Dynamic Capacity Management with Substitution: Online Appendix

Robert A. Shumsky Fuqiang Zhang

January, 2008

1 Propositions, lemmas, and selected proofs

Proposition 1 $\Pi^{NV}(\mathbf{X}) \leq \Pi^{DYN}(\mathbf{X}) \leq \Pi^{STC}(\mathbf{X})$.

Proof. The first inequality follows from the fact that any allocation of capacity that is feasible in NV is also feasible in DYN, while DYN has the additional freedom to substitute products. The second follows from the fact that for a given demand realization $\mathbf{D}^1, \mathbf{D}^2, ..., \mathbf{D}^T$, any allocation decision available in DYN is also a feasible allocation in STC. In addition, there are allocation opportunities in STC that are not feasible in DYN because in STC, capacity is allocated to customers after the firm observes all demand. ■

Lemma 1 $\Theta^t(\mathbf{X})$ *is monotonically increasing in* $\mathbf{X}$

Proof. First consider $\Theta^T(\mathbf{X}|\mathbf{D}^T)$, the optimal value function at time T, given period-T demand. Because $\Theta^{T+1} \equiv 0$, there is an optimal solution with $\mathbf{X}^{T+1} = 0$, and $\Theta^T(\mathbf{X}|\mathbf{D}^T)$ may be reduced to $H^T(\mathbf{X}|\mathbf{D}^T)$. Consider the capacity constraint in DYN,

$$\sum_i y_{ij}^t \leq y_j^t \qquad j = 1, 2, ..., N.$$

In the dual problem of H^T, the variable associated with this constraint is nonnegative, and therefore the marginal value of each element of $\mathbf{X}$ in the primal problem is nonnegative. This is true for any realized demand $\mathbf{D}^T$ and therefore monotonicity is preserved under expectation and $\Theta^T(\mathbf{X})$ is monotonically increasing in $\mathbf{X}$.

Now assume that $\Theta^{t+1}(\mathbf{X})$ is monotonically increasing. Let $\mathbf{X}' > \mathbf{X}$ and let $\mathbf{Y}^*$ be the optimal capacity offered to H^t in $\Theta^t(\mathbf{X}|\mathbf{D}^t)$. Therefore,

$$\begin{aligned}
\Theta^t(\mathbf{X}'|\mathbf{D}^t) &= \operatorname*{Max}_{\substack{\mathbf{Y}^t+\mathbf{X}^{t+1}=\mathbf{X}' \\ \mathbf{Y}^t\in\mathbb{R}_N^+,\mathbf{X}^{t+1}\in\mathbb{R}_N^+}} [H^t(\mathbf{Y}^t|\mathbf{D}^t) + \Theta^{t+1}(\mathbf{X}^{t+1})] && (1)\\
&\geq H^t(\mathbf{Y}^*|\mathbf{D}^t) + \Theta^{t+1}(\mathbf{X}' - \mathbf{Y}^*) && (2)\\
&\geq H^t(\mathbf{Y}^*|\mathbf{D}^t) + \Theta^{t+1}(\mathbf{X} - \mathbf{Y}^*) && (3)\\
&= \Theta^t(\mathbf{X}|\mathbf{D}^t). && (4)
\end{aligned}$$

This inequality holds for any realization of $\mathbf{D}^t$ and therefore $\Theta^t(\mathbf{X})$ is monotonically increasing. ■

Lemma 2 $\Theta^t(\mathbf{X})$ *is concave in* $\mathbf{X}$

Proof. Because $\Theta^{T+1} \equiv 0$, $\Theta^T(\mathbf{X}^T|\mathbf{D}^T)$ is equivalent to $H^T(\mathbf{X}^T|\mathbf{D}^T)$, and $H^T(\mathbf{X}^T|\mathbf{D}^T)$ is concave in $\mathbf{X}^T$ because a linear program is jointly concave in variables that determine the right-hand-side of its constraints. Therefore $\Theta^T(\mathbf{X}^T)$ is concave because concavity is preserved over the expectation operator on $\mathbf{D}^t$ (see van Slyke and Wets, 1966, Proposition 7).

Now assume that $\Theta^{t+1}(\mathbf{X}^{t+1})$ is concave in $\mathbf{X}^{t+1}$. In time period t, the function $H^t(\mathbf{Y}^t|\mathbf{D}^t)$ is concave in $\mathbf{Y}^t$ because, again, $\mathbf{Y}^t$ determines the right-hand-sides of constraints in H^t. Therefore, $\Theta^t(\mathbf{X}^t|\mathbf{D}^t)$ is the maximum value of the sum of two concave functions, $H^t(\mathbf{Y}^t|\mathbf{D}^t) + \Theta^{t+1}(\mathbf{X}^{t+1})$, with the constraints $\mathbf{Y}^t + \mathbf{X}^{t+1} = \mathbf{X}^t, \mathbf{Y}^t \in \mathbb{R}_N^+$, and $\mathbf{X}^{t+1} \in \mathbb{R}_N^+$. By theorems 5.3 and 5.4 in Rockafeller (1970) this maximal value is concave in $\mathbf{X}^t$. Again, concavity is preserved when taking an expectation over $\mathbf{D}^t$, so that $\Theta^t(\mathbf{X}^t)$, is also concave in $\mathbf{X}^t$. ■

Lemma 3 *The following algorithm solves* $H^t(\mathbf{Y}|\mathbf{D})$*:*
(i) $y_{ii} = d_i \wedge y_i, i = 1...N$
(ii) $y_{i+1,i} = (d_{i+1} - y_{i+1})^+ \wedge (y_i - d_i)^+, i = 1...N-1.$

Proof. Given that capacity $\mathbf{Y}$ is available for sale in period t, and given demand realization $\mathbf{D}$, $H^t(\mathbf{Y}|\mathbf{D})$ is a transportation problem with a cost structure defined by assumptions (A1) - (A3). Bassok et al. (1999) point out that the cost structure of H^t corresponds to a Monge sequence so that steps (i) and (ii) solve the problem (Hoffman, 1963). ■

Lemma 4 *Suppose that at time* t *after completing Step 1 of PRA, net capacity* $n_i^t \leq 0, i = k+1, \cdots, k+j$*, so that the capacities of these products have been depleted. Then the optimization problem can be separated into two independent subproblems: an upper part consisting of products* 1 *to* $k+1$*, and a lower part consisting of products* $k+j+1$ *to* N*.*

Proof. Given that only single-step upgrading is profitable, products with indices $1, 2, ..., k$ will not be used to satisfy demand by classes $k + j + 1, ..., N$. Therefore, the assignment of products in one group does not affect the capacity or profits of the other group, and the global optimization problem is separable into the two subproblems. ■

Lemma 5 *Suppose that Θ^{t+1} has the following properties:*

1. *The PRA solves $\Theta^{t+1}(\mathbf{X})$*
2. $\delta_k^- \Theta^{t+1}(\mathbf{X}) \leq \alpha_{kk}$
3. $\Theta^{t+1}(\mathbf{X})$ *is concave in* $\mathbf{X}$

   Then properties (1)-(3) hold for Θ^t.

Proof. See the main paper, Shumsky and Zhang (2007). ■

Proposition 2 *The PRA is an optimal policy from among all admissible policies.*

Proof. See the main paper, Shumsky and Zhang (2007). ■

Proposition 3 *If $\mathbf{X}^1$ and demand vectors $\mathfrak{D}^1, ..., \mathfrak{D}^T$ are integer-valued, then there exists an optimal integer rationing policy $\left(\tilde{P}^1, ..., \tilde{P}^T\right)$.*

Proof. First we define *Concave and Linear Between Integers* (CLBI) functions. A function $f(x)$ is CLBI if it is concave and piecewise linear with changes in slope only at integer values of the domain (see Brumelle and McGill, 1993 for more details). A CLBI function $f(x)$ satisfies the following property.

Covering property: if c is a constant such that $\delta^+ f(s_2) < c < \delta^- f(s_1)$ for some $s_1 \leq s_2$, then there exists an integer $n \in [s_1, s_2]$ such that $c \in \delta f(n)$.

Now consider $t = T$. Suppose $\mathbf{X}^T$ and $\mathfrak{D}^T$ are integer-valued. Since T is the last period, all leftover products after parallel assignment should be used for upgrading (if there is such a need), so the optimal protection limits are zero. By Proposition 2 we know that

$$\Theta^T(\mathbf{X}^T) = E_{\mathfrak{D}^T} \left\{ \begin{array}{c} \sum_{i=1}^{N} \alpha_{ii}(d_i^T \wedge x_i^T) \\ + \sum_{i=1}^{N-1} \alpha_{i+1,i}[\left(d_{i+1}^T - x_{i+1}^T\right)^+ \wedge \left(x_i^T - d_i^T\right)^+] \end{array} \right\}. \quad (5)$$

The terms of (5) that include x_i are,

$$E\{\alpha_{ii}(d_i^T \wedge x_i^T) + \alpha_{i+1,i}[\left(d_{i+1}^T - x_{i+1}^T\right)^+ \wedge \left(x_i^T - d_i^T\right)^+] + \alpha_{i-1,i}[\left(d_i^T - x_i^T\right)^+ \wedge \left(x_{i-1}^T - d_{i-1}^T\right)^+]\}, \quad (6)$$

where the second term disappears when $i = N$ and the third term disappears when $i = 1$. All of the terms in (6) are CLBI in x_i since $\mathbf{X}_{-i}^T$ and $\mathbf{D}^T$ are integer-valued and the derivatives of these terms change value only when x_i is an integer. Thus we know that if $\mathbf{X}^T$ is integer-valued, then there exists an optimal integer rationing policy $\tilde{P}^T \equiv 0$ and $\Theta^T(\mathbf{X}^T)$ is CLBI in x_i.

Now consider any period t. Suppose that if $\mathbf{X}^{t+1}$ is integer-valued, then there exists an optimal integer rationing policy $\tilde{P}^{t+1}$ and $\Theta^{t+1}(\mathbf{X}^{t+1})$ is CLBI in x_i. Next we show that if $\mathbf{X}^t$ is integer-valued, then 1) there exists an optimal integer rationing policy $\tilde{P}^t$ and 2) $\Theta^t(\mathbf{X}^t)$ is CLBI in x_i.

Without loss of generality, consider the following upgrading subproblem for a given demand realization $\mathbf{D}^t$ in period t: there is $n_j > 0, n_{j+1} > 0, ..., n_k > 0, n_{k+1} < 0$ after parallel allocation. Note that $n_j, n_{j+1}, ..., n_{k+1}$ are all integers because $\mathbf{X}^t$ and $\mathbf{D}^t$ are both integer-valued. By Lemma 5, any $\widetilde{p}_k$ satisfying

$$\alpha_{k+1,k} \in \delta_k \Theta^{t+1}\left(n_j, n_{j+1}, ..., n_{k-1}, \widetilde{p}_k\right)$$

is an optimal protection level. Since $n_j, n_{j+1}, ..., n_{k-1}$ are integers and thus $\Theta^{t+1}\left(n_j, n_{j+1}, ..., n_{k-1}, \widetilde{p}_k\right)$ is CLBI in x_k by the induction assumption, the covering property implies that there exists an integer $\widetilde{p}_k$ that is optimal. So we have shown that there exists an optimal integer rationing policy $\tilde{P}^t$ in period t if $\mathbf{X}^t$ is integer-valued.

To show that $\Theta^t(\mathbf{X}^t)$ is CLBI in x_i,we can write

$$\Theta^t(\mathbf{X}^t|\mathbf{D}^t) = H(\mathbf{Y}^*|\mathbf{D}^t) + \Theta^{t+1}(\mathbf{X}^t - \mathbf{Y}^*),$$

where $\mathbf{Y}^*$ is the optimal capacity vector offered for sale in period t. Because $\mathbf{D}^t$ is integer-valued and there exists an optimal integer rationing policy $\tilde{P}^t$, there exists a $\mathbf{Y}^*$ that is integer-valued. If $\mathbf{X}^t$ is integer-valued, then $\mathbf{X}^t - \mathbf{Y}^*$ is also integer-valued. This implies that $\Theta^t(\mathbf{X}^t|\mathbf{D}^t)$ is CLBI in x_i since $\Theta^{t+1}(\mathbf{X}^t - \mathbf{Y}^*)$ is CLBI in x_i by the induction assumption. Therefore, we know that that $\Theta^t(\mathbf{X}^t) = E_{\mathfrak{D}^t}[\Theta^t(\mathbf{X}^t|\mathbf{D}^t)]$ is CLBI in x_i.

Therefore, for $t = 1...T$ the PRA is optimal and, if $\mathbf{X}^t$ and $\mathfrak{D}^t...\mathfrak{D}^T$ are integer-valued, there exists an optimal integer rationing policy $\tilde{P}^t$ for step 2 of the PRA. Now suppose that we begin with integer capacity $\mathbf{X}^1$ and demands are integer-valued for $t = 1...T$. Now we need only show that $\mathbf{X}^t$ is integer valued for $t = 2...T$. Integrality in capacity is preserved within each period because if the starting capacity in each period is integer, there exists optimal integer protection limits and the use of the PRA with integer protection limits passes integer capacities to the next period. Therefore, by forward induction from period 1 to T, $\mathbf{X}^t$ is integer-valued and there exists an optimal integer rationing policy $\left(\tilde{P}^1, ..., \tilde{P}^T\right)$. ■

Proposition 4 *Given capacity* $\mathbf{Z} \leq \mathbf{X}^*$ *at the beginning of a replenishment interval, an optimal replenishment policy is to order up to* $\mathbf{X}^*$ *and the PRA is an optimal rationing policy within the interval.*

Proof. Throughout this proof we will use the index $k \in (1..R)$ to identify replenishment intervals and $t \in (1..T)$ to identify demand periods within each interval. Bold-face symbols ($\mathbf{c}, \mathbf{X}$, etc.) represent column vectors and primes denote the transpose, so that the inner product of vectors $\mathbf{c}$ and $\mathbf{X}$ is $\mathbf{c}'\mathbf{X}$. As defined in the main paper, $\Pi(\mathbf{X}; \mathbf{l})$ represents $\Pi^{DYN}(\mathbf{X})$, solved with the vector of salvage values $\mathbf{l}$. Similarly, let $\Theta^1(\mathbf{X}; \mathbf{l})$ be the within-interval rationing problem, as defined in equation (2) in the main paper, given salvage values $\mathbf{l}$. Let $V_k(\mathbf{Z})$ be the expected present value at the beginning of interval k, before replenishment, given capacity $\mathbf{Z}$.

The proof is by induction. We first assume that $V_{k+1}(\mathbf{Z})$ has the following three properties:
(1) $V_{k+1}(\mathbf{Z})$ is concave in $\mathbf{Z}$.
(2) At the beginning of interval $k+1$ if capacity $\mathbf{Z} \leq \mathbf{X}^*$, an optimal policy is to order up to $\mathbf{X}^*$and the PRA is an optimal rationing policy within interval $k+1$.
(3) $V_{k+1}(\mathbf{Z})$ is affine in the starting state $\mathbf{Z}$, with slope $\mathbf{c}$.

We will show that if $\mathbf{Z} \leq \mathbf{X}^*$at the beginning of interval k, an optimal policy is to order up to $\mathbf{X}^*$and the PRA is an optimal rationing policy within interval k. We will also show that properties (1) to (3) are preserved in interval k under optimization and that all three properties hold for the last interval R. First, the Bellman equation for interval k is,

$$V_k(\mathbf{Z}) = \underset{\mathbf{X} \geq \mathbf{Z}}{\text{Max}} \left[\Theta^1(\mathbf{X}; -\mathbf{h}) - \mathbf{c}'(\mathbf{X} - \mathbf{Z}) + \gamma V_{k+1}(\mathbf{X}^{T+1})\right] \tag{7}$$

$$= \underset{\mathbf{X} \geq \mathbf{Z}}{\text{Max}} \left[\Pi(\mathbf{X}; -\mathbf{h}) + \mathbf{c}'\mathbf{Z} + \gamma V_{k+1}(\mathbf{X}^{T+1})\right] \tag{8}$$

where $\mathbf{X}^{T+1}$ is the capacity left-over after demand period T in interval k. Note that this is a slight abuse of notation, for $\mathbf{X}^{T+1}$ is a function of $\mathbf{X}$ as well as the solution to the rationing problem in Π.

Let

$$G_k(\mathbf{X}) = \Pi(\mathbf{X}; -\mathbf{h}) + \mathbf{c}'\mathbf{Z} + \gamma V_{k+1}(\mathbf{X}^{T+1}) \tag{9}$$

To show that property (1) in conserved in interval k, we repeatedly apply the property of concavity preservation under maximization to show that $G_k(\mathbf{X})$ is concave in $\mathbf{X}$. Specifically, suppose that we have reached the beginning of demand period T within interval k, and capacity $\mathbf{Y} \leq \mathbf{X}$ has been allocated to fulfill demand thus far in the interval. Therefore, the present value is,

$$\underset{\mathfrak{D}^T}{\text{E}} \left\{ \underset{\mathbf{Y}^T + \mathbf{X}^{T+1} = \mathbf{X} - \mathbf{Y}}{\text{Max}} [H^T(\mathbf{Y}^T | \mathbf{D}^T) + \gamma V_{k+1}(\mathbf{X}^{T+1})] \right\}. \tag{10}$$

Because both H^T and V_{k+1} are concave, by theorems 5.3 and 5.4 in Rockafeller (1970), the maximal value inside the expectation is concave in $\mathbf{X}$. Concavity is preserved when taking the expectation over $\mathfrak{D}^T$, so that present value (10) is concave in $\mathbf{X}$. Working backwards, an identical argument applies to the sum of H^{T-1} and (10), and the argument can then be applied to $t = T-2, T-3, .., 1$. Therefore, $G_k(\mathbf{X})$ is concave in $\mathbf{X}$, and another application of concavity preservation under maximization shows that $V_k(\mathbf{Z})$ is concave.

For property (2),

$$\begin{aligned} G_k(\mathbf{X}) &= \Pi(\mathbf{X}; -\mathbf{h}) + \mathbf{c}'\mathbf{Z} + \gamma V_{k+1}(\mathbf{X}^{T+1}) && (11) \\ &= \Pi(\mathbf{X}; -\mathbf{h}) + \mathbf{c}'\mathbf{Z} + \gamma V_{k+1}(\mathbf{X}^*) - \gamma \mathbf{c}'(\mathbf{X}^* - \mathbf{X}^{T+1}) && (12) \\ &= \Pi(\mathbf{X}; -\mathbf{h}) + \boldsymbol{\gamma}\mathbf{c}'\mathbf{X}^{T+1} + \mathbf{c}'\mathbf{Z} + \gamma V_{k+1}(\mathbf{X}^*) - \gamma \mathbf{c}'\mathbf{X}^* && (13) \\ &= \Pi(\mathbf{X}; \gamma\mathbf{c} - \mathbf{h}) + \mathbf{c}'\mathbf{Z} + \gamma V_{k+1}(\mathbf{X}^*) - \gamma \mathbf{c}'\mathbf{X}^* && (14) \end{aligned}$$

where (12) follows from property 3 and (14) follows by incorporating the additional salvage-value $\gamma\mathbf{c}'\mathbf{X}^{T+1}$ into problem Π. By definition $\mathbf{X}^*$ is a maximizer of $\Pi(\mathbf{X}; \gamma\mathbf{c} - \mathbf{h})$, and we have shown that the PRA maximizes Π, given any initial capacity $\mathbf{X}$. Therefore, if $\mathbf{Z} \leq \mathbf{X}^*$ an optimal policy is to order up to $\mathbf{X}^*$ and to use the PRA within interval k.

For property (3), note that under the optimal policy,

$$V_k(\mathbf{Z}) = \Pi(\mathbf{X}^*; \gamma\mathbf{c} - \mathbf{h}) + \mathbf{c}'\mathbf{Z} + \gamma V_{k+1}(\mathbf{X}^*) - \gamma\mathbf{c}'\mathbf{X}^*,$$

which is affine in $\mathbf{Z}$ with slope $\mathbf{c}$.

For interval R, $V_{R+1}(\mathbf{Z}) \equiv \mathbf{c}'\mathbf{Z}$. Therefore, repeated applications of the property of concavity preservation under maximization, as described above, show that $V_R(\mathbf{Z})$ is also concave. To show that property (2) holds for interval R,

$$\begin{aligned} G_R(\mathbf{X}) &= \Pi(\mathbf{X}; -\mathbf{h}) + \mathbf{c}'\mathbf{Z} + \gamma\mathbf{c}'\mathbf{X}^{T+1} && (15) \\ &= \Pi(\mathbf{X}; \gamma\mathbf{c} - \mathbf{h}) + \mathbf{c}'\mathbf{Z}. && (16) \end{aligned}$$

Therfore, if $\mathbf{Z} \leq \mathbf{X}^*$an optimal policy is to order up to $\mathbf{X}^*$and to use the PRA within interval R. Finally,

$$V_R(\mathbf{Z}) = \Pi(\mathbf{X}^*; \gamma\mathbf{c} - \mathbf{h}) + \mathbf{c}'\mathbf{Z}. \quad (17)$$

Therefore, property (3) holds for interval R. ■

Proposition 5 *The optimal protection limit $\widetilde{p}^t$ is decreasing in the state vector $\mathbf{X}^t$.*

Proof. Consider two subproblems in time period t, and without loss of generality assume that the subproblem's product indices are $1, ..., k+1$. Before step 1, the first subproblem has capacities $\mathbf{X}^t$, where $x_i^t > 0, i = 1, ..., k$, and $x_{k+1}^t = 0$. The second subproblem has capacities $\hat{\mathbf{X}}^t = \mathbf{X}^t + \mathbf{e}_j$ with $1 \leq j \leq k-1$. Let $\Delta_k^t(\mathbf{X}^t)$ be the marginal value of an additional unit of product k in time-period t, given capacity $\mathbf{X}^t$. To prove that the proposition is true, we proceed by backwards induction, with two induction assumptions: (i) the optimal protection limit $\widetilde{p}^t$ is decreasing in the capacity vector $\mathbf{X}^t$ (this is the Proposition) and (ii) in the *next* time-period, the marginal value of product k is decreasing in the capacity vector. That is, $\Delta_k^{t+1}(\hat{\mathbf{X}}^{t+1}) \leq \Delta_k^{t+1}(\mathbf{X}^{t+1})$ for $\hat{\mathbf{X}}^{t+1} = \mathbf{X}^{t+1} + \mathbf{e}_j, 1 \leq j \leq k-1$.

Before showing that the induction assumptions are true for all t, we first prove that assumption (ii) implies (i). Recall that the protection limit $\widetilde{p}^t$ solves a concave optimization problem in one variable, with the solution specified by the condition,

$$\alpha_{k+1,k} \geq \Delta_k^{t+1}(n_1^t, ..., n_{k-1}^t, p,). \tag{18}$$

The right-hand-side of (18) is the marginal value of an increase in the quantity of product k made available in the next period. Therefore, the protection limit rises or falls as the marginal value of product k in the next period rises or falls. Furthermore, if $\hat{\mathbf{X}}^t = \mathbf{X}^t + \mathbf{e}_j$ for some $1 \leq j \leq k-1$, then $\hat{x}_j^{t+1} \geq x_j^{t+1}$, because the extra capacity of the higher-level product is either passed along or used to satisfy demand in period t. Therefore, given induction assumption (ii), an increase in $\mathbf{X}^t$ may lead to a decrease in the marginal value of product k in the next period, and $\widetilde{p}^t$ is decreasing in $\mathbf{X}^t$.

Now consider the rationing problem at time T. We first prove that for the optimal objective function, the marginal value of one extra unit of a product, $\Delta_k^T(\mathbf{X})$, is decreasing in the quantity of any other product (i.e., the objective function is submodular). First, the optimal allocation is to (i) make all possible parallel assignments and then (ii) make all possible one-step upgrades. For $k = 2...N-1$,an additional unit of product k costs c_k and may be used for a parallel assignment, may be used for an upgrade to a $k+1$ customer, and may prevent an upgrade from k to $k-1$. Therefore,

$$\begin{aligned}&\Delta_k^T(\mathbf{X}) = \alpha_{kk} \Pr(\mathfrak{d}_k > x_k)\\ &+ \alpha_{k+1,k} \Pr(\mathfrak{d}_k \leq x_k, \mathfrak{d}_k + \mathfrak{d}_{k+1} > x_k + x_{k+1}) - \alpha_{k,k-1} \Pr(\mathfrak{d}_k > x_k, \mathfrak{d}_{k-1} + \mathfrak{d}_k \leq x_{k-1} + x_k) - c_k\,.\end{aligned}$$

Therefore, for $\hat{\mathbf{X}} = \mathbf{X} + \mathbf{e}_j$, $\Delta_k^T(\hat{\mathbf{X}}) - \Delta_k^T(\mathbf{X}) = 0$ for $j < k-1$ and $j > k+1$. For $j = k+1$,

$$
\begin{aligned}
& \Delta_k^T(\hat{\mathbf{X}}) - \Delta_k^T(\mathbf{X}) && (19)\\
= \;& \alpha_{k+1,k}\left[\Pr(\mathfrak{d}_k \le x_k, \mathfrak{d}_k + \mathfrak{d}_{k+1} > x_k + x_{k+1} + 1) - \Pr(\mathfrak{d}_k \le x_k, \mathfrak{d}_k + \mathfrak{d}_{k+1} > x_k + x_{k+1})\right] && (20)\\
\le \;& 0 && (21)
\end{aligned}
$$

For $j = k-1$,

$$
\begin{aligned}
& \Delta_k^T(\hat{\mathbf{X}}) - \Delta_k^T(\mathbf{X}) && (22)\\
= \;& -\alpha_{k,k-1}\left[\Pr(\mathfrak{d}_k > x_k, \mathfrak{d}_{k-1} + \mathfrak{d}_k \le x_{k-1} + x_k + 1) - \Pr(\mathfrak{d}_k > x_k, \mathfrak{d}_{k-1} + \mathfrak{d}_k \le x_{k-1} + x_k)\right] && (23)\\
\le \;& 0 && (24)
\end{aligned}
$$

For $k = 1, \Delta_1^T(\mathbf{X} + \mathbf{e}_2) - \Delta_1^T(\mathbf{X})$ is also equal to (20) and for $k = N, \Delta_N^T(\mathbf{X} + \mathbf{e}_{N-1}) - \Delta_N^T(\mathbf{X})$ is also equal to (23). Therefore, $\Delta_k^T(\hat{\mathbf{X}}) \le \Delta_k^T(\mathbf{X})\ \forall j \ne k$. From the discussion in the last paragraph, this also implies that the optimal protection limit $\widetilde{p}^{T-1}$ is decreasing in the capacity vector $\mathbf{X}^{T-1}$.

Assume that induction assumptions (i) and (ii) hold for periods t and $t+1$, respectively, and we will show that (ii) is true for t and therefore (i) is true for $t-1$. Given a realization of demand in period t, $\mathbf{D}^t$, after Step 1 we are left with the net capacity vectors $\mathbf{N}^t = \mathbf{X}^t - \mathbf{D}^t$ and $\hat{\mathbf{N}}^t = \hat{\mathbf{X}}^t - \mathbf{D}^t$ (note $\mathbf{N}^t$ and $\hat{\mathbf{N}}^t$ only differ in the jth element, and by one unit). To find the marginal value of an extra unit of product k, we must consider a variety of scenarios. In each of these cases, an extra unit of product k may be used for one of three things. The unit may be used for a parallel assignment to a customer of class k (denoted by '$\underrightarrow{k}$' and '$\widehat{\underrightarrow{k}}$' given $\mathbf{N}^t$ and $\hat{\mathbf{N}}^t$, respectively), it may be used to upgrade a customer of class $k+1$ (denoted '$\underrightarrow{k+1}$' and '$\widehat{\underrightarrow{k+1}}$') and it may not be used in period t but passed along to period $t+1$ (denoted '$\underrightarrow{t+1}$' and '$\widehat{\underrightarrow{t+1}}$'). Before cataloguing an exhaustive list of scenarios, we consider the following observation:

Observation: Suppose that in period t, $n_k^t > 0$, and that the extra unit of product k is not allocated in period t but is passed along to the next period ('$\underrightarrow{t+1}$'). Then one of the following must be true:

Case A: We have the event $\underrightarrow{t+1}$ because all excess type-$(k+1)$ demand has been upgraded and the protection limit has not yet been reached. In this case $\Delta_k^t(\mathbf{X}^t) \le \alpha_{k+1,k}$ because the quantity of available capacity is larger than the protection limit.

Case B: We have the event $\underrightarrow{t+1}$ even though there is still excess type-$(k+1)$ demand to be upgraded. In this case, the protection limit *has* been reached. Here we can also make a somewhat surprising conclusion: there were *no* upgrades in period t. This can be shown by contradiction.

Suppose that there were upgrades in period t. Then there was one type-$(k+1)$ customer who hit the protection limit during the period and was not upgraded. But if we add an extra unit of type-k product, then this unit will be used to upgrade that customer, and we have $\underrightarrow{k+1}$, instead of the assumed event, $\underrightarrow{t+1}$. Also, in this case, $\Delta_k^t(\mathbf{X}^t) \geq \alpha_{k+1,k}$ because the protection limit has been reached.

The same reasoning can be applied when we have residual capacity $\hat{\mathbf{N}}^t$ and event $\widehat{\underrightarrow{t+1}}$: only **Case A** and **Case B** are possible.

Now we are ready to list all possible sample paths and examine, for each path, the marginal value of an extra unit of product k given capacities $\mathbf{X}^t$and $\hat{\mathbf{X}}^t$. We begin by looking at a relatively simple case in which our subproblem 'splits' because we run out of capacity for a high-level product:

(1) $\hat{n}_i^t \leq 0$ for some $j \leq i \leq k-1$, so that the demand for some product in the chain between j and $k-1$ is greater than the corresponding capacity $\hat{\mathbf{X}}^t$ (thus also $\mathbf{X}^t$). Then, the allocation problem separates in period $t+1$ and the one extra unit of product j in $\hat{\mathbf{X}}^t$ has no impact on the marginal value of product k. Therefore, $\Delta_k^t(\hat{\mathbf{X}}^t) = \Delta_k^t(\mathbf{X}^t)$.

(2) For the remaining scenarios we assume that $\hat{n}_i^t > 0$ for all $j \leq i \leq k-1$. We define subcases according to the value of $\hat{n}_k^t = n_k^t$, the amount of product k available after Step 1. We consider (2.1) $\hat{n}_k^t \geq 0$ and (2.2) $\hat{n}_k^t < 0$. Unfortunately, each of these cases will also have subcases, and subsubcases!

(2.1) $\hat{n}_k^t = n_k^t \geq 0$. Here there are two subcases, $n_{k+1}^t = 0$ and $n_{k+1}^t < 0$ (we cannot have $n_{k+1}^t > 0$, according to the definition of the subproblem).

(2.1.1) If $\hat{n}_{k+1}^t = n_{k+1}^t = 0$ then there will be no upgrading and $\hat{n}_k^t = n_k^t$ will be passed to period $t+1$. Therefore, by the induction assumption, we know $\Delta_k^t(\hat{\mathbf{X}}^t) \leq \Delta_k^t(\mathbf{X}^t)$.

(2.1.2) If $\hat{n}_{k+1}^t = n_{k+1}^t < 0$, then the extra unit of product k *may* be used to upgrade demand for product $k+1$. This is the most complex case because the extra unit may be used differently, given $\mathbf{X}^t$ and $\hat{\mathbf{X}}^t$ (recall that the protection limit may be lower under $\hat{\mathbf{X}}^t$). Because $\hat{n}_k^t = n_k^t \geq 0$ there is no type-k demand remaining, so we cannot have $\underrightarrow{k}$ or $\widehat{\underrightarrow{k}}$. Therefore, we have four cases: $(\underrightarrow{k+1}, \widehat{\underrightarrow{k+1}})$, $(\underrightarrow{t+1}, \widehat{\underrightarrow{t+1}})$, $(\underrightarrow{k+1}, \widehat{\underrightarrow{t+1}})$, and $(\underrightarrow{t+1}, \widehat{\underrightarrow{k+1}})$.

(2.1.2.1) $(\underrightarrow{k+1}, \widehat{\underrightarrow{k+1}})$: In this case, $\Delta_k^t(\hat{\mathbf{X}}^t) = \Delta_k^t(\mathbf{X}^t) = \alpha_{k+1,k}$.

(2.1.2.2) $(\underrightarrow{t+1}, \widehat{\underrightarrow{t+1}})$: From the Observation above, the same amount of product k is passed to period $t+1$ under $\mathbf{X}^t$ and $\hat{\mathbf{X}}^t$. For **Case A**, all demand for product $k+1$ is upgraded, and the same quantity $n_k^t - d_{k+1}^t$ is passed to period $t+1$ under both $\mathbf{X}^t$ and $\hat{\mathbf{X}}^t$. For **Case B**, there is no upgrading, so n_k^t is passed to period $t+1$ under both $\mathbf{X}^t$ and $\hat{\mathbf{X}}^t$. Then by the induction assumption, we know $\Delta_k^t(\hat{\mathbf{X}}^t) \leq \Delta_k^t(\mathbf{X}^t)$.

(2.1.2.3) $(\underrightarrow{k+1}, \widehat{\underrightarrow{t+1}})$: Here the additional unit in $\mathbf{X}^t$ is used for upgrading, for a marginal value of $\alpha_{k+1,k}$. Under $\hat{\mathbf{X}}^t$, we are passing along the extra unit, and for **Case A** we know that $\Delta_k^t(\hat{\mathbf{X}}^t) \leq \alpha_{k+1,k} = \Delta_k^t(\mathbf{X}^t)$. **Case B** implies an upgrade occurred under $\mathbf{X}^t$ while the same unit of capacity was protected under $\hat{\mathbf{X}}^t$, implying a larger protection limit under $\hat{\mathbf{X}}^t$. But the induction assumption indicates that protection limits are decreasing under $\hat{\mathbf{X}}^t$. Therefore, **Case B** cannot occur.

(2.1.2.4) $(\underrightarrow{t+1}, \widehat{\underrightarrow{k+1}})$: Under $\mathbf{X}^t$ we again consider **Case A** and **Case B**. For **Case A**, we observed that all demand must have been upgraded and that there is more capacity than the protection limit. However, we also know that under $\hat{\mathbf{X}}^t$ the protection limit is the same, or smaller, than under $\mathbf{X}^t$ so both $\underrightarrow{t+1}$ and $\widehat{\underrightarrow{k+1}}$ cannot occur simultaneously, and **Case A** is impossible. Given **Case B**, under $\hat{\mathbf{X}}^t$ the extra unit of product k is used for upgrading, with marginal value $\alpha_{k+1,k}$. Under $\mathbf{X}^t$ we know the marginal value of the additional unit is at least as high as $\alpha_{k+1,k}$ because the unit is passed to the next period even though there is an upgrading opportunity. Again, we have $\Delta_k^t(\hat{\mathbf{X}}^t) \leq \Delta_k^t(\mathbf{X}^t)$.

(2.2) $\hat{n}_k^t = n_k^t < 0$. Because it is always optimal to complete parallel allocations (Step 1), this case implies events $\underrightarrow{k}$ and $\widehat{\underrightarrow{k}}$: we always assign an extra unit of product k to unmet k demand. However, to calculate the marginal value of this assignment, we have to consider whether this 'marginal' customer had already been satisfied by an upgrade to capacity $k-1$. Therefore, we consider four cases:

(2.2.1) For both $\mathbf{X}^t$ and $\hat{\mathbf{X}}^t$, the additional unit of product k satisfies a type-k customer who otherwise would have been turned away. In this case, $\Delta_k^t(\hat{\mathbf{X}}^t) = \Delta_k^t(\mathbf{X}^t)$.

(2.2.2) For both $\mathbf{X}^t$ and $\hat{\mathbf{X}}^t$, the additional unit of product k satisfies a type-k customer who otherwise would have been upgraded to product $k-1$. In this case, $\Delta_k^t(\hat{\mathbf{X}}^t) = \Delta_k^t(\mathbf{X}^t)$.

(2.2.3) Under $\mathbf{X}^t$ the customer would not have been upgraded (would have been turned away), but under $\hat{\mathbf{X}}^t$ the additional unit of product k satisfies a type-k customer who otherwise would have been upgraded to product $k-1$. In this case, $\Delta_k^t(\hat{\mathbf{X}}^t) = \alpha_{kk} - \alpha_{k,k-1} + \Delta_{k-1}^{t+1}(\hat{\mathbf{X}}^t)$. Because the last unit of product $k-1$ had been used for upgrading, we know $\Delta_{k-1}^{t+1}(\hat{\mathbf{X}}^t) \leq \alpha_{k,k-1}$. Therefore, $\Delta_k^t(\hat{\mathbf{X}}^t) \leq \Delta_k^t(\mathbf{X}^t) = \alpha_{kk}$.

(2.2.4) Under the last scenario, the marginal customer would have been upgraded under $\mathbf{X}^t$ but not upgraded under $\hat{\mathbf{X}}^t$. However, our induction assumption states that under $\hat{\mathbf{X}}^t$ the protection limit is the same, or smaller, than under $\mathbf{X}^t$. Therefore, this scenario cannot occur.

We have shown that for all possible scenarios $\Delta_k^t(\hat{\mathbf{X}}^t) \leq \Delta_k^t(\mathbf{X}^t)$ and, therefore, the protection limit is decreasing in the state vector. ■

Proposition 6 *The optimal protection limit* $\widetilde{p}^t$ *is decreasing in* t.

Proof. Consider two rationing problems with the same state vector $\mathbf{N} = (n_1, n_2, ..., n_{k+1})$. Let problem 1 be in period t_1, while problem 2 is in period t_2, and $t_1 < t_2$. Let $\widetilde{p}^1$ and $\widetilde{p}^2$ be the optimal protection limits for product k in the two problems, respectively. To prove $\widetilde{p}^1 \geq \widetilde{p}^2$ we first show that the marginal value of product k, passed to the next period, is higher in problem 1 than that in problem 2. In particular, we show that this is true for any sample path between t_1 and t_2.

Suppose that an extra unit of product k is passed to $t_1 + 1$ in problem 1 and to $t_2 + 1$ in problem 2, and consider the demand arriving in problem 1 during periods $t_1 + 1$ to t_2. There are two possible cases. First, if no demand for any product is satisfied during those periods. In this case, problem 1 is equivalent to problem 2 at period $t_2 + 1$. Second, if a positive amount of demand is satisfied during those periods. Then at period $t_2 + 1$, the capacity vector of problem 1 is strictly smaller than that of problem 2. By the reasoning in the proof of Proposition 5, the marginal value of a unit of product k passed to the next period is higher for problem 1 than for problem 2. From the rationing optimality condition (18), $\widetilde{p}^1 \geq \widetilde{p}^2$. ■

Proposition 7 *For a subproblem with* k *products,*
$\widetilde{p}(\mathbf{X}(0,\infty)) \leq \widetilde{p}(\mathbf{X}(1,\infty)) \leq ... \leq \widetilde{p}(\mathbf{X}(k-1,\infty))$
$\leq \widetilde{p}(\mathbf{X})$
$\leq \widetilde{p}(\mathbf{X}(k-1,0)) \leq \widetilde{p}(\mathbf{X}(k-2,0)) \leq ... \leq \widetilde{p}(\mathbf{X}(0,0))$.

Proof. See the main paper, Shumsky and Zhang (2007). ■

2 Protection limit bounds: numerical experiments

This Section describes details of the experiments to test the quality of the bounds on the protection limits. In all of these experiments we have 5 products ($k = 5$), 10 time periods ($T = 10$), and a maximum initial capacity of 30 ($\widehat{x} = 30$) for each product. There are two major subsets of experiments, one using Poisson distributions that are independent between demand periods and between products, and another using the multivariate normal distribution (truncated at 0 and rounded to the nearest integer), with within-period correlation among demands. We summarize the parameter sets for these experiments in Table 1.

	Poisson demand (288 experiments)	Multivariate normal demand (120 experiments)
demand distibutions	12 scenarios (see Appendix)	Mean demand=2 units for every product in every period, coefficient of variation=1, correlation coefficients=(-0.25,0,0.25,0.5,0.9)
initial capacity	4 realistic scenarios, 4 extreme scenarios (see Appendix)	
contribution margins	3 realistic scenarios, 3 extreme scenarios (see Appendix)	

Table 1: Summary of the parameters for the numerical tests of the tightness of the bounds

For the Poisson experiments we define 12 demand scenarios, including 'flat' (demands the same in each period for all products), 'low then high' (demands for low-value products decrease over time while demands for high-value products increase), and 'alternate' (product 1 has demand 0, 10, 0, 10... while product 2 has demand 8, 0, 8, 0..., etc.). These scenarios also feature varying quantities of total demand (over all 10 periods) for each product. The mean demands for all 12 scenarios are included in the Appendix, below.

For the multivariate normal experiments, all products in every period have an average demand of 2 units and a coefficient of variation equal to 1 (the standard deviation for each product is 2 in the underlying normal distribution, before truncation and rounding). We then vary the coefficient of correlations among all demands from -0.25 to 0.9 (specifically, using -0.25, 0, 0.25, 0.5 and 0.9). That is, in one set of experiments the correlation between any two of the five products is -0.25, in the next set the correlation is 0, etc.

For both the Poisson and normal experiments, we defined two sets of parameters for the contribution margins $\alpha_{i,j}$ and initial capacities $\mathbf{X}^1$, roughly categorized as *realistic* and *extreme* parameters. Tables containing the complete parameter sets are included in the Appendix, below. For the realistic scenarios we define 3 sets of contribution margins $\alpha_{i,j}$ and 4 sets of initial capacities. For these realistic scenarios all upgrade margins are approximately 15-50% of the parallel margins. The initial capacities for all products are close together, within 10 units of each adjacent product. For each product, the initial capacities are usually close to the total demand over all 10 periods. The extreme scenarios also include 3 sets of margins $\alpha_{i,j}$ and 4 sets of initial capacities. For these extreme scenarios, the upgrade margins can be nearly equal to the parallel margins, or nearly 0, and the initial capacities for certain products can be double the total demand, or nearly 0. In total, there are 12 Poisson demand patterns, leading to $12 \times 3 \times 4 = 144$ 'realistic' and 144 'extreme' parameter combinations, producing 288 Poisson scenarios in total. For the normal experiments there are 5 correlation coefficients, leading to $5 \times 3 \times 4 = 60$ 'realistic' and 60 'extreme' parameter combinations, producing 120 multivariate normal scenarios in total.

3 Protection limits and optimal capacity for the static and dynamic 2X2 model

The single-period (static) model has been a popular framework for exploring the impact of flexibility on the optimal level of capacity investment. Using a single-period model, Bassok et al. (1999) and Netessine et al. (2002) show that the optimal level of flexible, class-1 capacity is higher than the optimal level if that product were not available for upgrades (i.e., higher than the newsvendor quantity). Likewise, they show that the optimal level of the lowest-class capacity is lower than the equivalent newsvendor quantity, because customers for the lowest-class product can be upgraded. This section compares optimal capacities for the static model, STC, the dynamic model, DYN, and the newsvendor quantities. In this Section, we assume that each period's demand and capacities are non-negative real numbers: $\mathbf{D}^t \in \mathbb{R}_2^+$ and $\mathbf{X}^t \in \mathbb{R}_2^+$.

3.1 Protection limits in the 2x2 model

Because it is prohibitively unwieldy to derive and analyze expressions for rationing policies and optimal capacities of the N-product, T-period model, here we examine the simplest possible model that retains both the product flexibility and the dynamic nature of the general model: a model with two products and two time-periods (the '2x2 model'). In this section, analysis of this model leads to an understanding of how the protection limit changes with product contribution margins, the distribution of product demand, and the correlation between demand distributions.

First we derive first-order conditions for p^*, the optimal protection limit for product 1 in the first period (no protection limit is needed in the last period). After the parallel assignment prescribed by Step 1, suppose that there is excess type-1 capacity, $n_1^1 > 0$, and surplus demand from type-2 customers, $n_2^1 < 0$ (otherwise, no rationing decision is necessary). If we upgrade type-2 customers until we reach the protection limit p, the 2nd-period profit $\Gamma(p)$ is

$$\Gamma(p) = \underset{\mathbf{D}^2}{E}\left[\alpha_{11}\min(d_1^2, p) + \alpha_{21}\min[d_2^2, (p - d_1^2)^+]\right]. \tag{25}$$

From the discussion in the main body of the paper, the optimal protection p^* limit must satisfy the property $\alpha_{21} \in \delta\Gamma(p^*)$. Under the assumption that capacity is continuous and that $\Gamma(p)$ is differentiable, this is equivalent to the first-order condition that $\Gamma'(p^*) - \alpha_{21} = 0$. Taking the derivative of (25) with respect to p, we find,

$$\alpha_{11}P(d_1^2 > p^*) + \alpha_{21}P(d_1^2 \leq p^*, d_1^2 + d_2^2 > p^*) - \alpha_{21} = 0. \tag{26}$$

Using the identity $P(d_1^2 \leq p^*, d_1^2 + d_2^2 > p^*) = P(d_1^2 \leq p^*) - P(d_1^2 + d_2^2 \leq p^*)$, the first-order condition can be rewritten as

$$\frac{P(d_1^2 + d_2^2 \leq p^*)}{P(d_1^2 > p^*)} = \frac{\alpha_{11} - \alpha_{21}}{\alpha_{21}}. \tag{27}$$

One might think of the ratio $\beta \equiv \alpha_{11}/\alpha_{21}$ as a measure of the cost of supply cannibalization. Because the left-hand side of (27) is increasing in p^*, and because the right-hand side of (27) is equal to $\beta - 1$, p^* increases with β. This makes sense: as the cost of supply cannibalization increases, the protection limit should increase, as well.

Recall that $a_{11} = p_1 + v_1 - u_1$ and $a_{21} = p_2 + v_2 - u_1$. The following proposition also follows directly from (27).

Proposition 8 *Given the $2X2$ case, and if all other parameters are held constant,*

(1) The optimal protection limit p^* rises with price $p_1 + v_1$ and falls with price $p_2 + v_2$;

(2) The optimal protection limit p^* rises with variable cost u_1, while variable cost u_2 has no impact on p^*.

Point (2) in the proposition indicates that as the usage cost u_1 rises, the firm is less willing to release expensive capacity to less-lucrative customers.

From Equation (27) we know that the optimal protection limit depends on the demand distributions in the second period (demand distributions in the first period have no impact on p^*). How will p^* change if there is a change in the demand distributions? In particular, what happens if demand shifts higher, or lower, in the next period? Here, a 'demand shift' is indicated by the usual stochastic order, written as $F^1 \leq_{st} F^2$. We say that distribution F^1 is stochastically dominated by distribution F^2 if $F^1(x) \geq F^2(x)$ for all x. In the following proposition and proof we suppress the time superscript, given that all distributions refer to the period-2 demands.

Proposition 9 *In the $2X2$ model, if demands for product 1 and product 2 in period 2 are independent and if all other parameters are held constant,*

(1) If the second-period distribution of product 1 changes from F_1^1 to F_1^2 with $F_1^1 \leq_{st} F_1^2$, then p^* increases;

(2) If the second-period distribution of product 2 changes from F_2^1 to F_2^2 with $F_2^1 \leq_{st} F_2^2$, then p^* increases;

(3) If both distributions change, with $F_1^1 \leq_{st} F_1^2$ and $F_2^1 \leq_{st} F_2^2$, then p^* increases.

Proof. Equation (27) can be rewritten as:

$$\frac{\int F_1^i(p^* - u)dF_2^i(u)}{1 - F_1^i(p^*)} = \frac{\alpha_{11} - \alpha_{21}}{\alpha_{21}}$$

where $i = 1$ or 2 indicates the relevant demand distribution. To see that (1) is true, note that when demand moves from F_1^1 to F_1^2, $F_1^1(p^* - u) \geq F_1^2(p^* - u)$ and $1 - F_1^1(p^*) \leq 1 - F_1^2(p^*)$. Therefore, the left-hand side of this equality will decrease unless p^* increases. To see (2), the numerator of the left-hand side can be written as $\int F_2^i(p^* - u)dF_1^i(u)$, and we can apply a similar analysis. To prove (3), note that stochastic dominance is preserved under convolution. Therefore, when moving to the new distribution the numerator of (27) declines, and the denominator rises, and p^* must increase to satisfy the equality. ■

Next we examine the impact on p^* of changes to the correlation between the second-period demands. Let ρ be the correlation coefficient between demand for product 1 and product 2 in period 2.

Proposition 10 *(1) $sign(dp^*/d\rho) = -sign\left[dP(d_1^2 + d_2^2 \leq p^*)/\partial\rho\right]$*

If the second-period demands of products 1 and 2 are distributed according to a bivariate normal distribution with means μ_1 and μ_2, respectively, then

(2) $\frac{dp^*}{d\rho} > 0$ if $p^* > \mu_1 + \mu_2$,

(3) $\frac{dp^*}{d\rho} < 0$ if $p^* < \mu_1 + \mu_2$,

(4) $\frac{dp^*}{d\rho} = 0$ if $p^* = \mu_1 + \mu_2$.

Proof. Equation (27) can be rewritten as:

$$P(d_1^2 \leq p^*) = 1 - \frac{\alpha_{21}}{\alpha_{11} - \alpha_{21}} P(d_1^2 + d_2^2 \leq p^*).$$

Therefore, an increase in $P(d_1^2 + d_2^2 \leq p^*)$ must lead to a decrease in p^*. To prove (2), let products 1 and 2 have normal marginal distributions $N(u_1, \sigma_1^2)$ and $N(u_2, \sigma_2^2)$, respectively, and let $d_T = d_1^2 + d_2^2$ so that d_T is also normally distributed with $u_T = u_1 + u_2$ and $\sigma_T^2 = \sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2$ so that increasing ρ increases σ_T^2. If $p^* > u_T$ then $P(d_T \leq p^*)$ decreases as ρ and σ_T^2 increases. From (1), this implies that p^* increases. Similar logic leads to (3) and (4).

■

Results (2)-(4) may be clearer if one thinks of the type-1 rationing problem as a newsvendor problem, where the demand stream in period 2 includes both type-1 and type-2 customers. Increasing ρ increases the variance of the total demand for type-1 products, and in the newsvendor problem, increasing the variance increases the optimal order quantity if the critical fractile is above the mean ($p^* > \mu_1 + \mu_2$) and decreases the optimal quantity if the critical fractile is below the mean.

3.2 Optimal capacities in the 2X2 model

One might think of the static model as a best case, for the firm is able to gather all demand information and then allocate capacity optimally. Because in the dynamic model the firm is forced to make allocation decisions before all customers have arrived, flexibility may not be used optimally. Therefore, a reasonable prediction is that the solution to the dynamic model should have equal or smaller investments in the highest-class capacity and larger investments in the lowest-class capacity, as compared to the static model. In general, our analysis and numerical experiments confirm this prediction, although there can be exceptions. In fact, given certain parameters, it may be optimal to have more class-1 capacity in the dynamic case than in the static case.

Let (x_1^{STC}, x_2^{STC}) and (x_1^{DYN}, x_2^{DYN}) be the optimal capacities for the STC and DYN models, respectively. These capacities maximize the following two objective functions. The objective function of STC is

$$\Pi^{STC}(x_1, x_2) = \underset{\mathbf{D}^1, \mathbf{D}^2}{\mathrm{E}} \left\{ \begin{array}{l} \alpha_{11} \min(d_1^1 + d_1^2, x_1) + \alpha_{22} \min(d_2^1 + d_2^2, x_2) \\ +\alpha_{21} \min\left[(d_2^1 + d_2^2 - x_2)^+, (x_1 - d_1^1 - d_1^2)^+\right] - c_1 x_1 - c_2 x_2 \end{array} \right\}.$$

The partial derivative, with respect to x_2, for STC is,

$$\frac{\partial \Pi^{STC}}{\partial x_2} = \alpha_{22} P(d_2^1 + d_2^2 > x_2) - \alpha_{21} P(d_2^1 + d_2^2 > x_2, d_T \leq x_T) - c_2. \tag{28}$$

Now, the objective function of the dynamic 2x2 model is,

$$\Pi^{DYN}(x_1, x_2)$$

$$= \underset{\mathfrak{D}^1,\mathfrak{D}^2}{\mathrm{E}} \left\{ \begin{array}{l} \alpha_{11} \min(d_1^1, x_1) + \alpha_{22} \min(d_2^1, x_2) + \alpha_{21} \min \left[(d_2^1 - x_2)^+, (x_1 - p - d_1^1)^+ \right] \\ +\alpha_{11} \min \left[d_1^2, (x_1 - d_1^1)^+ - \min \left[(d_2^1 - x_2)^+, (x_1 - p - d_1^1)^+ \right] \right] \\ +\alpha_{22} \min \left[d_2^2, (x_2 - d_2^1)^+ \right] \\ +\alpha_{21} \min \left[\begin{array}{c} \left\{ d_2^2 - (x_2 - d_2^1)^+ \right\}^+, \\ \left\{ \begin{array}{c} (x_1 - d_1^1)^+ - \\ \min \left[(d_2^1 - x_2)^+, (x_1 - p - d_1^1)^+ \right] - d_1^2 \end{array} \right\}^+ \end{array} \right] \\ -c_1 x_1 - c_2 x_2 \end{array} \right\}.$$

For convenience let $x_T = x_1 + x_2$, $x_{T-p} = x_1 - p + x_2$,and $d_T = d_1^1 + d_2^1 + d_1^2 + d_2^2$. Using techniques similar to those described by Netessine and Rudi (2003), we find the following partial derivative with respect to x_2:

$$\frac{\partial \Pi^{DYN}}{\partial x_2} = \alpha_{22} P(d_2^1 > x_2) \tag{29}$$

$$- \alpha_{21} P(d_2^1 > x_2, d_1^1 + d_2^1 \leq x_{T-p}) \tag{30}$$

$$+ \alpha_{11} P(d_2^1 > x_2, d_1^1 + d_2^1 \leq x_{T-p}, d_1^1 + d_2^1 + d_1^2 > x_T) \tag{31}$$

$$+ \alpha_{22} P(d_2^1 \leq x_2, d_2^1 + d_2^2 > x_2) \tag{32}$$

$$+ \alpha_{21} P(d_2^1 > x_2, d_1^1 + d_2^1 \leq x_{T-p}, d_1^1 + d_2^1 + d_1^2 \leq x_T, d_T > x_T) \tag{33}$$

$$- \alpha_{21} P(d_2^1 \leq x_2, d_2^1 + d_2^2 > x_2, d_T \leq x_T) \tag{34}$$

$$- c_2 \tag{35}$$

The term (31) with coefficient α_{11} on the right-hand-side merits special attention. This is the incremental profit when an additional unit of type-2 capacity leads to fewer upgrades in the first period, and thus more type-1 sales in the second period; this term is the marginal benefit due to a reduction in the cannibalization of capacity. These first-order conditions lead to the following result.

Proposition 11 *For the 2X2 model, $\partial \Pi^{DYN}(x_1, x_2)/\partial x_2 \geq \partial \Pi^{STC}(x_1, x_2)/\partial x_2$ for any capacities x_1 and x_2.*

Proof. In the expression for $\partial \Pi^{DYN}/\partial x_2$ there are two probability terms multiplied by the constant α_{22},term (29) and term (32):

$$\alpha_{22} \left\{ P(d_2^1 > x_2) + P(d_2^1 \leq x_2, d_2^1 + d_2^2 > x_2) \right\} = \alpha_{22} P(d_2^1 + d_2^2 > x_2).$$

By comparison with $\partial\Pi^{STC}/\partial x_2$,equation (28), the terms with the coefficient α_{22} are equal in $\partial\Pi^{DYN}/\partial x_2$ and $\partial\Pi^{STC}/\partial x_2$. In addition, both expressions include the term '$-c_2$'. Let Ψ denote the remaining terms in $\partial\Pi^{DYN}/\partial x_2$ (terms 30, 31, 33, and 34). We now show that Ψ is greater than or equal to the remaining term $-\alpha_{21}P(d_2^1 + d_2^2 > x_2, d_T \leq x_T)$ in $\partial\Pi^{STC}/\partial x_2$:

$$\Psi \geq \alpha_{21} \left\{ \begin{array}{l} -P(d_2^1 + d_2^2 > x_2, d_T \leq x_T) \\ +P(d_2^1 > x_2, d_T \leq x_T) \\ -P(d_2^1 > x_2, d_1^1 + d_2^1 \leq x_{T-p}, d_1^1 + d_2^1 + d_1^2 \leq x_T, d_T \leq x_T) \end{array} \right\} \tag{36}$$

$$\geq -\alpha_{21}P(d_2^1 + d_2^2 > x_2, d_T \leq x_T). \tag{37}$$

where the inequality (36) follows by replacing α_{11} with α_{21} and rearranging the probability terms. This result applies for any protection level p, including the optimal protection level p^*. Therefore, $\partial\Pi^{DYN}(x_1, x_2)/\partial x_2 \geq \partial\Pi^{STC}(x_1, x_2)/\partial x_2$. ■

Proposition 11 indicates that the marginal value of an additional unit of type-2 capacity is at least as valuable in the dynamic environment than in the static environment. The terms of the partial derivative $\partial\Pi^{DYN}/\partial x_2$ above suggest why: extra type-2 capacity can be useful for protecting against 'supply cannibalization,' upgrades of type-2 customers in the first period that lead to a shortage of type-1 capacity for type-1 customers in the second period. While Proposition 11 is not sufficient to show that $x_2^{DYN} \geq x_2^{STC}$, we have conducted thousands of numerical experiments using a wide variety of parameters and two types of distribution functions (truncated normal and uniform), and in every case, $x_2^{DYN} \geq x_2^{STC}$. We describe examples of these experiments below.

There is no analogue of Proposition 11 for type-1 capacity: $\partial\Pi^{DYN}/\partial x_1 \lesseqgtr \partial\Pi/\partial x_1$. In addition, we will see examples below in which $x_1^{DYN} \leq x_1^{STC}$ and $x_1^{DYN} > x_1^{STC}$.

In the following numerical experiments we assume that all demands are normally distributed and truncated at 0, although the coefficient of variation will be sufficiently small so that truncation does not significantly affect the results. For the STC model, we assume that the total type-1 and type-2 demands are distributed with mean $\mu_i^1 + \mu_i^2 = 100$ and standard deviations $\sigma(D_i^1 + D_i^2) = 30, i = 1, 2$. For DYN, when we split demand between the first and second periods, we will hold these total-demand parameters constant. Specifically, if a proportion r of type-i demand occurs in the first period, then $D_i^1 \sim N(100r, 30\sqrt{r})$ and $D_i^2 \sim N(100(1-r), 30\sqrt{(1-r)})$, so that the standard deviation of the *total* demand is 30. In the first set of experiments described here, the contribution margin and cost parameters are $\alpha_{11} = 40, \alpha_{21} = 15, \alpha_{22} = 20$, $c_1 = 12$, and $c_2 = 10$. These parameters imply that the newsvendor critical ratios for type-1 and type-2 are 0.7 and 0.5, respectively.

The numerical experiments examine four models: NV, STC, DYN, and a *Greedy heuristic*, the dynamic model with no rationing (protection level $p = 0$). The first-order conditions for STC and DYN are described above, and the solution to the newsvendor problem is well known. The optimal capacities of each model were found numerically, using Monte Carlo Integration and a simple search procedure (for details on the search procedure, see Section 4.1).

We find that optimal capacities for the static and dynamic models diverge significantly when (i) a majority of type-2 demand occurs in the first period and (ii) a majority of type-1 demand occurs in the second period. Therefore, in the dynamic model we 'unbalance' the demand to emphasize this point. Given that r is the proportion of type-2 demand in the first period and $1 - r$ is the proportion of type-1 demand in the first period, we vary r from 0.4 to 1.

For example, when $r = 0.5$, demands for both products are distributed equally between periods. In this case there is almost always insufficient demand in the first period of the dynamic model to require any upgrading, so that there is little risk of supply cannibalization, type-1 capacity is rarely rationed, the particular rationing policy does not matter, and there is little difference between the static and dynamic models. However, as r rises, the early appearance of type-2 demand and the late appearance of type-1 demand forces the firm to either upgrade type-2 demand or ration type-1 products. The model with $r = 1$ is analogous to the standard yield management problem, in which low-fare passengers arrive first, followed by high-fare passengers.

Figures 1 and 2 show the optimal type-1 and type-2 capacity values, respectively, for each model. In Figure 1 the dynamic model's optimal type-1 capacity, x_1^{DYN} is consistently below the optimal capacity from the static model, x_1^{STC}, although we have found that the opposite can be true (see below). A more pronounced pattern is shown in Figure 2, where we see that the optimal type-2 capacities can be significantly higher in the dynamic model ($x_2^{DYN} \geq x_2^{STC}$). The extra type-2 capacity acts as a buffer to prevent cannibalization of more lucrative type-1 capacity. This role for type-2 capacity is particularly important when there is no rationing, thus inflating the optimal type-2 capacity.

To see that it is possible to have $x_1^{DYN} > x_1^{STC}$, consider an experiment with the following margin and cost parameters: $\alpha_{21} = 4, \alpha_{22} = 5$, and $c_2 = 1$ (we will try a variety of values for both α_{11} and

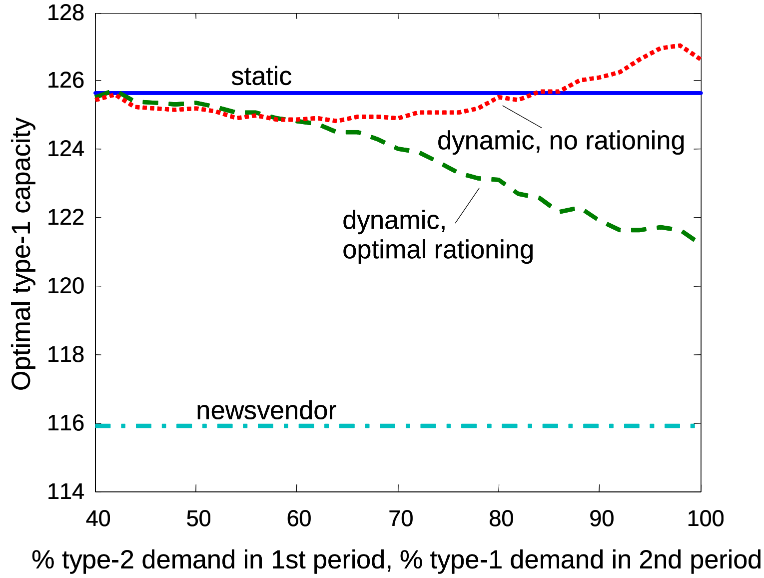

Figure 1: Optimal type-1 capacity

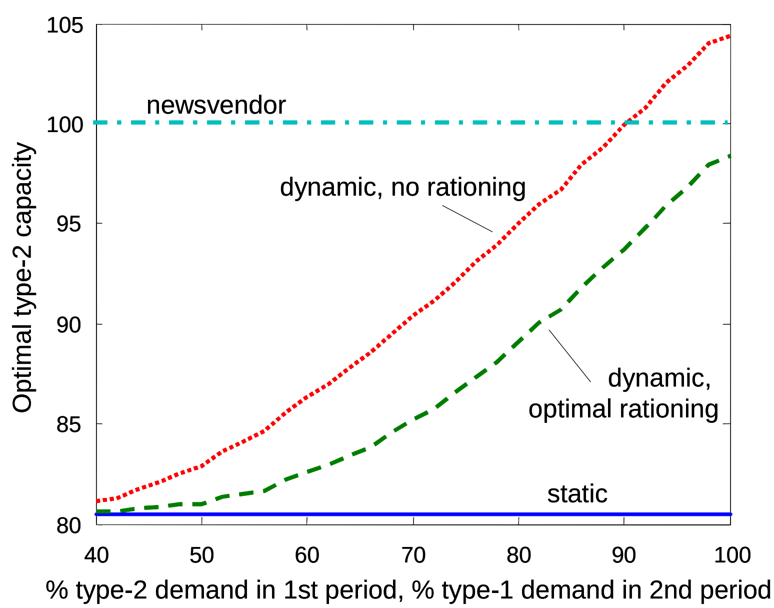

Figure 2: Optimal type-2 capacity

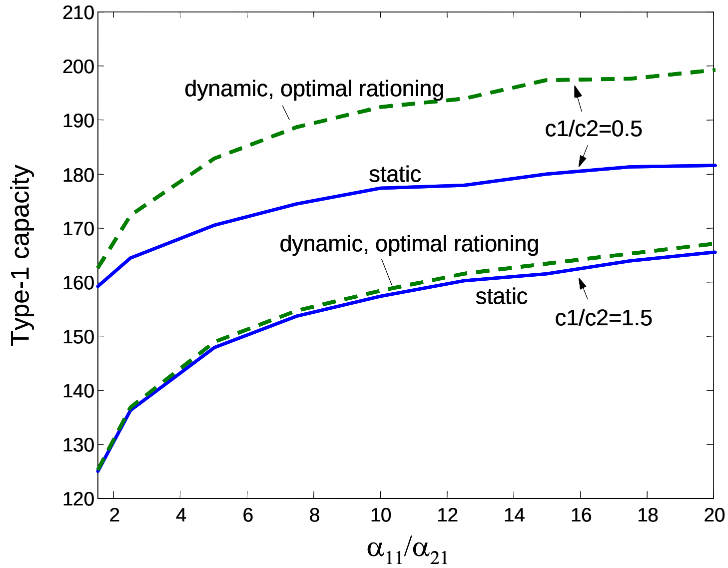

Figure 3: Optimal type-1 capacity can be larger in the dynamic model

c_1). The total demands are still $N(100, 30)$, and we assume that $r = 1$, so that in the dynamic model there is no type-1 demand in the first period and no type-2 demand in the second period. The parameters α_{22} and c_2 imply that the newsvendor problem's critical ratio is 0.8 for product 2. This ratio will be substantially higher for product 1 in the following examples, for we will vary α_{11} from 5 to 80 and will use two low values of c_1: 1.5 and 0.5. The second value indicates that the initial purchase cost of product 1 is less than the cost of product 2, although the usage cost may be significantly greater for product 1 than product 2.

Figure 3 shows the optimal type-1 capacities from the dynamic and static procedures, x_1^{DYN} and x_1^{STC}. Here the optimal dynamic type-1 capacities are higher than the optimal static capacities. This difference is again caused by the problem of supply cannibalization in the dynamic case. For demand realizations in which cannibalization occurs, an additional unit of type-1 product always has the marginal value $\alpha_{11} - \alpha_{21}$ in the dynamic case, but may have no value in the static case. This effect is largest when the profitability of a type-1 sale is greatest, i.e., when α_{11} is large and when c_1 is low. In addition, this risk of supply cannibalization is even greater when protection limits are lowered. If there is no rationing, the differences between the optimal dynamic and static capacities are consistently larger than the differences seen in Figure 3.

4 The value of optimal capacity and allocation: numerical experiments

This section describes in detail the results from numerical studies designed to understand how the parameters of the model affect two quantities, (i) the value of optimal upgrading and (ii) the value of using the capacity that is strictly optimal, given that optimal upgrading will be used (rather than using capacity that is optimal for the simpler, static model). Here we calculate the value of optimal upgrading as the difference between the profit generated from the DYN model and the profit generated from two simpler heuristics, the NV model and a *Greedy heuristic* in which $y^t_{k+1,k} = \left[\left(d^t_{k+1} - x^t_{k+1}\right)^+ \wedge \left(x^t_k - d^t_k\right)^+\right]$ for $k = 1...N-1$, i.e., the protection limits are 0 and all possible upgrading is performed in each period. We calculate the value of strictly optimal capacity as the difference between the profits generated by DYN and a *Hybrid heuristic* in which the initial capacity is optimal for the STC problem and then optimal upgrading is used once customers begin arriving.

We assess the impact of model parameters on the quantities (i) and (ii) described above. In particular, we examine the impact of three attributes of the model:

1. *Availability of advance demand information.* In the one-period model (STC), all demand information is available when all allocation occurs, so that capacity may be assigned to customers without any possibility of cannibalization. In practice, demand information may become available in small increments over time, and we examine the impact of the incremental release of demand information by varying the number of periods in the DYN model.

2. *Economic parameters*, the contribution margins α_{ij} and the initial capacity costs c_j.

3. *Demand parameters*, the variance and within-period correlations of the demand.

Below we first describe a large number of experiments with a 2-product model - we evaluated the profits generated by NV, STC, DYN, the Greedy heuristic and the Hybrid heuristic for almost 5000 parameter combinations. From these we assessed the impact of the model parameters described above. Then we tested a smaller number of 5-product models, and found that the insights developed from the 2-product model for attributes (1) and (2), above, applied to these models with larger numbers of products as well. In all experiments we chose parameter ranges that were bounded either by the assumptions of the model (specifically, assumptions A1-A3), or by limits imposed by real-world applications (e.g., the unit cost of product 1 should be greater than the unit cost of product 2, $c_1 > c_2$).

4.1 Finding the optimal capacity

The STC model, the DYN model, and the Greedy heuristic all begin by finding the optimal integral initial capacities $\mathbf{X}^1$. To find these capacities we use a neighborhood search algorithm that begins at the newsvendor solution (the solution to NV), and then evaluates the objective function at each neighbor around that solution. We define a 'neighbor' as a capacity vector with 1 less, the same, or 1 more unit of capacity for each product; a capacity vector with all elements greater than 0 has $3^N - 1$ neighbors. After evaluating the profit function at each neighbor, the algorithm moves to the neighbor with the highest value. This process is repeated until no neighbor has a higher value.

Although DYN and STC are concave functions when the capacities and protection levels are continuous (see Lemma 2), our algorithm only evaluates the functions on the integer lattice and therefore may not find the true optimal capacity. In addition, the continuous version of the Greedy heuristic may not necessarily be concave. To determine the effectiveness of the search procedure, we searched exhaustively for the true optimal capacities for 625 of the 2-product experiments described below. In every case the neighborhood search algorithm found the optimal capacity vectors for DYN, STC and the Greedy heuristic. For the remaining 2-product problems and for the higher-dimensional problems, we have no reason to believe that the capacity solutions found with this heuristic are not equal to, or close to, the optimum.

4.2 Parameters for the 2-product experiments

We conduct two sets of experiments with 2 products, one focusing on the financial parameters and the number of periods, another focusing on the demand parameters. We will call the first set the 'economic scenarios' and the second set the 'demand scenarios.' For both sets of experiments, the total demand over all periods is 60 for each product. Demand for product 1 rises linearly over the horizon, i.e., x is the mean demand for period 1, $2x$ is the mean demand for period 2, etc., and demand for product 2 falls linearly. For example, if there are 5 time periods, then the mean period-by-period demand for product 1 is [4, 8, 12, 16, 20] and the mean demand for product 2 is [20, 16, 12, 8, 4]. As is assumed throughout the paper, demand is independent across periods. In addition, in all experiments α_{22} is normalized to 1.

For the economic scenarios, demand follows a Poisson distribution within each period. The following table describes the remaining parameters for the first set of experiments. Because we evaluated the models with all combinations of all parameters, there are $4 * 5 * 5 * 5 * 5 = 2500$

scenarios.

parameter	values	comments
T	$2, 5, 10, 20$	demand within each period is adjusted so that total demand equals 60 for both products
α_{11}	$1.2, 1.4, 1.6, 1.8, 2.0$	
α_{21}	$\alpha_{21} = \gamma\alpha_{22}$, where $\gamma = \{0.5, 0.6, 0.7, 0.8, 0.9\}$	As γ increases, the value of upgrading increases.
c_2	$c_2 = \beta\alpha_{22}$, where $\beta = \{0.5, 0.6, 0.7, 0.8, 0.9\}$	As β increases, the cost of type-2 capacity increases (relative to the profit margin)
c_1	$c_1 = c_2 + \delta(\alpha_{11} - c_2)$, where $\delta = \{0.3, 0.4, 0.5, 0.6, 0.7\}$	As δ increases, the cost of type-1 capacity increases. The formula ensures that c_1 is within a range from c_2, a lower bound, to α_{11}, an upper bound.

Note that because demand follows a Poisson distribution, the variance of the total demand remains constant as the number of periods changes.

For the demand scenarios, demand in each period follows a bivariate normal distribution. Demand values were truncated at 0 and rounded to the nearest integer. In all experiments, $\alpha_{22} = 1$ and $\alpha_{11} = 1.5$. The following table describes the remaining parameters. We evaluated the models with all combinations of all parameters, so that there are $3 * 4 * 7 * 3 * 3 * 3 = 2268$ scenarios.

parameter	values	comments
T	$2, 5, 10$	demand within each period is adjusted so that total demand equals 60 for both products
coefficient of variation of each demand within each period	$0.1, 0.2, 0.3, 0.4$	
correlation coefficient between demands within each period	$-0.9, -0.6, -0.3, 0, 0.3, 0.6, 0.9$	
α_{21}	$\alpha_{21} = \gamma\alpha_{22}$, where $\gamma = \{0.5, 0.7, 0.9\}$	As γ increases, the value of upgrading increases.
c_2	$c_2 = \beta\alpha_{22}$, where $\beta = \{0.5, 0.7, 0.9\}$	As β increases, the cost of type-2 capacity increases (relative to revenue)
c_1	$c_1 = c_2 + \delta(\alpha_{11} - c_2)$, where $\delta = \{0.3, 0.5, 0.7\}$	As δ increases, the cost of type-1 capacity increases. The formula ensures that c_1 is within a range from c_2, a lower bound, to α_{11}, an upper bound.

4.3 The value of using optimal capacity in the dynamic model

We first compare the profits generated by DYN and the Hybrid heuristic. This is of interest because finding the optimal capacity for DYN is significantly more difficult than finding the optimal capacity for STC, because the capacity optimization in DYN must take the future dynamic rationing policy into account and therefore must evaluate the full dynamic program, given any initial capacity level. This can be cumbersome, even when taking advantage of the bounds described in Proposition 6. The value of STC given any capacity level, however, requires few relatively simple calculations.

For each scenario we found the percentage increase in the expected profit due to using the optimal capacity for DYN rather than using the Hybrid heuristic with the STC-optimal capacity. That is,

$$\text{value of using optimal capacity} \equiv \frac{\Pi^{DYN}(\mathbf{X}^{DYN}) - \Pi^{DYN}(\mathbf{X}^{STC})}{\Pi^{DYN}(\mathbf{X}^{DYN})}$$

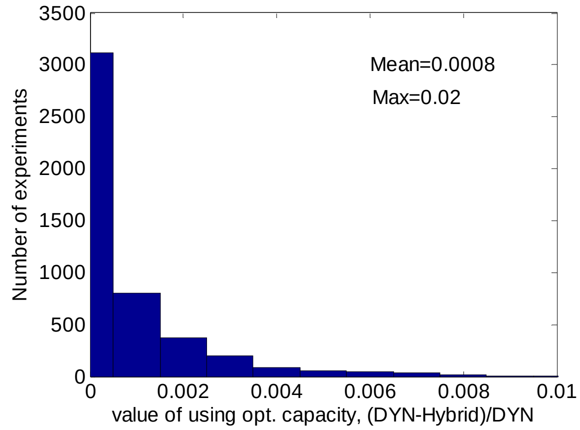

Figure 4: Value of using optimal capacity for DYN vs. using the Hybrid heuristic

where $\Pi^X(\mathbf{X}^Y)$ is the profit from model X when starting with capacity that is optimal for model Y. Figure 4 shows a histogram of this value from all $2500 + 2268 = 4769$ scenarios.

For 48% of the scenarios, the DYN and STC capacities were identical. Overall, the average difference between the DYN and Hybrid profits was 0.0008 (0.08%) of the DYN profit. The 90th percentile of the differences among scenarios is 0.2%, and the maximum difference is 2%. Therefore, ignoring the dynamic rationing policy when finding the initial capacity by using STC for capacity optimization, and then using optimal rationing, almost always performs as well as the much more complex capacity optimization in DYN.

To avoid convolving multiple affects, however, for the remainder of this Section we will not use the Hybrid heuristic and instead will compare the optimal dynamic profit, $\Pi^{DYN}(\mathbf{X}^{DYN})$, with profits from heuristics that do not use optimal upgrading.

4.4 The value of optimal upgrading

In this section we compare the profit from DYN with the profit from the NV model (no upgrading) and the Greedy heuristic (myopic upgrading). First we examine how DYN and the heuristics perform as the number of periods changes. This provides us with information about the value of advanced demand information. Then we examine the impacts of financial and demand parameters.

One method for describing the effects of the parameters would be to choose a scenario, vary one parameter, examine whether the effect on profits follows a particular pattern (say, monotonicity), and then repeat with another scenario, and another. If the pattern is consistent over a large number of scenarios, then we have good evidence for the general applicability of the effect. In our problem, however, patterns such as monotonicity are disrupted by the effects of integrality, i.e. in general profit might go down as a parameter rises, but integrality in capacity or protection levels may force profit to go *up* for one particular parameter value in one particular scenario. Therefore, below we report aggregate statistics such as medians and percentiles that are collected across hundreds of scenarios, given certain levels of a particular parameter; see Figures 5-11 below. By pooling the results of many scenarios, we see patterns that might be hidden within particular scenarios by the effects of integrality.

4.4.1 Advance demand information

The model in DYN is equivalent to STC if the firm has a perfect demand forecast: if the firm knows exactly who is coming and when, then it can optimally allocate capacity among customers as if all customers had arrived in the same time period, as in STC. In this section we examine the question, how valuable is advance demand information?

As the number of periods in DYN increases, information availability may decrease and the allocation may be less effective than the allocation in STC. In a problem with 2 periods, 50% of the allocation decisions are made in the first period with 50% of the demand information available, and the other 50% are made with 100% of the demand information, so that one could say that the average information available for allocation is $(0.5)(0.5) + (0.5)(1) = 0.75$. With 3 periods, the average is $(1/3)(1/3) + (1/3)(2/3) + (1/3)(1) = 2/3$. In general, for a T-period problem,

$$\text{average information available for allocation}=\frac{T(T+1)}{2T^2},$$

so that the average information available declines with T, from 1 when $T = 1$ to 1/2 as $T \to \infty$.

The change in profits as T increases, however, is not clear, for stockouts are less likely in early periods and subdividing periods may create beneficial combinations of additional demand information and upgade opportunities. Consider a change from $T = 2$ to $T = 3$. Instead of one decision epoch at $T/2$,there are now two decision epochs at $T/3$ and $2T/3$. Under certain demand scenarios stockouts may be unlikely at the $T/3$ epoch, so that most upgrading decisions are made at the $2T/3$ epoch, when there is *more* demand information than at the $T/2$ epoch.[1]

[1] We are grateful to an anonymous reviewer for noting that profits may not necessarily decline in T and for providing us with this counterexample.

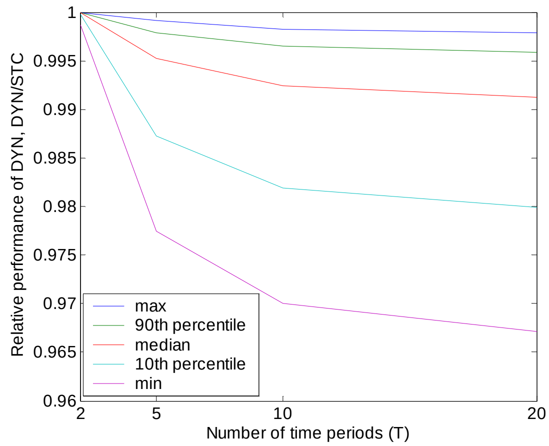

Figure 5: Effect of the number of time periods (T) on the relative performance of DYN vs. STC

In our numerical experiments, however, we do find that profits tend to decline as T rises. Figure 5 displays statistics for $\Pi^{DYN}(\mathbf{X}^{DYN})/\Pi^{STC}(\mathbf{X}^{STC})$ for $T = 2, 5, 10$ and 20 from the 2500 economic scenarios (a plot for $T = 2$, 5,and 10 for the 2268 demand scenarios is similar). As the number of periods rises the performance of DYN relative to the upper bound from STC decreases. The performance is quite good even when $T = 20$, however. The median performance of DYN is less than 1% from STC, and the worst-case performance is within 3.5%. This implies that, under the assumptions stated in this paper, optimal rationing policies can be as valuable as nearly perfect demand information.

We next ask whether the optimal upgrading itself is valuable as T varies, that is, when is it worth implementing the DYN policy rather than the simpler NV or Greedy heuristics? Figure 6 shows the median and minimum performances of the two heuristics over the economic scenarios (again, the results from the demand scenarios are similar). The NV performance does not vary with the number of time periods because, with no upgrading, there is no risk of cannibalization between time-periods. The NV median performance is 98.6% of STC, quite close to DYN's median performance, which approaches 99% as the number of time-periods increase. For certain economic scenarios, however, DYN can be close to 100% while NV is close to 96%, and the differences can be even greater for certain demand scenarios. Below we examine when, and why, those differences occur.

Finally, note that the Greedy heuristic does well with just 2 time periods, but its performance is much worse with $T = 5, 10$ and 20. In fact, with $T = 2$, Greedy performs better than NV in 93% of the scenarios. However, with $T = 5, 10$ and 20, NV outperforms Greedy in over 99% of

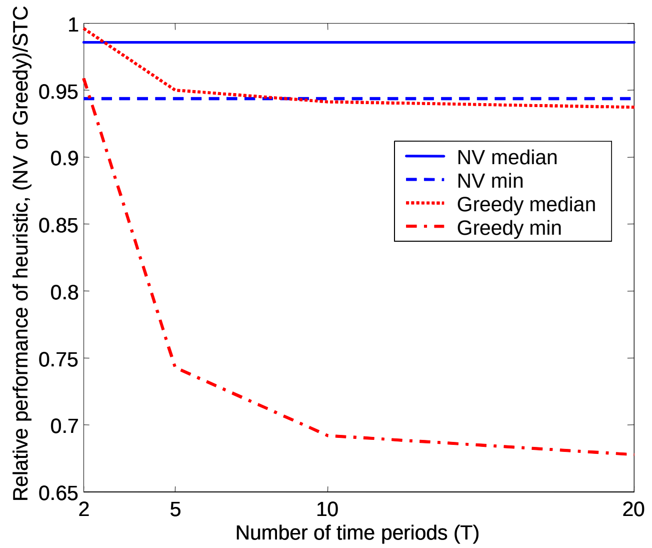

Figure 6: Effect of the number of time periods (T) on the relative performance of the heuristics NV and Greedy vs. STC

the scenarios. This is because with $T = 2$, half of the allocation decisions are made with perfect information, limiting the impact of cannibalization, and giving an advantage to the Greedy heuristic. Because the performances of all the algorithms (STC, DYN, NV and Greedy) are nearly identical with $T = 2$, from now on we will restrict our attention to scenarios with $T > 2$ (1,875 from the economic scenarios and 1,512 from the demand scenarios), and we will focus on the difference between DYN and NV profits, given our observation that the Greedy algorithm performs poorly when T is large.

4.4.2 Impact of economic parameters

We now examine in more detail the question, when is it worthwhile to implement the DYN policy rather NV? We compare the profits from DYN and NV (as a proportion of STC profits) for each economic scenario, to assess how the value of optimal upgrading changes in response changes in the financial parameters of the model. That is, we evaluate for each scenario,

$$\text{value of using optimal upgrading} \equiv \frac{\Pi^{DYN}(\mathbf{X}^{DYN}) - \Pi^{NV}(\mathbf{X}^{NV})}{\Pi^{STC}(\mathbf{X}^{STC})}.$$

First we examine the impact of α_{11}, the margin for the high-value product. Figure 14 shows the value of using optimal upgrading as α_{11} changes. The statistics (min, 10th percentile, median, etc.) are calculated from the 1,875 economic scenarios with $T > 2$. In general, the value of optimal

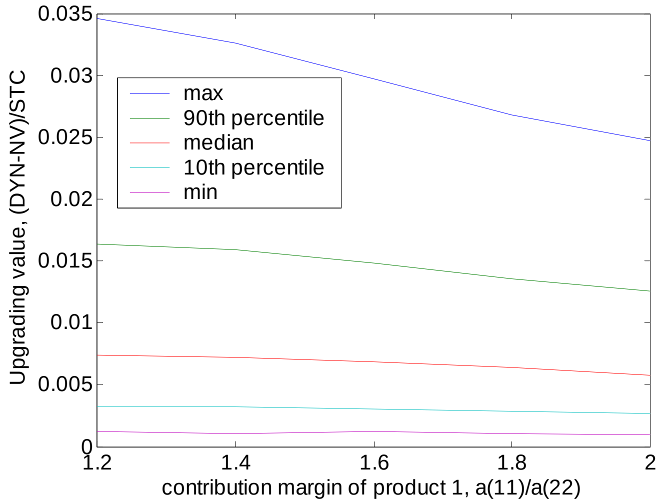

Figure 7: Value of optimal upgrading as the contribution margin of product 1 varies

upgrading declines as α_{11} rises. A large α_{11} implies that parallel revenues from sales of product 1 dominate profits, so that the value of upgrading is relatively small. Therefore, the difference between DYN and NV is smaller when α_{11}/α_{22} is large.

Figure 8 shows the impact of α_{21}, the profit from the upgrade. It is intuitive that as the value of the upgrade rises, the value of optimal upgrading will increase, compared to the no-upgrading policy of the NV heuristic.

Figure 9 shows how the value of optimal upgrading decreases in $c_1 - c_2$ (to produce this plot, the value of $c_1 - c_2$ for each scenario was rounded to the nearest 10th, and then the sets of scenarios centered at each 10th were used to calculate the statistics for 'max', '90th percentile', etc.). When c_1 is large (and therefore $c_1 - c_2$ is large), the optimal initial capacity of the type-1 product is low. Therefore upgrading is risky because it may cannibalize precious, and rare, type-1 capacity. On the other hand, when c_2 is large (and therefore $c_1 - c_2$ is small), the optimal initial capacity of the type-2 product is low. In this case, there will be many upgrading opportunities, so the value of upgrading is high. A very similar plot can be generated with $c1/c2$ on the horizontal axis.

4.4.3 Impact of demand parameters

Here we examine how demand variance and correlation affect the value of optimal upgrading by examining the demand scenarios described in the second table, above. Figure 10 shows that the value of upgrading increases in the coefficient of variation (CV). To understand this, it is helpful

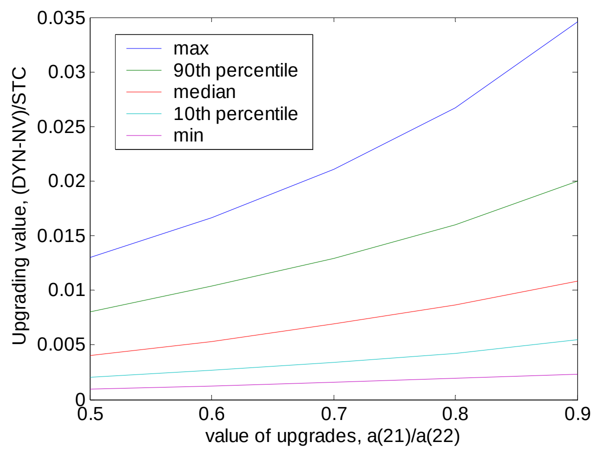

Figure 8: Value of optimal upgrading as the contribution margin of upgrading varies

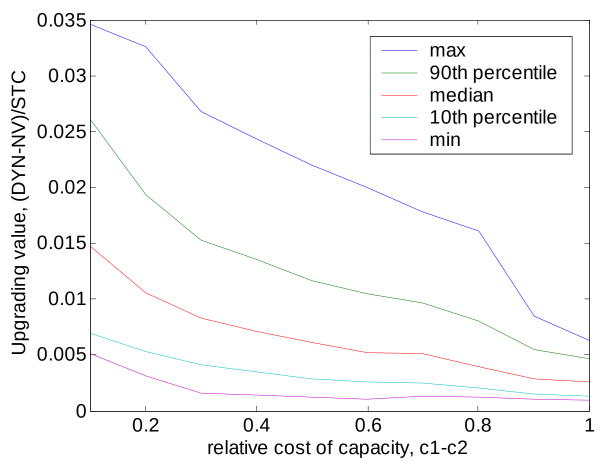

Figure 9: Value of optimal upgrading as the relative costs of products 1 and 2 vary

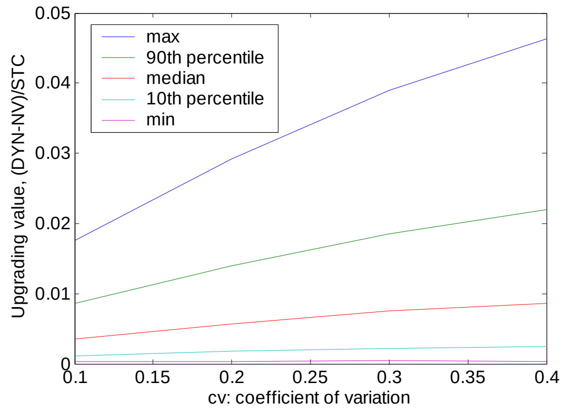

Figure 10: Value of optimal upgrading as the coefficient of variation (CV) of demand varies

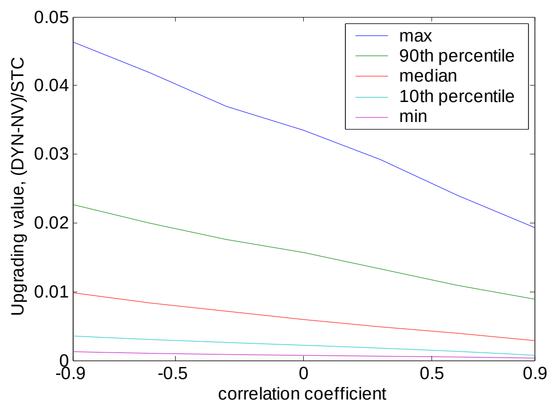

Figure 11: Value of optimal upgrading as the coefficient of correlation between products 1 and 2 varies

to consider the extreme case: with a CV=0, demand is perfectly predictable, every customer will be assigned capacity of the correct type, and no upgrading is needed. As demand uncertainty increases, the option to upgrade becomes more valuable.

Figure 11 shows that the value of upgrading decreases in ρ. A large, positive ρ implies that demands for the products move together and therefore either (i) both demands are low, so that no upgrades are needed or (ii) both demands are high, so that upgrades are needed but type-1 capacity is not available to satisfy type-2 customers. Negative correlation, on the other hand, provides more opportunities for upgrading.

To summarize, the optimal upgrading policy generated by DYN is most valuable when,

1. α_{11}/α_{22} is close to 1. If parallel contribution margins of products 1 and 2 are close together, then type-1 parallel revenues do not dominate and optimal upgrading is valuable.

2. α_{21}/α_{11} is close to 1. If upgrades have a relatively high value, then using optimal upgrading provides significant profits above the newsvendor solution with no upgrading.

3. $c1 - c2$ is close to 0. If c_1 is low and c_2 is high, then it is optimal to invest in large amounts of type-1 and few type-2's, so that there are many opportunities to upgrade.

4. demand variance is high. With uncertain demand, mismatches between demand and capacity are more likely to occur, so that optimal upgrading becomes useful.

5. demand correlation between products in low. Under low correlation, it is more likely that a stock-out for a low-type product is paired with a surplus of a higher-type product, increasing the value of optimal upgrading.

4.5 Experiments with 5 products

We next describe experiments with a 5-product model and see that these results are consistent with the results described above. Note, however, that these observations are based on a significantly smaller number of experiments. We begin with a base-case scenario and then change each parameter and examine the impact as we move away from the base case, so that each parameter is changed one-at-a-time. That is, we do not look at all combinations of parameters as we did for the 2-product scenarios, above.

For all experiments with 5 products, we assume the following:

- Number of periods: $T = 10$

- Average demand: 20 units for each product in total over all 10 periods.

- Demand distribution: Poisson. For product 1, demand is linearly increasing over time, as described in Section 4.2, above. Demand for product 2 increases at half the rate of product 1 (half the total demand is distributed evenly to every period, and there is a small increase over time on top of that). For product 5, demand is linearly decreasing, as described in Section 4.2. Demand for product 4 decreases at half the rate of product 5 (half the total demand is distributed evenly to every period, and there is a small decrease over time on top of that). Finally, demand for product 3 is stable over all 10 periods.

- Parallel contribution margins: $\alpha_{11} = 2$, $\alpha_{55} = 1$, and $\alpha_{ii}, i = 2, 3$ and 4,are linearly interpolated between α_{11} and α_{55}.

- Upgrade contribution margins: $\alpha_{i+1,i}/\alpha_{i,i}$=0.5. Note that with 5 products and the margins α_{ii} described above, the ratio $\alpha_{i,i-1}/\alpha_{i,i}$ can be at most 5/9 without violating assumptions A1-A3.

- Unit capacity costs: we let $c_i = \beta_i \alpha_{i,i}$, and for the base case, $\beta_1 = 0.55$, $\beta_5 = 0.9$, and $\beta_i, i = 2, 3$ and 4, are linearly interpolated between β_1 and β_5. This implies that in the base case $c_1 = 1.1$, $c_5 = 0.9$, and $c_i\,, i = 2, 3$ and 4, are distributed between these two.

4.5.1 Calculating protection limits and capacities with 5 products

To find the protection limits in DYN we used the one-product bounds described in Section 5.2 of the paper. In these experiments the bounds were usually tight: out of 27,000 protection levels calculated, over 99% of the gaps $\nabla_1(\mathbf{X})$ were 0, and the maximum gap was 1. Therefore, the approximate protection levels we used in the 5-product experiments were essentially optimal. Given these protection levels, we used the neighborhood search algorithm described in Section 4.1 to find the optimal capacities for DYN, STC, and the Greedy heuristic. All converged quickly, and based on our experiments with 2 products, we believe that the search algorithm found capacities that are, or are close to, the globally optimal capacities.

4.5.2 Results of the experiments with 5 products

In the base case, profits from the DYN and the Hybrid heuristic are identical, and both achieve profits that are over 99% of the upper-bound profit from STC. Using no upgrades (the NV solution) achieves 96% of STC profits, and the Greedy heuristic achieves 92% of STC.

Figures 12-15 show the results when parameters T, $\alpha_{i+1,i}/\alpha_{i,i}$, α_{11}, and $c_1 - c_5$ are varied from the base case. Following are a few general observations that apply to all of these experiments. These observations are consistent with the results of the 2-product experiments.

1. Profits from DYN and the Hybrid heuristic are nearly identical, so that using optimal capacity for STC, when paired with optimal upgrading, produces results that are close to results when using the optimal capacity for DYN;

2. Profits from DYN (or the Hybrid heuristic) are consistently within 1% of STC, so that perfect demand information has relatively little value as long as optimal upgrading is used;

3. For $T \geq 5$, profits from NV dominate profits from the Greedy heuristic.

Figure 12 demonstrates all three of these points: relative profits from DYN and the Hybrid heuristic are close together at the top of the figure, both are within 1% of STC, and the results from NV are superior to the Greedy heuristic for $T \geq 5$.

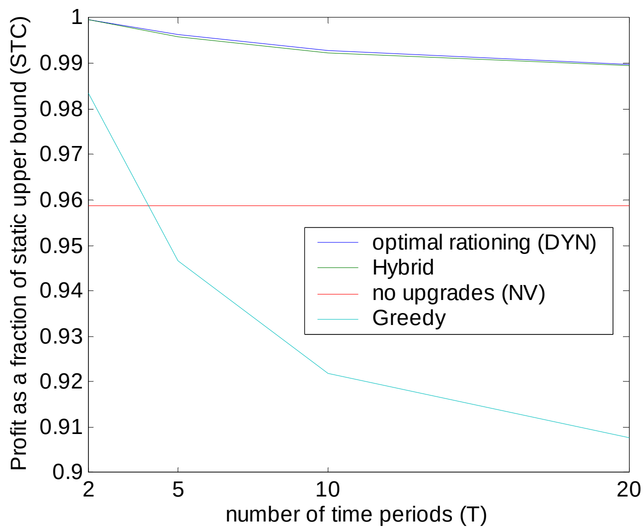

Figure 12: Value of optimal upgrading in the 5-product problem as the number of periods, T, varies

Figure 13 displays profits as a fraction of STC as we vary the relative value of upgrades, $\alpha_{i+1,i}/\alpha_{i,i}$. As the value of upgrades rises, the relative performance of the NV solution declines, so that the value of optimal upgrading increases.

To generate Figure 14, we held $\alpha_{55} = 1$, varied α_{11} from 1.1 to 5, and for each of these scenarios calculated $\alpha_{ii}, i = 2, 3$ and 4 by linearly interpolating between α_{11} and α_{55}. Although the effect is weak here, we see that as α_{11}/α_{55} declines, the NV profit declines (relative to STC), and the value of optimal upgrading rises. As we discussed above, when α_{11}/α_{55} is large, parallel revenues from high-value products represent a larger proportion of the profit, so that optimal upgrading provides relatively less value.

Finally, to generate Figure 15, we set $\beta_1 = 0.5(1+\varepsilon)$, $\beta_5 = 1-\varepsilon$, and $\varepsilon = \{0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$. As a result, $c_1 - c_5$ varies from 0.1 to 1. (For each value of ε we also set $\beta_i, i = 2, 3$ and 4, to be

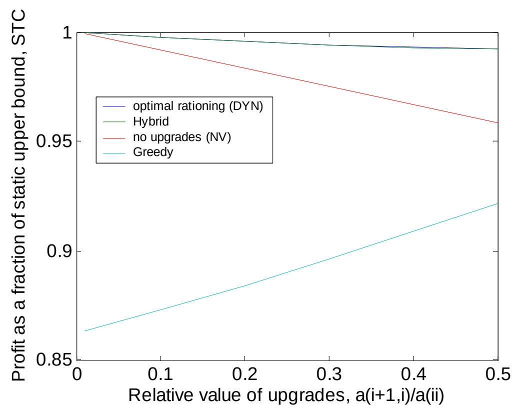

Figure 13: Value of optimal upgrading in the 5-product problem as the contribution margin of upgrading varies

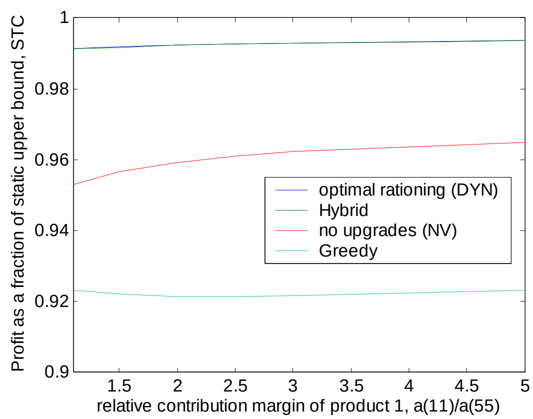

Figure 14: Value of optimal upgrading in the 5-product example as the contribution margin of product 1 varies

linearly interpolated between β_1 and β_5). As was true for the 2-product experiments, when $c_1 - c_2$ is low, the optimal initial capacities of the high-value products are relatively high, while the optimal initial capacities of the low-value products are relatively low. Therefore there will be many upgrading opportunities and the value of optimal upgrading is relatively high, when compared to the profit from the NV model.

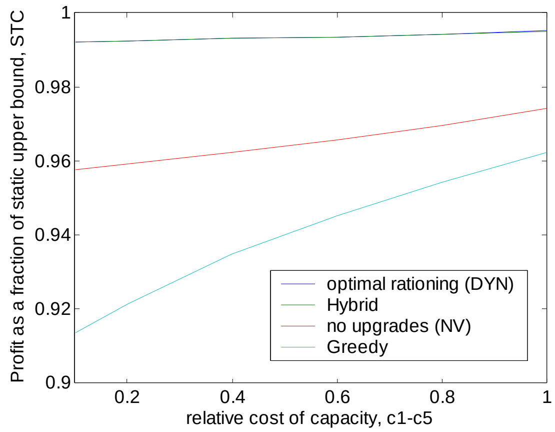

Figure 15: Value of optimal upgrading in the 5-product example as the relative cost of product 1 varies

Appendix: Parameters of the experiments to evaluate protection limit bounds

Demand Parameters

These 12 sets of demand parameters are used for both realistic and extreme sets of scenarios:

1. Mean demands = 2 units for all products in all 10 periods. Therefore, the total demand for each product is 20 units.

2.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	0	0	1	1	2	2	3	3	4	4	20
2	1	1	1	2	2	2	2	3	3	3	20
3	2	2	2	2	2	2	2	2	2	2	20
4	3	3	3	2	2	2	2	1	1	1	20
5	4	4	3	3	2	2	1	1	0	0	20

3.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	0	0	1	1	2	2	3	3	4	4	20
2	0	0	1	1	2	2	3	3	4	4	20
3	1	1	1	2	2	2	2	3	3	3	20
4	1	1	1	2	2	2	2	3	3	3	20
5	4	4	3	3	2	2	1	1	0	0	20

4.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	4	0	4	0	4	0	4	0	4	0	20
2	0	4	0	4	0	4	0	4	0	4	20
3	4	0	4	0	4	0	4	0	4	0	20
4	0	4	0	4	0	4	0	4	0	4	20
5	4	0	4	0	4	0	4	0	4	0	20

5.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	5	5	5	5	5	5	5	5	5	5	50
2	4	4	4	4	4	4	4	4	4	4	40
3	3	3	3	3	3	3	3	3	3	3	30
4	2	2	2	2	2	2	2	2	2	2	20
5	1	1	1	1	1	1	1	1	1	1	10

6.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	0	0	2	2	4	4	8	8	10	12	50
2	1	2	3	3	4	4	5	5	6	7	40
3	3	3	3	3	3	3	3	3	3	3	30
4	4	4	3	3	2	2	1	1	0	0	20
5	2	2	2	1	1	1	1	0	0	0	10

7.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	0	0	2	2	4	4	8	8	10	12	50
2	0	2	2	2	2	4	4	6	8	10	40
3	0	1	1	1	2	3	4	4	6	8	30
4	0	0	1	1	2	2	3	3	4	4	20
5	2	2	2	1	1	1	1	0	0	0	10

8.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	0	10	0	10	0	10	0	10	0	10	50
2	8	0	8	0	8	0	8	0	8	0	40
3	0	6	0	6	0	6	0	6	0	6	30
4	4	0	4	0	4	0	4	0	4	0	20
5	0	2	0	2	0	2	0	2	0	2	10

9.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	1	1	1	1	1	1	1	1	1	1	10
2	2	2	2	2	2	2	2	2	2	2	20
3	3	3	3	3	3	3	3	3	3	3	30
4	4	4	4	4	4	4	4	4	4	4	40
5	5	5	5	5	5	5	5	5	5	5	50

10.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	0	0	0	1	1	1	1	2	2	2	10
2	0	0	1	1	2	2	3	3	4	4	20
3	3	3	3	3	3	3	3	3	3	3	30
4	7	6	5	5	4	4	3	3	2	1	40
5	12	10	8	8	4	4	2	2	0	0	50

11.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	0	0	0	1	1	1	1	2	2	2	10
2	0	0	1	1	2	2	3	3	4	4	20
3	0	1	1	1	2	3	4	4	6	8	30
4	0	2	2	2	2	4	4	6	8	10	40
5	12	10	8	8	4	4	2	2	0	0	50

12.

	time period										
product	1	2	3	4	5	6	7	8	9	10	total demand
1	0	2	0	2	0	2	0	2	0	2	10
2	4	0	4	0	4	0	4	0	4	0	20
3	0	6	0	6	0	6	0	6	0	6	30
4	8	0	8	0	8	0	8	0	8	0	40
5	0	10	0	10	0	10	0	10	0	10	50

Initial Capacity: Realistic Scenarios

	parameter set			
product	1	2	3	4
1	20	30	10	10
2	20	20	20	30
3	20	10	30	20
4	20	20	20	20
5	20	20	20	20

Initial Capacity: Extreme Scenarios

	parameter set			
product	1	2	3	4
1	20	40	1	1
2	20	1	45	25
3	20	40	1	45
4	20	1	45	25
5	20	20	1	1

Contribution Margins: Realistic Scenarios

1.

$\alpha_{11} = 15$				
$\alpha_{21} = 7$	$\alpha_{22} = 14$			
	$\alpha_{32} = 6$	$\alpha_{33} = 13$		
		$\alpha_{43} = 6$	$\alpha_{44} = 12$	
			$\alpha_{54} = 5$	$\alpha_{55} = 10$

2.

$\alpha_{11} = 15$				
$\alpha_{21} = 3$	$\alpha_{22} = 14$			
	$\alpha_{32} = 3$	$\alpha_{33} = 13$		
		$\alpha_{43} = 2$	$\alpha_{44} = 12$	
			$\alpha_{54} = 2$	$\alpha_{55} = 10$

3.

$\alpha_{11} = 15$				
$\alpha_{21} = 2$	$\alpha_{22} = 14$			
	$\alpha_{32} = 2$	$\alpha_{33} = 13$		
		$\alpha_{43} = 2$	$\alpha_{44} = 12$	
			$\alpha_{54} = 7$	$\alpha_{55} = 10$

Contribution Margins: Extreme Scenarios

1.

$\alpha_{11} = 15$				
$\alpha_{21} = 12$	$\alpha_{22} = 14$			
	$\alpha_{32} = 1$	$\alpha_{33} = 13$		
		$\alpha_{43} = 10$	$\alpha_{44} = 12$	
			$\alpha_{54} = 1$	$\alpha_{55} = 10$

2.

$\alpha_{11} = 15$				
$\alpha_{21} = 1$	$\alpha_{22} = 14$			
	$\alpha_{32} = 11$	$\alpha_{33} = 13$		
		$\alpha_{43} = 1$	$\alpha_{44} = 12$	
			$\alpha_{54} = 9$	$\alpha_{55} = 10$

3.

$\alpha_{11} = 15$				
$\alpha_{21} = 12$	$\alpha_{22} = 14$			
	$\alpha_{32} = 1$	$\alpha_{33} = 13$		
		$\alpha_{43} = 1$	$\alpha_{44} = 12$	
			$\alpha_{54} = 10$	$\alpha_{55} = 11$

References

Bassok, Y., R. Anupindi, R. Akella. 1999. Single-period multiproduct inventory models with substitution. *Operations Research* 47(4): 632-642.

Brumelle, S.L., J.I. McGill. 1993. Airline seat allocation with multiple nested fare classes. *Operations Research* 41(1): 127-137.

Hoffman, A.J. 1963. On simple linear programming problems. V. Klee, ed., *Convexity: Proceedings of Symposia in Pure Math.* Vol. 7.

Netessine, S., G. Dobson, R.A. Shumsky. 2002. Flexible service capacity: optimal investment and the impact of demand correlation. *Operations Research* Vol. 50, No. 2, pgs. 375-388.

Netessine, S., N. Rudi. 2003. Centralized and competitive inventory models with substitution. *Operations Research* 51(2): 329-335.

Rockafellar, R. T. 1970. *Convex Analysis.* Princeton University Press, Princeton, New Jersey.

Shumsky, R.A., F. Zhang. 2007. Dynamic capacity management with substitution. Available at http://mba.tuck.dartmouth.edu/pages/faculty/robert.shumsky

Van Slyke R., R. Wets. 1966. Programming under uncertainty and stochastic optimal control. *SIAM Journal on Control* 4(1): 179-193.