $r\leq i'$ such that $\sum_{s=r}^{i'}(n_s-d_s)>0$. Given $\hat{n}_{i'+1}>d_{i'+1}\geq 0$ and Lemma EC.6, there is $\sum_{s=r}^{i'}\hat{n}_s\geq\sum_{s=r}^{i'}n_s$ and $\sum_{s=r}^{i'}(\hat{n}_s-d_s)\geq\sum_{s=r}^{i'}(n_s-d_s)>0$, which is a contradiction since $\sum_{s=r}^{i'}(\hat{n}_s-d_s)\leq 0$. Therefore, class i is the lowest separable class in $\hat{\mathbf{N}}-\mathbf{D}$.

The necessary condition can be similarly proved. This completes the proof. $\square$

For any demand realization $\mathbf{D}$ in period t $(1\leq t\leq T)$, let $\hat{\mathbf{N}}$ be the effective state of $\mathbf{N}$, Proposition EC.2 gives the relation between the outcomes of applying the PSR policy to initial states $(\mathbf{N},\mathbf{D})$ and $(\hat{\mathbf{N}},\mathbf{D})$.

PROPOSITION EC.2. *Suppose* $\hat{\mathbf{N}}=(\hat{n}_1,\cdots,\hat{n}_N)$ *is the effective state of* $\mathbf{N}=(n_1,\cdots,n_N)$*, and the PSR policy solves the general upgrading problem in period* t $(1\leq t\leq T)$*. For any demand realization* $\mathbf{D}$ *in period* t*, let* $\mathbf{N}'=(n'_1,\cdots,n'_N)$ *and* $\hat{\mathbf{N}}'=(\hat{n}'_1,\cdots,\hat{n}'_N)$ *be the effective states of the outcomes of applying the PSR policy to* $(\mathbf{N},\mathbf{D})$ *and* $(\hat{\mathbf{N}},\mathbf{D})$*, respectively. Then,* $\mathbf{N}'=\hat{\mathbf{N}}'$ *if classes* $(1,\cdots,N)$ *are not separable under* $\mathbf{N}-\mathbf{D}$*. Especially,* $\mathbf{N}'$ *and* $\hat{\mathbf{N}}'$ *are the outcomes of applying the PSR policy to* $(\mathbf{N},\mathbf{D})$ *and* $(\hat{\mathbf{N}},\mathbf{D})$ *in period* T*.*

Proof. From Proposition EC.1, classes $(1,\cdots,N)$ are also not separable under $\hat{\mathbf{N}}-\mathbf{D}$.

First, we must have $\mathbf{N}'\geq 0$. Suppose to the contrary that class k is the highest class with $n'_k<0$. Since $\mathbf{N}'$ is the effective state of the outcome of applying the PSR to $(\mathbf{N},\mathbf{D})$, there is $n'_1=\cdots=n'_{k-1}=0$. Note that any allocation decision is a transfer between two classes, which is true in both the PSR and the greedy upgrading. Thus, we have $0>\sum_{s=r}^{k}n'_s\geq\sum_{s=r}^{k}(n_s-d_s)$ for any class $r<k$, where the second inequality is strict if there is any upgrade between classes $(1,\cdots,r-1)$ and $(r,\cdots,k)$ when applying the PSR or generating the effective state. Hence, class k is separable, which contradicts the assumption. Similarly, we know $\hat{\mathbf{N}}'\geq 0$.

Next, we show $\mathbf{N}'=\hat{\mathbf{N}}'$. Let class k be the lowest class such that $\hat{n}'_k\neq n'_k$. From the above argument, we have $\sum_{s=k}^{N}n'_s=\sum_{s=k}^{N}(n_s-d_s)$ if there is no upgrade between classes $(1,\cdots,k-1)$ and $(k,\cdots,N)$ in either solving $(\mathbf{N},\mathbf{D})$ by the PSR or generating the effective state $\mathbf{N}'$. Furthermore, such an upgrade exists only if $n'_k=0$ by the optimality of the PSR and the definition of the greedy upgrading, in which case $\sum_{s=k}^{N}n'_s>\sum_{s=k}^{N}(n_s-d_s)$. The same argument can be applied to $(\hat{\mathbf{N}},\mathbf{D})$. With these observations, we derive contradictions for all possible cases.

1. $\sum_{s=k}^{N}\hat{n}_s=\sum_{s=k}^{N}n_s$.

(a) $\hat{n}'_k > 0, n'_k > 0$: For both initial states $(\mathbf{N}, \mathbf{D})$ and $(\hat{\mathbf{N}}, \mathbf{D})$, there is no upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ in either applying the PSR or generating the effective state. Moreover, $\hat{n}_k \geq \hat{n}'_k > 0$ implies that $\sum_{s=k}^N \hat{n}_s = \sum_{s=k}^N n_s$ by Lemma EC.6. Thus,

$$\sum_{s=k}^N (\hat{n}_s - d_s) = \sum_{s=k}^N \hat{n}'_s \neq \sum_{s=k}^N n'_s = \sum_{s=k}^N (n_s - d_s),$$

which is a contradiction;

(b) $\hat{n}'_k > n'_k = 0$: Similar to the previous case, since $\hat{n}_k \geq \hat{n}'_k > 0$, then

$$\sum_{s=k}^N \hat{n}'_s = \sum_{s=k}^N (\hat{n}_s - d_s) = \sum_{s=k}^N (n_s - d_s) \leq \sum_{s=k}^N n'_s,$$

which violates the assumption of class k;

(c) $n'_k > \hat{n}'_k = 0$: From Lemma EC.6, we similarly have

$$\sum_{s=k}^N \hat{n}'_s \geq \sum_{s=k}^N (\hat{n}_s - d_s) \geq \sum_{s=k}^N (n_s - d_s) = \sum_{s=k}^N n'_s,$$

which is also a contradiction;

2. $\sum_{s=k}^N \hat{n}_s > \sum_{s=k}^N n_s$. In this case, $\hat{n}_k = 0$ by the definition of the effective state $\hat{\mathbf{N}}$. Since $\hat{n}'_k \geq 0$ from the previous discussion, there is $\hat{n}_k = \hat{n}'_k = 0$. Meanwhile, $n'_k \neq \hat{n}'_k$ by the assumption of k. From $n'_k \geq 0$, we must have $n'_k > \hat{n}'_k = 0$, and

$$\sum_{s=k}^N \hat{n}'_s \geq \sum_{s=k}^N (\hat{n}_s - d_s) > \sum_{s=k}^N (n_s - d_s) = \sum_{s=k}^N n'_s,$$

where the first inequality is from the fact that there might be upgrade between classes $(1, \cdots, k-1)$ and $(k, \cdots, N)$ while solving $(\hat{\mathbf{N}}, \mathbf{D})$ by the PSR and generating the effective state $\hat{\mathbf{N}}'$. This is a contradiction since $n'_k > \hat{n}'_k$ and $\sum_{s=k+1}^N n'_s = \sum_{s=k+1}^N \hat{n}'_s$ by assumption of k.

Therefore, $\mathbf{N} = \hat{\mathbf{N}}'$. Note that the PSR optimally solves the general upgrading problem with protection levels being 0 in period T by Lemma EC.3. Since the greedy upgrading is equivalent to the PSR with protection levels being 0, we know $\mathbf{N}'$ and $\hat{\mathbf{N}}'$ are the outcomes of applying the PSR policy to $(\mathbf{N}, \mathbf{D})$ and $(\hat{\mathbf{N}}, \mathbf{D})$ in period T, which completes the proof. ☐

Lemma EC.7 considers a general upgrading problem with special states in period T, which can be used to simplify the proof of Lemma EC.3.

LEMMA EC.7. *Consider an N-class general upgrading problem with states $\mathbf{N} = (n_1, \cdots, n_N)$ and demand realization $\mathbf{D}$ in period T. Suppose classes $(1, \cdots, N)$ are not separable based on $\mathbf{N} - \mathbf{D}$, $(n_{i+1}, \cdots, n_{j-1}) \leq 0$ and $n_j < 0$. Then, $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})$ and $\Delta_{ij}^{-+}\Theta^T(\mathbf{N}|\mathbf{D})$ are independent of the values of $(n_j, \cdots, n_N)$.*

Proof. Since $\Theta^T(\mathbf{N}|\mathbf{D})$ is piecewise linear and concave (see Murty 1983), both $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})$ and $\Delta_{ij}^{-+}\Theta^T(\mathbf{N}|\mathbf{D})$ exist.

We focus on the proof of Δ_{ij}^{+-}, and the same argument applies to Δ_{ij}^{-+}. We consider the dual form of the general upgrading problem with initial state $(\mathbf{N},\mathbf{D})$, and let the dual variables be $(\lambda_1,\cdots,\lambda_N)$, where λ_i corresponds to the constraint of class i. The dual problem is

$$\begin{aligned} \min_{(\lambda_1,\cdots,\lambda_N)\geq 0} \quad & \sum_{s=1}^{N}|n_s-d_s|\lambda_s \\ \text{s.t.} \quad & \lambda_s+\lambda_r \geq \alpha_{sr}, \qquad s,r\in\{s,r|(\mathbf{N}-\mathbf{D})_s\geq 0,(\mathbf{N}-\mathbf{D})_r<0,\ 1\leq s<r\leq N\}. \end{aligned} \tag{EC.23}$$

1. $n_i \geq 0$: By Linear Programming theory, there is

   $$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D}) = \begin{cases} \lambda_i^*+\lambda_j^*-g_j, & \text{if } n_i\geq d_i \\ -\lambda_i^*+\lambda_j^*+\alpha_{ii}-g_j, & \text{if } n_i<d_i, \end{cases}$$

   where $\lambda^*=(\lambda_1^*,\cdots,\lambda_N^*)$ is optimal in the dual problem (EC.23).

   (a) $n_i>d_i$: Given classes $(1,\cdots,N)$ are not separable, we have $y_{kj}^*>0$ for some class k $(1\leq k\leq i)$. By the complementary slackness in the linear program, $\lambda_k^*+\lambda_j^*=\alpha_{kj}$. Assume without loss of generality that $i+1$ is the highest class s $(i+1\leq s\leq j-1)$ with $n_s-d_s<0$. Since it is optimal to first use class i's remaining capacity n_i-d_i to satisfy demands from $(i+1,\cdots,j)$, there is $y_{i,i+1}^*>0$, which implies $\lambda_i^*+\lambda_{i+1}^*=\alpha_{i,i+1}$. By examining constraints $\lambda_i^*+\lambda_j^*\geq\alpha_{ij}$, $\lambda_k^*+\lambda_{i+1}^*\geq\alpha_{k,i+1}$ in the dual problem (EC.23), as well as the assumption $\alpha_{kj}+\alpha_{i,i+1}=\alpha_{ij}+\alpha_{k,i+1}$, we have $\lambda_i^*+\lambda_j^*=\alpha_{ij}$ and $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})=\lambda_i^*+\lambda_j^*-g_j=\alpha_{ij}-g_j$. Note that $y_{ij}>0$ if $n_s-d_s=0$ for all classes s $(i+1\leq s\leq j-1)$, which implies $\lambda_i^*+\lambda_j^*-g_j=\alpha_{ij}-g_j$.

   (b) $n_i<d_i$: The non-separable assumption implies that there exist classes r $(r<i)$ and k $(k<i)$ such that $y_{ri}^*>0$ and $y_{kj}^*>0$. Thus, $\lambda_k^*+\lambda_j^*=\alpha_{kj}$ and $\lambda_r^*+\lambda_i^*=\alpha_{ri}$. We similarly have $\lambda_r^*+\lambda_j^*=\alpha_{rj}$ and $\lambda_k^*+\lambda_i^*=\alpha_{ki}$ by using the constraints in (EC.23) and the assumption $\alpha_{kj}+\alpha_{ri}=\alpha_{rj}+\alpha_{ki}$. Thus, $-\lambda_i^*+\lambda_j^*=\alpha_{rj}-\alpha_{ri}$ and $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})=-\lambda_i^*+\lambda_j^*+\alpha_{ii}-g_j=\alpha_{ij}-g_j$.

   Since $\Theta^T(\mathbf{N}|\mathbf{D})$ is piecewise linear in n_i and n_j, then $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})=\alpha_{ij}-g_j$ when $n_i\geq 0$.

2. $n_i<0$: In this case,

   $$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})=-\lambda_i^*+\lambda_j^*+g_i-g_j.$$

   Note that this is similar to the case when $0\leq n_i<d_i$. Hence, $-\lambda_i^*+\lambda_j^*=\alpha_{rj}-\alpha_{ri}$ and $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})=r_j-r_i$.

Hence, $\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})$ is independent of the values of $(n_j,\cdots,n_N)$, which concludes the proof. □

Suppose the PSR is optimal in period t. Then similar to Proposition EC.2, the following proposition shows the relation between the outcomes of $\mathbf{N}$ and its effective state $\hat{\mathbf{N}}$ after applying the PSR given any demand realization $\mathbf{D}$.

Proposition EC.3. *Suppose $\hat{\mathbf{N}} = (\hat{n}_1, \cdots, \hat{n}_N)$ is the effective state of $\mathbf{N} = (n_1, \cdots, n_N)$. If the PSR policy solves the general upgrading problem in period t, and the protection levels in period t have the independence property. For any demand realization $\mathbf{D}$, let $\mathbf{N}' = (n'_1, \cdots, n'_N)$ and $\hat{\mathbf{N}}' = (\hat{n}'_1, \cdots, \hat{n}'_N)$ be the outcomes of applying the PSR policy to $(\mathbf{N}, \mathbf{D})$ and $(\hat{\mathbf{N}}, \mathbf{D})$, respectively.*

Let k be the highest class in $\mathbf{N}'$ such that $(n'_k, \cdots, n'_N) \geq 0$ and $n'_k > 0$, where $k = N+1$ if such a class does not exist in $\mathbf{N}'$. $\hat{k}$ is similarly defined in $\hat{\mathbf{N}}'$. Then, $k = \hat{k}$ and $(\mathbf{N}')_{k,\cdots,N} = (\hat{\mathbf{N}}')_{k,\cdots,N}$ if classes $(1, \cdots, N)$ are not separable under $\mathbf{N} - \mathbf{D}$.

Proof. We first show $(\hat{\mathbf{N}}')_{k,\cdots,N} = (\mathbf{N}')_{k,\cdots,N}$. Let i $(k \leq i \leq N)$ be the lowest class such that $\hat{n}'_i \neq n'_i \geq 0$. There are three cases.

1. $\hat{n}'_i > n'_i \geq 0$: Since capacity i may be used when applying the PSR, we have $\hat{n}_i \geq \hat{n}'_i > 0$, which implies $\sum_{s=i}^N \hat{n}_s = \sum_{s=i}^N n_s$ by Lemma EC.6. Because $\hat{n}'_i > 0$, there is no upgrade from classes $(1, \cdots, i-1)$ to $(i, \cdots, N)$ when applying the PSR to $(\hat{\mathbf{N}}, \mathbf{D})$, and $\sum_{s=i}^N (\hat{n}_s - d_s) = \sum_{s=i}^N \hat{n}'_s$. On the other hand, $n'_i \geq 0$ implies that there could be upgrade from classes $(1, \cdots, i-1)$ to $(i, \cdots, N)$ when solving $(\mathbf{N}, \mathbf{D})$, thus $\sum_{s=i}^N n'_s \geq \sum_{s=i}^N (n_s - d_s)$. From the above, we have
$$\sum_{s=i}^N n'_s \geq \sum_{s=i}^N (n_s - d_s) = \sum_{s=i}^N (\hat{n}_s - d_s) = \sum_{s=i}^N \hat{n}'_s.$$
This is a contradiction given the assumption of class i.
2. $n'_i > \hat{n}'_i$ and $n'_i > 0$: In this case, $n'_i > 0$ implies that there is no upgrade from classes $(1, \cdots, i-1)$ to $(i, \cdots, N)$ when applying the PSR to $(\mathbf{N}, \mathbf{D})$. Thus, $\sum_{s=i}^N n'_s = \sum_{s=i}^N (n_s - d_s)$. However, there could be upgrade between classes $(1, \cdots, i-1)$ and $(i, \cdots, N)$ when generating $\hat{\mathbf{N}}$ as well as applying the PSR to $(\hat{\mathbf{N}}, \mathbf{D})$, thus
$$\sum_{s=i}^N n'_i = \sum_{s=i}^N (n_s - d_s) \leq \sum_{s=i}^N (\hat{n}_s - d_s) \leq \sum_{s=i}^N \hat{n}'_s, \tag{EC.24}$$
which is a contradiction.
3. $n'_i = 0 > \hat{n}'_i$ and $i > k$: From (EC.24), we only need to consider the case when $\sum_{s=i}^N n'_i > \sum_{s=i}^N (n_s - d_s)$, i.e., there is upgrade from classes $(1, \cdots, i-1)$ to $(i, \cdots, N)$ when applying the PSR to $(\mathbf{N}, \mathbf{D})$. Without loss of generality, we assume that $i-1$ is the highest class that upgrades the demands in classes $(i, \cdots, N)$ under initial state $(\mathbf{N}, \mathbf{D})$, and l $(l \geq i)$ is the lowest class being upgraded by capacity $i-1$. Since there is no upgrade from classes $(1, \cdots, i-2)$ to $(i-1, \cdots, N)$ when solving $(\mathbf{N}, \mathbf{D})$, similar to (EC.24), there is
$$\sum_{s=i-1}^N n'_i = \sum_{s=i-1}^N (n_s - d_s) \leq \sum_{s=i-1}^N (\hat{n}_s - d_s) \leq \sum_{s=i-1}^N \hat{n}'_s. \tag{EC.25}$$

Since $n'_i = 0 > \hat{n}'_i$ and $(\hat{\mathbf{N}}')_{i+1,\cdots,N} = (\mathbf{N}')_{i+1,\cdots,N} \geq 0$ by assumption of class i, (EC.25) implies $\hat{n}'_{i-1} > n'_{i-1} \geq 0$. Moreover, $\hat{n}_{i-1} > 0$ if $\hat{n}'_{i-1} > 0$.

Next, we show that the profit can be increased by upgrading demand i by capacity $i-1$ under $(\hat{\mathbf{N}}, \mathbf{D})$, which violates the optimality assumption of the PSR. Since $\hat{n}'_{i-1} > 0$ and the assumption of class $i-1$, there is no upgrade between classes $(1, \cdots, i-2)$ and $(i-1, \cdots, N)$ when generating the effective state $\hat{\mathbf{N}}$ as well as applying the PSR to both $(\hat{\mathbf{N}}, \mathbf{D})$ and $(\mathbf{N}, \mathbf{D})$. From Proposition EC.2, given classes $(1, \cdots, N)$ are not separable under $\mathbf{N} - \mathbf{D}$, the effective states of $(\hat{\mathbf{N}}')_{1,\cdots,i-2}$ are the same as those of $(\mathbf{N}')_{1,\cdots,i-2}$.

If $l = i$, from the independence property of the protection levels, $p_{i-1,i}$ is the same for both $(\hat{\mathbf{N}}, \mathbf{D})$ and $(\mathbf{N}, \mathbf{D})$. Because $\hat{n}'_{i-1} > n'_{i-1}$ and capacity $i-1$ upgrades demand i under $(\mathbf{N}, \mathbf{D})$, it is also optimal to upgrade demand i by capacity $i-1$ under $(\hat{\mathbf{N}}, \mathbf{D})$.

If $l > i$, we have $(\hat{\mathbf{N}}')_{i+1,\cdots,l} = (\mathbf{N}')_{i+1,\cdots,l}$ by the assumption of class i. Moreover, $(\mathbf{N}')_{i+1,\cdots,l} = 0$ since capacity $i-1$ upgrades demand l under initial state $(\mathbf{N}, \mathbf{D})$, and n'_{i-1} is the remaining capacity after such upgrading. From the PSR, there is

$$\begin{aligned}
\alpha_{i-1,l} \geq & \Delta^{+-}_{i-1,l} \Theta^{t+1}(\mathbf{N}') \\
= & \Delta^{+-}_{i-1,l} \Theta^{t+1}\left((\mathbf{N}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\mathbf{N}')_{l+1,\cdots,N}\right) \\
= & \Delta^{+-}_{i-1,l} \Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\
\geq & r_l - r_i + \partial^+_{i-1} \Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\
& - \partial^-_i \Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\
= & r_l - r_i + \Delta^{+-}_{i-1,i} \Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}),
\end{aligned} \tag{EC.26}$$

where the second equality follows from the independence property of Θ^{t+1} and the fact that the effective states of $(\mathbf{N}')_{1,\cdots,i-2}$ and $(\hat{\mathbf{N}}')_{1,\cdots,i-2}$ are the same, and the last inequality is because of Lemma EC.2.

Since $\alpha_{i-1,i} - \alpha_{i-1,l} = r_i + g_i - r_l - g_l$, where $g_i > g_l$, $\hat{n}'_{i-1} > n'_{i-1}$, and Θ^{t+1} is concave, we have

$$\begin{aligned}
\alpha_{i-1,i} > & \Delta^{+-}_{i-1,i} \Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, n'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\
\geq & \Delta^{+-}_{i-1,i} \Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, \hat{n}'_{i-1}, n'_{i-1} - \hat{n}'_{i-1}, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}) \\
= & \Delta^{+-}_{i-1,i} \Theta^{t+1}((\hat{\mathbf{N}}')_{1,\cdots,i-2}, \hat{n}'_{i-1}, \hat{n}'_i, 0, \cdots, 0, (\hat{\mathbf{N}}')_{l+1,\cdots,N}).
\end{aligned} \tag{EC.27}$$

Thus, the profit can be increased by upgrading demand i with capacity $i-1$ under $(\hat{\mathbf{N}}, \mathbf{D})$, which contradicts the optimality assumption of the PSR policy.

Hence, $(\hat{\mathbf{N}}')_{k,\cdots,N} = (\mathbf{N}')_{k,\cdots,N}$. Similarly, we know $(\mathbf{N}')_{\hat{k},\cdots,N} = (\hat{\mathbf{N}}')_{\hat{k},\cdots,N}$, which concludes the proof. □

E.1. Monotonicity

To prove the monotonicity result in Proposition 5, we start with a basic property.

Under certain conditions, the following lemma states that the marginal values, $\Delta_{ij}^{+-}\Theta^t$ $(i<j)$ and $\Delta_{ij}^{-+}\Theta^t$, remain the same if capacity k $(k<i)$ is used to "optimally" upgrade the backlogged demand i. Note that such an upgrade can go beyond class k as long as there is unmet demand i.

Lemma EC.8. *Suppose* $(\hat{n}_1,\cdots,\hat{n}_{i-1})$ *is the effective state of* $(n_1,\cdots,n_{i-1})$*, and there exists class* k $(1\le k<i)$ *such that* $\hat{n}_k>0$ *and* $\hat{n}_{k+1}=\cdots=\hat{n}_{i-1}=0$*. If* $(n_i,\cdots,n_j)\le 0$*,* $\delta>0$*, and* $n_i+\delta\le 0\le \hat{n}_k-\delta$*, then*

$$\Delta_{ij}^{+-}\Theta^t(n_1,\cdots,n_N)=\Delta_{ij}^{+-}\Theta^t(\hat{n}_1,\cdots,\hat{n}_{k-1},\hat{n}_k-\delta,0,\cdots,0,n_i+\delta,n_{i+1},\cdots,n_N) \tag{EC.28}$$

and

$$\Delta_{ij}^{-+}\Theta^t(n_1,\cdots,n_N)=\Delta_{ij}^{-+}\Theta^t(\hat{n}_1,\cdots,\hat{n}_{k-1},\hat{n}_k-\delta,0,\cdots,0,n_i+\delta,n_{i+1},\cdots,n_N). \tag{EC.29}$$

Proof. It is sufficient to prove the equality in (EC.28). From Proposition 2, there is

$$\Delta_{ij}^{+-}\Theta^t(n_1,\cdots,n_N)=\Delta_{ij}^{+-}\Theta^t(\hat{n}_1,\cdots,\hat{n}_k,0,\cdots,0,n_i,\cdots,n_N).$$

Thus, for any demand realization $\mathbf{D}$ in period t, we use induction to show

$$\begin{aligned}&\Delta_{ij}^{+-}\Theta^t(\hat{n}_1,\cdots,\hat{n}_k,0,\cdots,0,n_i,\cdots,n_N|\mathbf{D})\\ =&\Delta_{ij}^{+-}\Theta^t(\hat{n}_1,\cdots,\hat{n}_{k-1},\hat{n}_k-\delta,0,\cdots,0,n_i+\delta,n_{i+1},\cdots,n_N|\mathbf{D})\end{aligned} \tag{EC.30}$$

under the conditions given in Lemma EC.8. To simplify our notations, let

$$\begin{aligned}\mathbf{N}&=(\hat{n}_1,\cdots,\hat{n}_k,0,\cdots,0,n_i,\cdots,n_N)\\ \bar{\mathbf{N}}&=(\hat{n}_1,\cdots,\hat{n}_{k-1},\hat{n}_k-\delta,0,\cdots,0,n_i+\delta,n_{i+1},\cdots,n_N).\end{aligned}$$

In period T, let r^* $(1\le r^*\le k)$ be the lowest class such that $n_i+\sum_{s=r^*}^{k}\hat{n}_s\ge\sum_{s=r^*}^{i}d_s$, i.e., r^* is the lowest class that satisfies the last unit of demand i. We analyze (EC.30) based on following cases.

1. r^* does not exist: Then $n_i+\sum_{s=r}^{k}\hat{n}_s<\sum_{s=r}^{i}d_s$ and $n_i-\delta+\sum_{s=r}^{k}\hat{n}_s<-\delta+\sum_{s=r}^{i}d_s$ for all class r $(1\le r\le k)$. After applying the PSR, there is unmet demand i in both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$. Thus, given $(n_{i+1},\cdots,n_j)\le 0$, we have

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})=\Delta_{ij}^{+-}\Theta^T(\bar{\mathbf{N}}|\mathbf{D})=g_i-g_j.$$

2. r^* does exist: For both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$, since $n_i-\delta+\sum_{s=r^*}^k \hat{n}_s \geq -\delta+\sum_{s=r^*}^i d_s$, the last unit of demand i is fulfilled by capacity r^*. And the states of classes $(r^*,\cdots,i)$ after the last unit of demand i being satisfied are $(n_i+\sum_{s=r^*}^k \hat{n}_s-\sum_{s=r^*}^i d_s,0,\cdots,0)$. Hence,

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}|\mathbf{D})=\Delta_{ij}^{+-}\Theta^T(\bar{\mathbf{N}}|\mathbf{D})=g_i-\alpha_{r^*i}+\frac{\partial}{\partial n_i^+}\Theta^T(\tilde{\mathbf{N}}|\tilde{\mathbf{D}})-\frac{\partial}{\partial n_j^-}\Theta^T(\tilde{\mathbf{N}}|\tilde{\mathbf{D}}),$$

where $\tilde{\mathbf{N}}=(\hat{n}_1,\cdots,\hat{n}_{r^*-1},n_i+\sum_{s=r^*}^k \hat{n}_s-\sum_{s=r^*}^i d_s,0,\cdots,0,n_{i+1},\cdots,n_N)$ and $\tilde{\mathbf{D}}=((\mathbf{D})_{1,\cdots,r^*-1},0,\cdots,0,(\mathbf{D})_{i+1,\cdots,N})$.

Hence, (EC.30) holds in period T.

In period $t<T$, we apply the PSR policy to the general upgrading problem with initial states $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$, and denote $\mathbf{N}'$ and $\bar{\mathbf{N}}'$ as the corresponding outcomes. We examine (EC.30) based on the states of class i in $\mathbf{N}'$ and $\bar{\mathbf{N}}'$.

1. $(\mathbf{N}')_i=(\bar{\mathbf{N}}')_i=0$: From the above analysis, the last unit of demand i is satisfied by class r^* in both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$, and we assume $r^*=k$ without loss of generality. Hence,

$$\Delta_{ij}^{+-}\Theta^t(\mathbf{N}|\mathbf{D})=\Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}|\mathbf{D})=g_i-\alpha_{ki}+\frac{\partial}{\partial n_i^+}\Theta^t(\tilde{\mathbf{N}}|\tilde{\mathbf{D}})-\frac{\partial}{\partial n_j^-}\Theta^t(\tilde{\mathbf{N}}|\tilde{\mathbf{D}}),$$

where $\tilde{\mathbf{N}}=(\hat{n}_1,\cdots,\hat{n}_{k-1},n_i+\hat{n}_k-\sum_{s=k}^i d_s,0,\cdots,0,n_{i+1},\cdots,n_N)$ and $\tilde{\mathbf{D}}=((\mathbf{D})_{1,\cdots,k-1},0,\cdots,0,(\mathbf{D})_{i+1,\cdots,N})$.

2. $(\mathbf{N}')_i<0$ and $(\bar{\mathbf{N}}')_i<0$: If there is no class r^* $(1\leq r^*\leq k)$ such that $n_i+\sum_{s=r^*}^k \hat{n}_s\geq \sum_{s=r^*}^i d_s$, demand i and j will never be satisfied in the remaining periods for both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$, which means

$$\Delta_{ij}^{+-}\Theta^t(\mathbf{N}|\mathbf{D})=(T-t+1)(g_i-g_j)=\Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}|\mathbf{D}).$$

Hence, we only need to consider the case when class r^* does exist. In this case, we assume $r^*=k$ without loss of generality. From Proposition 2, since $(\mathbf{N}')_i<0$ and $(\bar{\mathbf{N}}')_i<0$, MP_{ij} will not affect the optimal allocation decisions in period t under both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$. Thus,

$$\begin{aligned}\Delta_{ij}^{+-}\Theta^t(\mathbf{N}|\mathbf{D})&=g_i-g_j+\frac{\partial}{\partial n_i^+}\Theta^{t+1}(\mathbf{N}')-\frac{\partial}{\partial n_j^-}\Theta^{t+1}(\mathbf{N}'),\\ \Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}|\mathbf{D})&=g_i-g_j+\frac{\partial}{\partial n_i^+}\Theta^{t+1}(\bar{\mathbf{N}}')-\frac{\partial}{\partial n_j^-}\Theta^{t+1}(\bar{\mathbf{N}}').\end{aligned} \tag{EC.31}$$

By the definition of class k, there is no upgrade between classes $(1,\cdots,k-1)$ and $(k,\cdots,N)$ when applying the PSR under both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$; otherwise, all capacity k should have been depleted before performing the aforementioned upgrade, which means there is no unmet demand i. Since the initial states of classes $(1,\cdots,k-1)$ are the same, the effective states of classes $(1,\cdots,k-1)$ in $\mathbf{N}'$ are the same as those in $\bar{\mathbf{N}}'$. Note that $(\mathbf{N}')_{i+1,\cdots,j}=(\bar{\mathbf{N}}')_{i+1,\cdots,j}=$

$(n_{i+1}-d_{i+1},\cdots,n_j-d_j)\le 0$ and $n_i+\hat{n}_k-\sum_{s=k}^{i} d_s>0$ by assumption. Applying the induction assumption, we have

$$\begin{aligned}
&\frac{\partial}{\partial n_i^+}\Theta^{t+1}(\bar{\mathbf{N}}')-\frac{\partial}{\partial n_j^-}\Theta^{t+1}(\bar{\mathbf{N}}')\\
=&\Delta_{ij}^{+-}\Theta^{t+1}\Big((\mathbf{N}')_{1,\cdots,k-1},n_i+\hat{n}_k+\delta-\sum_{s=k}^{i}d_s,0,\cdots,0,-\delta,(\mathbf{N}')_{i+1,\cdots,N}\Big)\\
=&\Delta_{ij}^{+-}\Theta^{t+1}\Big((\bar{\mathbf{N}}')_{1,\cdots,k-1},n_i+\hat{n}_k+\delta-\sum_{s=k}^{i}d_s,0,\cdots,0,-\delta,(\bar{\mathbf{N}}')_{i+1,\cdots,N}\Big)\\
=&\frac{\partial}{\partial n_i^+}\Theta^{t+1}(\bar{\mathbf{N}}')-\frac{\partial}{\partial n_j^-}\Theta^{t+1}(\bar{\mathbf{N}}'),
\end{aligned}\tag{EC.32}$$

where $0<\delta<-\max\big((\mathbf{N}')_i,\ (\bar{\mathbf{N}}')_i\big)$ and the second equality follows from Proposition 2. This is a contradiction. Hence, $\Delta_{ij}^{+-}\Theta^t(\mathbf{N}|\mathbf{D})=\Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}|\mathbf{D})$ from (EC.31) and (EC.32).

3. $(\mathbf{N}')_i=0$ and $(\bar{\mathbf{N}}')_i<0$: In this case, there exists a class r^*, which can be assumed as $r^*=k$ without loss of generality. Moreover, the last unit of demand i is upgraded by capacity k when the PSR solves $(\mathbf{N},\mathbf{D})$.

   Given $(\bar{\mathbf{N}}')_i<0$, we must have $(\bar{\mathbf{N}}')_k>(\mathbf{N}')_k\ge 0$ since the total unmet demand after parallel allocation in classes $(k,\cdots,i)$ is the same for both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$. When the last unit of demand i is upgraded by capacity k in $(\mathbf{N},\mathbf{D})$, from the PSR, the upgrading decisions between classes $(1,\cdots,k-1)$ and $(i+1,\cdots,N)$ have not been considered yet. At that moment, the effective state of classes $(1,\cdots,k-1)$ in $(\mathbf{N},\mathbf{D})$ is the same as that in $\bar{\mathbf{N}}'$ because there is also no upgrade between classes $(1,\cdots,k-1)$ and $(k,\cdots,N)$ in $(\bar{\mathbf{N}},\mathbf{D})$ when applying the PSR. Hence, the protection levels between class k and the lower classes are the same for both $(\mathbf{N},\mathbf{D})$ and $(\bar{\mathbf{N}},\mathbf{D})$ from Proposition 2. Let h $(k<h\le i)$ be the highest class with $(\bar{\mathbf{N}}')_h<0$. Similar to (EC.26) and (EC.27) in the proof of Proposition EC.3, we can show that the profit from solving $(\bar{\mathbf{N}},\mathbf{D})$ can be increased by upgrading demand h with capacity k. This contradicts the optimality of the PSR. Hence, this case cannot exist.

4. $(\mathbf{N}')_i<0$ and $(\bar{\mathbf{N}}')_i=0$: Similar to the previous case, this would lead to a contradiction.

Therefore, (EC.30) holds for any demand realization $\mathbf{D}$ in period t, and this completes the induction proof. □

The following lemma shows that the protection level p_{ij} $(1\le i<j\le N)$ in period $t-1$ is decreasing in the states of classes $(1,\cdots,,i-1)$ if the same monotonicity holds in period t.

LEMMA EC.9. *Consider an N-class upgrading problem in period t $(1\le t<T)$ with $(n_{i+1},\cdots,n_j)\le 0$. Let $\bar{\mathbf{N}}=\mathbf{N}+\epsilon\mathbf{e}_r$, where $1\le r<i$ and $\epsilon>0$. Then,*

$$\Delta_{ij}^{+-}\Theta^t(\mathbf{N})\ge\Delta_{ij}^{+-}\Theta^t(\bar{\mathbf{N}}),\qquad \Delta_{ij}^{-+}\Theta^t(\mathbf{N})\ge\Delta_{ij}^{-+}\Theta^t(\bar{\mathbf{N}})\tag{EC.33}$$

if the same inequality holds for Θ^{t+1}.

Proof. To prove (EC.33), it is sufficient to show

$$\Theta^t(\mathbf{N}_{ij}) - \Theta^t(\mathbf{N}) \geq \Theta^t(\bar{\mathbf{N}}_{ij}) - \Theta^t(\bar{\mathbf{N}}), \tag{EC.34}$$

where

$$\begin{aligned}
\mathbf{N}_{ij} &= (n_1, \cdots, n_{i-1}, n_i + 1, n_{i+1}, \cdots, n_{j-1}, n_j - 1, n_{j+1}, \cdots, n_N), \\
\bar{\mathbf{N}} &= (n_1, \cdots, n_{r-1}, n_r + 1, n_{r+1}, \cdots, n_N), \\
\bar{\mathbf{N}}_{ij} &= (n_1, \cdots, n_{r-1}, n_r + 1, n_{r+1}, \cdots, n_{i-1}, n_i + 1, n_{i+1}, \cdots, n_{j-1}, n_j - 1, n_{j+1}, \cdots, n_N).
\end{aligned}$$

In each period t, given any demand realization $\mathbf{D} = (d_1, \cdots, d_N)$, we next show

$$\Delta = \Theta^t(\mathbf{N}_{ij}|\mathbf{D}) - \Theta^t(\mathbf{N}|\mathbf{D}) \geq \Theta^t(\bar{\mathbf{N}}_{ij}|\mathbf{D}) - \Theta^t(\bar{\mathbf{N}}|\mathbf{D}) = \bar{\Delta}, \tag{EC.35}$$

which proves (EC.34).

To compare Δ and $\bar{\Delta}$, we consider upgrading decisions in period t. Denote $\mathbf{R}$ as the resulting states of classes $(1, \cdots, N)$ after applying the PSR under initial state $(\mathbf{N}, \mathbf{D})$, let h be the highest capacity class which upgrades the demand in classes $(i, \cdots, N)$, and l is the lowest demand class that is upgraded by classes $(1, \cdots, i)$. By the definition of classes h and l, we have $h \leq l$, and $h = l$ only if $h = i$ and $l = i$. From the PSR policy, there is neither unmet demand nor remaining capacity between classes h and l in $\mathbf{R}$, i.e. $(\mathbf{R})_{h+1, \cdots, l-1} = \mathbf{0}$. $\bar{\mathbf{R}}$, $\bar{h}$ and $\bar{l}$ are similarly defined under initial state $(\bar{\mathbf{N}}, \mathbf{D})$.

For any classes $1 \leq k < s \leq N$ in period t, the protection level p_{ks} defined in (8) are decreasing in $(n_1, \cdots, n_{k-1})$ since (EC.33) is true for Θ^{t+1}, thus upgrade is more likely to happen under initial state $(\bar{\mathbf{N}}, \mathbf{D})$ rather than $(\mathbf{N}, \mathbf{D})$, i.e., $l \leq \bar{l}$.

Switching from $\mathbf{N}$ $(\bar{\mathbf{N}})$ to $\mathbf{N}_{ij}$ $(\bar{\mathbf{N}}_{ij})$, we not only change the current revenues in period t, but also the result $\mathbf{R}$ $(\bar{\mathbf{R}})$, which is the initial states in period $t+1$. Denote $\mathbf{R}'$ and $\bar{\mathbf{R}}'$ as the outcomes after applying the PSR under $(\mathbf{N}_{ij}, \mathbf{D})$ and $(\bar{\mathbf{N}}_{ij}, \mathbf{D})$, respectively. Then,

$$\Delta = \delta + \Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}), \qquad \bar{\Delta} = \bar{\delta} + \Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}),$$

where δ and $\bar{\delta}$ are the corresponding differences of the current period revenues in period t under $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, respectively.

When the initial states change from $\mathbf{N}$ to $\mathbf{N}_{ij}$, there are four cases which differ in the allocation decisions in period t. Note that the analogy applies when initial states change form $\bar{\mathbf{N}}$ to $\bar{\mathbf{N}}_{ij}$. For simplicity, we assume without loss of generality that $(\mathbf{R})_{l+1} < 0$.

Case 1: An extra unit of demand l is satisfied when $l < j$.

Case 2: An extra unit of capacity h is passed along to period $t+1$.

Case 3: An extra unit of demand $l+1$ is satisfied when $l+1 < j$.

Case 4: An extra unit of demand j is satisfied if $l \geq j$.

Here, we explain the above cases in detail by recalling the "chain reaction" described in the proof of Proposition 2.

Case 1: There is unmet demand l in $\mathbf{R}$ in this case. Note that capacity h is the highest class that upgrades demand l under $(\mathbf{N}, \mathbf{D})$. And the upgrade between classes h and l is bounded because either there is no capacity h remaining or the protection level p_{hl} is reached. When the initial state changes from $\mathbf{N}$ to $\mathbf{N}_{ij}$, from the chain reaction, there is an additional unit of capacity h which will upgrade the remaining demand l.

Case 2: In this case, class l demand has been fully satisfied in $\mathbf{R}$; otherwise, the analysis of Case 1 gives a contradiction.

Case 3: If $l+1 < j$, similar to *Case 1*, it is possible that an additional unit of demand $l+1$ is upgraded by capacity h under $(\mathbf{N}_{ij}, \mathbf{D})$, in which case all demand l has been satisfied in $\mathbf{R}$.

Case 4: Suppose that k_j is the highest class that upgrades demand j under $(\mathbf{N}, \mathbf{D})$. Because increasing n_i simultaneously decreases n_j by the same amount, there will be an additional unit of both capacity k_j and unmet demand j from the chain reaction. From the PSR, it is optimal to upgrade such an additional demand j by capacity k_j, and the outcome $\mathbf{R}' = \mathbf{R}$ in this case.

To compare Δ and $\bar{\Delta}$, we start with Case 4, where $l \geq j$ and $\mathbf{R}' = \mathbf{R}$ from the above discussion.

1. $n_i < 0$: Suppose the last unit of demand i is upgraded by class k_i, then $\Delta = g_i - g_j - \alpha_{k_i i} + \alpha_{k_i j} = r_j - r_i$;
2. $0 \leq n_i < d_i$: Similar to the previous case, we have $\Delta = -g_j - \alpha_{k_i i} + \alpha_{k_i j} = r_j - r_i - g_i$;
3. $n_i \geq d_i$: Given the chain reaction, the overall effect is equivalent to upgrading demand j with capacity i. Then, $\Delta = -g_j + \alpha_{ij} = r_j - u_i$.

To summarize, if $l \geq j$, we have

$$\Delta = \begin{cases} r_j - r_i, & \text{if } n_i < 0 \\ r_j - r_i - g_i, & \text{if } 0 \leq n_i < d_i \\ r_j - u_i, & \text{if } n_i \geq d_i. \end{cases} \tag{EC.36}$$

Note that (EC.36) also holds for $\bar{\Delta}$ if $\bar{l} \geq j$. Therefore, $\Delta = \bar{\Delta}$ when $j \leq l \leq \bar{l}$.

Next, we compare Δ and $\bar{\Delta}$ when both $l < j$ and $\bar{l} < j$. We categorize different situations based on Case 1, Case 2 and Case 3 as follows:

1. Case 1 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$: Notice that class l here is similar to class j in (EC.36) in Case 4. Then, we have

$$\delta = \begin{cases} g_i - g_j + \alpha_{k_i l} - \alpha_{k_i i} = g_l - g_j - r_i + r_l, & \text{if } n_i < 0 \\ -g_j + \alpha_{k_i l} - \alpha_{k_i i} = -g_j - g_i + g_l - r_i + r_l, & \text{if } 0 \leq n_i < d_i \\ -g_j + \alpha_{il}, & \text{if } n_i \geq d_i, \end{cases} \tag{EC.37}$$

and

$$\bar{\delta} = \begin{cases} g_i - g_j + \alpha_{\bar{k}_i \bar{l}} - \alpha_{\bar{k}_i i} = g_{\bar{l}} - g_j - r_i + r_{\bar{l}}, & \text{if } n_i < 0 \\ -g_j + \alpha_{\bar{k}_i \bar{l}} - \alpha_{\bar{k}_i i} = -g_j - g_i + g_{\bar{l}} - r_i + r_{\bar{l}}, & \text{if } 0 \le n_i < d_i \\ -g_j + \alpha_{i\bar{l}}, & \text{if } n_i \ge d_i. \end{cases} \tag{EC.38}$$

Thus, $\delta - \bar{\delta} = r_l + g_l - r_{\bar{l}} - g_{\bar{l}}$.

Furthermore, $\bar{\mathbf{R}}' = \bar{\mathbf{R}}_{\bar{l}j}$ by the assumption of this case. We next show

$$\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}) = \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{R}}) = \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{N}}). \tag{EC.39}$$

From the assumption, there is no upgrade between classes $(1, \cdots, \bar{h}-1)$ and $(\bar{h}, \cdots, N)$ when applying the PSR under $(\bar{\mathbf{N}}, \mathbf{D})$, whose result is $\bar{\mathbf{R}}$. Thus, the effective states of classes $(1, \cdots, \bar{h}-1)$ in $\bar{\mathbf{R}}$ are the same as those in $\bar{\mathbf{N}}$ by Proposition EC.2. Moreover, note that $\bar{h}$ is the highest class upgrading demand $\bar{l}$ by assumption. Without loss of generality, we assume $\bar{h}$ is also the lowest class upgrading demand $\bar{l}$, then the effective state of classes $(\bar{h}, \cdots, \bar{l}-1)$ in $\bar{\mathbf{N}}$ is $((\bar{\mathbf{R}})_{\bar{h}} + y_{\bar{h}\bar{l}}, 0, \cdots, 0)$, where $y_{\bar{h}\bar{l}}$ is the upgrade between classes $\bar{h}$ and $\bar{l}$ under initial state $(\bar{\mathbf{N}}, \mathbf{D})$. Thus, classes $\bar{h}$ and $\bar{l}$ correspond to classes k and i in Lemma EC.8, which proves (EC.39).

Similarly, since $\mathbf{R}' = \mathbf{R}_{lj}$ in this case, there is

$$\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}) = \Theta^{t+1}(\mathbf{R}_{lj}) - \Theta^{t+1}(\mathbf{R}) = \Theta^{t+1}(\mathbf{N}_{lj}) - \Theta^{t+1}(\mathbf{N}). \tag{EC.40}$$

Moreover,

$$\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}) = \Theta^{t+1}(\mathbf{N}_{lj}) - \Theta^{t+1}(\mathbf{N}) \ge \Theta^{t+1}(\bar{\mathbf{N}}_{lj}) - \Theta^{t+1}(\bar{\mathbf{N}}) \tag{EC.41}$$

from the induction assumption.

To complete the proof in this case, from Lemma EC.2 and the fact that $l \le \bar{l}$, there is

$$\Theta^{t+1}(\bar{\mathbf{N}}_{lj}) - \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) \ge r_{\bar{l}} - r_l,$$

which implies $\Delta - \bar{\Delta} = \delta - \bar{\delta} + r_{\bar{l}} - r_l$ by (EC.39) and (EC.41). Since $\delta - \bar{\delta} = r_l + g_l - r_{\bar{l}} - g_{\bar{l}}$ by (EC.37) and (EC.38), we have $\Delta - \bar{\Delta} = g_l - g_{\bar{l}} \ge 0$.

In the remaining cases, we apply similar arguments to prove (EC.35). For simplicity, we will omit some details and only present the primary results.

2. Case 2 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$: We have

$$\delta = \begin{cases} g_i - g_j + (r_l + g_l - r_i - g_i) - \alpha_{hl} = -g_j - r_i + u_h, & \text{if } n_i < 0 \\ -g_j + (r_l + g_l - r_i - g_i) - \alpha_{hl} = -g_j - g_i - r_i + u_h, & \text{if } 0 \le n_i < d_i \\ -g_j + \alpha_{il} - \alpha_{hl} = -g_j - u_i + u_h, & \text{if } n_i \ge d_i \end{cases} \tag{EC.42}$$

and

$$\bar{\delta} = \begin{cases} g_i - g_j + (r_{\bar{l}} + g_{\bar{l}} - r_i - g_i) - \alpha_{\bar{h}\bar{l}} = -g_j - r_i + u_{\bar{h}}, & \text{if } n_i < 0 \\ -g_j + (r_{\bar{l}} + g_{\bar{l}} - r_i - g_i) - \alpha_{\bar{h}\bar{l}} = -g_j - g_i - r_i + u_{\bar{h}}, & \text{if } 0 \le n_i < d_i \\ -g_j + \alpha_{i\bar{l}} - \alpha_{\bar{h}\bar{l}} = -g_j - u_i + u_{\bar{h}}, & \text{if } n_i \ge d_i. \end{cases} \tag{EC.43}$$

Note that $\delta - \bar{\delta} = u_h - u_{\bar{h}}$ in all cases.

(a) $\bar{l} = l$: From the assumption, all backlogged demands in classes $(i, \cdots, l)$, which are the same for both initial states $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, have been satisfied in period t. Meanwhile, $\mathbf{D}$ is the same for both initial states in period t. Thus, the total demands satisfied are the same for both $(\mathbf{N}, \mathbf{D})$ and $(\bar{\mathbf{N}}, \mathbf{D})$, and we have $\bar{h} = h \ge r$ or $h \le \bar{h} \le r$.

By assumption, $\mathbf{R}' = \mathbf{R}_{hj}$ and $\bar{\mathbf{R}}' = \bar{\mathbf{R}}_{\bar{h}j}$ in this case, then

$$\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}) = \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}), \quad \Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R}) = \Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}). \tag{EC.44}$$

Moreover, we define $\tilde{\mathbf{R}}$ as follows:

$$\tilde{\mathbf{R}} = \begin{cases} \mathbf{R} + \mathbf{e}_r, & \text{if } r < h \\ \mathbf{R} + \mathbf{e}_h, & \text{if } r \ge h. \end{cases}$$

Note that $\tilde{\mathbf{R}} = \bar{\mathbf{R}}$ from the definition. If $r < h$, given $(\mathbf{R})_{h+1,\cdots,j-1} = (\mathbf{R}')_{h+1,\cdots,j-1} \le 0$, we have

$$\Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}) \ge \Theta^{t+1}(\tilde{\mathbf{R}}_{hj}) - \Theta^{t+1}(\tilde{\mathbf{R}}) \tag{EC.45}$$

from the induction assumption. On the other hand, if $r \ge h$, (EC.45) still holds because of the concavity in Proposition 1.

Since $\tilde{\mathbf{R}} = \bar{\mathbf{R}}$ and $h \le \bar{h}$, there is $\Theta^{t+1}(\tilde{\mathbf{R}}_{hj}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) \ge u_{\bar{h}} - u_h$ by Lemma EC.1. Therefore, from (EC.44) and (EC.45), we have

$$\Delta - \bar{\Delta} \ge \delta - \bar{\delta} + \Theta^{t+1}(\tilde{\mathbf{R}}_{hj}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) \ge 0,$$

where $\delta - \bar{\delta} = u_h - u_{\bar{h}}$ by (EC.42) and (EC.43).

(b) $l < \bar{l}$: From the above discussion of Case 2, $\bar{\mathbf{R}}'$ has one more unit of capacity $\bar{h}$ than $\bar{\mathbf{R}}$ after the chain reaction. Note that class $\bar{h}$ would have upgraded demand $\bar{l}$ if there exists unmet demand $\bar{l}$ under $\bar{\mathbf{R}}'$, which implies that the expected value of such a unit of capacity $\bar{h}$ is smaller than $\alpha_{\bar{h}\bar{l}}$. Thus,

$$\begin{aligned} &\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}}) = \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}) \\ =&\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}\bar{l}}) + \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}\bar{l}}) - \Theta^{t+1}(\bar{\mathbf{R}}) \\ \le&\alpha_{\bar{h}\bar{l}} + \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j}) - \Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}\bar{l}}), \end{aligned} \tag{EC.46}$$

Moreover, similar to (EC.39), we can apply Lemma EC.8 to (EC.46) as $(\bar{\mathbf{R}})_{\bar{h}+1,\cdots,\bar{l}}=0$, then

$$\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j})-\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}\bar{l}})=\Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j})-\Theta^{t+1}(\bar{\mathbf{N}})$$

and

$$\Theta^{t+1}(\bar{\mathbf{R}}')-\Theta^{t+1}(\bar{\mathbf{R}})\le\alpha_{\bar{h}\bar{l}}+\Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j})-\Theta^{t+1}(\bar{\mathbf{N}}). \tag{EC.47}$$

For initial state $(\mathbf{N},\mathbf{D})$, since $l<\bar{l}$ and $(\mathbf{R})_{l+1}<0$, after the chain reaction, there is an additional unit of capacity h in $\mathbf{R}'$ which can be used to upgrade demand $l+1$. However, upgrading demand $l+1$ by capacity h is not optimal under $\mathbf{R}'$, i.e., the expected value of such a unit of capacity h is higher than $\alpha_{h,l+1}$. Then,

$$\begin{aligned}\Theta^{t+1}(\mathbf{R}')-\Theta^{t+1}(\mathbf{R})&=\Theta^{t+1}(\mathbf{R}_{hj})-\Theta^{t+1}(\mathbf{R})\\ =&\Theta^{t+1}(\mathbf{R}_{hj})-\Theta^{t+1}(\mathbf{R}_{h,l+1})+\Theta^{t+1}(\mathbf{R}_{h,l+1})-\Theta^{t+1}(\mathbf{R})\\ \ge&\alpha_{h,l+1}+\Theta^{t+1}(\mathbf{R}_{hj})-\Theta^{t+1}(\mathbf{R}_{h,l+1}).\end{aligned} \tag{EC.48}$$

From the definition of $\tilde{\mathbf{R}}$ and the induction assumption, we have

$$\Theta^{t+1}(\mathbf{R}_{hj})-\Theta^{t+1}(\mathbf{R}_{h,l+1})\ge\Theta^{t+1}(\tilde{\mathbf{R}}_{hj})-\Theta^{t+1}(\tilde{\mathbf{R}}_{h,l+1})$$

because $(\mathbf{R})_{l+2,\cdots,j-1}=(\tilde{\mathbf{R}})_{l+2,\cdots,j-1}\le 0$. Moreover, $(\tilde{\mathbf{R}})_{h+1,\cdots,l}=(\mathbf{R})_{h+1,\cdots,l}=0$ by the assumption of this case, from Lemma EC.8, we similarly have

$$\Theta^{t+1}(\tilde{\mathbf{R}}_{hj})-\Theta^{t+1}(\tilde{\mathbf{R}}_{h,l+1})=\Theta^{t+1}(\bar{\mathbf{N}}_{l+1,j})-\Theta^{t+1}(\bar{\mathbf{N}}).$$

Thus,

$$\Theta^{t+1}(\mathbf{R}')-\Theta^{t+1}(\mathbf{R})\ge\alpha_{h,l+1}+\Theta^{t+1}(\bar{\mathbf{N}}_{l+1,j})-\Theta^{t+1}(\bar{\mathbf{N}}). \tag{EC.49}$$

Given $l<\bar{l}$, we have $\Theta^{t+1}(\bar{\mathbf{N}}_{l+1,j})-\Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j})\ge r_{\bar{l}}-r_{l+1}$ from Lemma EC.2. Since $\delta-\bar{\delta}=u_h-u_{\bar{h}}$ by (EC.42) and (EC.43), (EC.47) and (EC.49) imply that $\Delta\ge\bar{\Delta}$ as $g_{l+1}\ge g_{\bar{l}}$.

3. Case 3 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$: Since $l\le\bar{l}$, the same proof of “Case 1 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$” can be applied.
4. Case 1 for $\mathbf{N}$ and Case 2 for $\bar{\mathbf{N}}$: Note that (EC.41) and (EC.47) still hold, meanwhile, δ and $\bar{\delta}$ are given in (EC.37) and (EC.43), respectively. We have

$$\delta-(\bar{\delta}+\alpha_{\bar{h}\bar{l}})=r_l+g_l-r_{\bar{l}}-g_{\bar{l}}$$

and

$$\begin{aligned}&\left(\Theta^{t+1}(\mathbf{R}')-\Theta^{t+1}(\mathbf{R})\right)-\left(\Theta^{t+1}(\bar{\mathbf{R}}')-\Theta^{t+1}(\bar{\mathbf{R}})\right)\\ =&\left(\Theta^{t+1}(\mathbf{R}_{lj})-\Theta^{t+1}(\mathbf{R})\right)-\left(\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{h}j})-\Theta^{t+1}(\bar{\mathbf{R}})\right)\\ \ge&\left(\Theta^{t+1}(\bar{\mathbf{N}}_{lj})-\Theta^{t+1}(\bar{\mathbf{N}})\right)-\left(\alpha_{\bar{h}\bar{l}}+\Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j})-\Theta^{t+1}(\bar{\mathbf{N}})\right)\\ =&\Theta^{t+1}(\bar{\mathbf{N}}_{lj})-\Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j})-\alpha_{\bar{h}\bar{l}}.\end{aligned}$$

Given $l \le \bar{l}$, $\Theta^{t+1}(\bar{\mathbf{N}}_{lj}) - \Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) \ge r_{\bar{l}} - r_l$ by Lemma EC.2. Then, $\Delta \ge \bar{\Delta}$ since $g_l \ge g_{\bar{l}}$.

5. Case 1 for $\mathbf{N}$ and Case 3 for $\bar{\mathbf{N}}$: Note that $l \le \bar{l} < \bar{l}+1$, the same proof of "Case 1 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$" can be applied.
6. Case 2 for $\mathbf{N}$ and Case 1 for $\bar{\mathbf{N}}$: In this case, class $\bar{l}$ in $(\bar{\mathbf{N}}, \mathbf{D})$ still has unmet demand while demand l is fully satisfied in $(\mathbf{N}, \mathbf{D})$ by assumption. From the induction assumption, upgrade is more likely to happen under initial state $(\bar{\mathbf{N}}, \mathbf{D})$, thus $l < \bar{l}$.

   Given that (EC.38, EC.39, EC.42, EC.49) all hold, there is

$$(\delta + \alpha_{h,l+1}) - \bar{\delta} = r_{l+1} + g_{l+1} - r_{\bar{l}} - g_{\bar{l}}.$$

   Since $l+1 \le \bar{l}$,

$$\begin{aligned}
&\big(\Theta^{t+1}(\mathbf{R}') - \Theta^{t+1}(\mathbf{R})\big) - \big(\Theta^{t+1}(\bar{\mathbf{R}}') - \Theta^{t+1}(\bar{\mathbf{R}})\big) \\
=&\big(\Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R})\big) - \big(\Theta^{t+1}(\bar{\mathbf{R}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{R}})\big) \\
\ge&\big(\alpha_{h,l+1} + \Theta^{t+1}(\bar{\mathbf{N}}_{l+1,j}) - \Theta^{t+1}(\bar{\mathbf{N}})\big) - \big(\Theta^{t+1}(\bar{\mathbf{N}}_{\bar{l}j}) - \Theta^{t+1}(\bar{\mathbf{N}})\big) \\
\ge&\alpha_{h,l+1} + r_{\bar{l}} - r_{l+1}
\end{aligned}$$

   by Lemma EC.2. Then, $\Delta \ge \bar{\Delta}$ since $g_{l+1} \ge g_{\bar{l}}$.
7. Case 2 for $\mathbf{N}$ and Case 3 for $\bar{\mathbf{N}}$: Note that $l \le \bar{l} < \bar{l}+1$, the same proof of "Case 2 for $\mathbf{N}$ and Case 1 for $\bar{\mathbf{N}}$" can be applied.
8. Case 3 for $\mathbf{N}$ and Case 1 for $\bar{\mathbf{N}}$: To apply the same proof of "Case 1 for both $\mathbf{N}$ and $\bar{\mathbf{N}}$", we only need to show $l+1 \le \bar{l}$. Suppose to the contrary that $l+1 > \bar{l}$, then $l \ge \bar{l}$. Note that upgrade is more likely to happen under initial state $(\bar{\mathbf{N}}, \mathbf{D})$ by assumption. Recall the discussions about Case 1 and Case 3, there is unmet demand $\bar{l}$ remaining in $\bar{\mathbf{R}}'$, but all demand l has been satisfied in $\mathbf{R}$. This is a contradiction.
9. Case 3 for $\mathbf{N}$ and Case 2 for $\bar{\mathbf{N}}$: To apply the same proof of "Case 1 for $\mathbf{N}$ and Case 2 for $\bar{\mathbf{N}}$", we need to show $l+1 \le \bar{l}$. Similar to the above discussion, we suppose $l+1 > \bar{l}$. Note that all demand l has been satisfied in $\mathbf{R}$ and some of the lower class demand $l+1$ is also satisfied in $\mathbf{R}'$. Meanwhile, the demand lower than class $\bar{l}$ is not upgraded under both $\bar{\mathbf{R}}$ and $\bar{\mathbf{R}}'$. This is contradiction.

To complete this proof, we need to consider the case when $l < j$ and $\bar{l} \ge j$, where $\bar{l} \ge j$ means $\mathbf{R}' = \mathbf{R}$ and (EC.36) is true.

1. Case 1 for $\mathbf{N}$: From (EC.40),

$$\Theta^{t+1}(\mathbf{R}_{lj}) - \Theta^{t+1}(\mathbf{R}) \ge r_j - r_l$$

   by Lemma EC.2. Since $\bar{\Delta}$ is given in (EC.36), we have $\Delta \ge \bar{\Delta}$ from δ in (EC.37).

2. Case 2 for $\mathbf{N}$: From (EC.48) and the fact $l+1 \leq j$, there is

$$\Theta^{t+1}(\mathbf{R}_{hj}) - \Theta^{t+1}(\mathbf{R}_{h,l+1}) \geq r_j - r_{l+1}$$

   by Lemma EC.2. With $\bar{\Delta}$ in (EC.36) and δ in (EC.42), we have $\delta + \alpha_{h,l+1} + r_j + g_j - r_{l+1} - g_{l+1} = \bar{\Delta}$. Hence, $\Delta \geq \bar{\Delta}$.
3. Case 3 for $\mathbf{N}$: Note that $l+1 < j$ in this case. Then, the same proof of "Case 1 for $\mathbf{N}$" can be applied.

This completes the proof. □

The next lemma states that the protection level p_{ij} $(1 \leq i < j \leq N)$ in period $T-1$ decease in the states of classes $(1, \cdots, i-1)$.

Lemma EC.10. *Consider an N-class upgrading problem in period T with $(n_{i+1}, \cdots, n_j) \leq 0$. Let $\bar{\mathbf{N}} = \mathbf{N} + \epsilon \mathbf{e}_r$, where $1 \leq r < i$ and $\epsilon > 0$. Then,*

$$\Delta_{ij}^{+-}\Theta^T(\mathbf{N}) \geq \Delta_{ij}^{+-}\Theta^T(\bar{\mathbf{N}}), \qquad \Delta_{ij}^{-+}\Theta^T(\mathbf{N}) \geq \Delta_{ij}^{-+}\Theta^T(\bar{\mathbf{N}}).$$

Proof. Following the notations in the proof of Lemma EC.9, in this proof we only need to consider Case 1, Case 3 and Case 4 for Θ^T since the additional unit of capacity h $(\bar{h})$ will not be passed to the next period. Note that $\Delta = \delta$ and $\bar{\Delta} = \bar{\delta}$ since $\Theta^{T+1} \equiv 0$. Also, the protection levels are zero in period T.

Recall the similarity of Case 1 and Case 3. In the proof of Lemma EC.9, we have shown that $l+1 \leq \bar{l}$ if "Case 3 for $\mathbf{N}$ and Case 1 for $\bar{\mathbf{N}}$". Therefore, we only have three different cases in period T.

1. $j \leq l \leq \bar{l}$: Since (EC.36) still holds, we have $\Delta = \bar{\Delta}$.
2. $l \leq \bar{l} < j$: From (EC.37) and (EC.38), there is $\Delta - \bar{\Delta} = r_l + g_l - r_{\bar{l}} - g_{\bar{l}} \geq 0$ since $r_l \geq r_{\bar{l}}$ and $g_l \geq g_{\bar{l}}$.
3. $l < j \leq \bar{l}$: From (EC.36) and (EC.37), we have $\Delta - \bar{\Delta} = r_l + g_l - r_j - g_j > 0$ since $r_l > r_j$ and $g_l > g_j$.

Hence, the desired result holds in period T for any demand realization, which completes the proof. □

With the previous two lemmas, we can prove the monotonicity result.

Proposition 5 *The optimal protection level p_{ij} $(1 \leq i < j \leq N)$ in period t $(1 \leq t \leq T)$ are decreasing in $(n_1^t, \cdots, n_{i-1}^t)$.*

Proof. Given the definition of the protection level in (8), this proposition can be inductively proved using Lemmas EC.9 and EC.10. □

E.2. Non-concave Example

For the RCEC heuristic, although the allocation decisions in each period can be solved by a concave function $\bar{\Theta}^{t+1}_{\mathrm{RCEC}}$, $\Pi_{\mathrm{RCEC}}(\mathbf{X})$, the firm's total revenue with initial capacity $\mathbf{X}$ under the RCEC heuristic, is in general not concave or even quasi-concave. To illustrate this, we provide the following 2-product and 3-period example. In this example, the revenue is $(r_1, r_2) = (5.5, 2.5)$, the goodwill cost is $(g_1, g_2) = (1.5, 0)$, the usage cost is $(u_1, u_2) = (1.5, 0)$, and the capacity cost is $(c_1, c_2) = (0.2, 0.1)$. Hence, the profit margins are $(\alpha_{11}, \alpha_{22}, \alpha_{12}) = (5.5, 2.5, 1)$. We assume that there are only two possible demand realizations with equal probabilities $\frac{1}{2}$ in each period, i.e.,

$$\mathbf{D}^1 = \begin{cases} (1,4)^\intercal & \text{with probability } 0.5; \\ (1,0)^\intercal & \text{with probability } 0.5. \end{cases}$$

$$\mathbf{D}^2 = \begin{cases} (1,0)^\intercal & \text{with probability } 0.5; \\ (0,4)^\intercal & \text{with probability } 0.5. \end{cases}$$

$$\mathbf{D}^3 = \begin{cases} (1,4)^\intercal & \text{with probability } 0.5; \\ (0,0)^\intercal & \text{with probability } 0.5. \end{cases}$$

Let $x_2 = 1$, then Figure EC.1 illustrates the non-concavity of $\Pi_{\mathrm{RCEC}}(\mathbf{X})$ with respect to the initial capacity x_1.

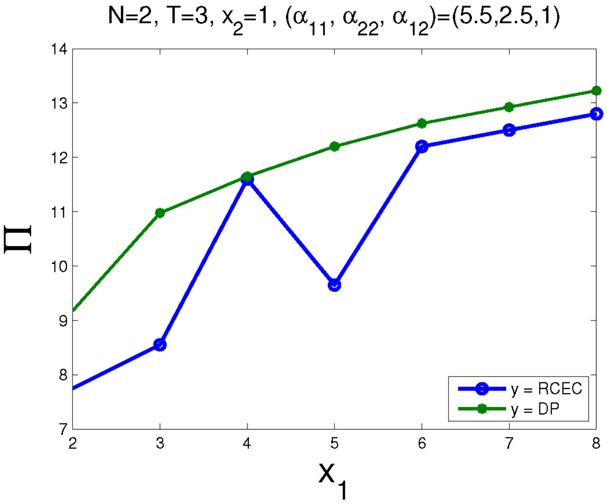

Figure EC.1 $\Pi_{\mathrm{RCEC}}(\mathbf{X})$