Online Appendix

EC.1. Proofs of Lemmas and Theorems

Proof of Lemma 1. For any $i = 1, ..., m$, we need to show that $E[g(\Xi_1, ...\Xi_{i-1}, \Xi_{i+1}, ..., \Xi_m)|\Xi_i = s]$ is increasing in s for any increasing function g. Let $f(X_1, ..., X_n) = g(\Xi_1, ...\Xi_{i-1}, \Xi_{i+1}, ..., \Xi_m)$. If g is an increasing function, then f is also an increasing function since all entries in A are nonnegative. According to Theorem 1 in Efron (1965) we have that $E[f(X_1, ..., X_n)|a_{i1}X_1 + ... + a_{in}X_n = s]$ is increasing in s for any increasing function f. Then the desired result holds. □

Proof of Theorem 1. In the following we provide a proof of part (ii). Proof of part (i) follows similar but simpler arguments and is omitted. Let π^* be the optimal objective value of problem (4). Since for any $u \in \mathcal{F}^n$, $(u, (v(\xi), \xi \in \mathcal{X})) = (u, (u \wedge \xi, \xi \in \mathcal{X}))$ is feasible for problem (4), we have $\pi^* \leq \tau^*$.

It remains to show that $\tau^* \leq \pi^*$. When $\pi^* = \infty$ this clearly holds. As a result, we assume that $\pi^* < \infty$, which together with assumption (T1.a) implies that problems (2), (4), and all optimization problems below admit finite optimal solutions.

We first show that it is true for $n = 1$. Let $\hat{u} \in \arg\min_u f(u)$ and $u^* \in \arg\min_u \{l(u) + E[f(u \wedge \Xi)]\}$. Given any solution of problem (4), denoted by $(u', (v'(\xi), \xi \in \mathcal{X}))$, we are going to show that

$$l(u^*) + E[f(u^* \wedge \Xi)] \leq l(u') + E[f(v'(\Xi))].$$

For this purpose, we discuss two cases.

Case (1): if $u' \leq \hat{u}$, then

$$l(u^*) + E[f(u^* \wedge \Xi)] \leq l(u') + E[f(u' \wedge \Xi)] \leq l(u') + E[f(v'(\Xi))].$$

The first inequality is due to the optimality of u^*. In the following we prove the second inequality. If $\xi \leq u'$, then $f(u' \wedge \xi) = f(\xi) \leq f(v'(\xi))$ since $v'(\xi) \leq \xi \leq u' \leq \hat{u}$ and $f(\cdot)$ is convex. If $\xi > u'$, then $f(u' \wedge \xi) = f(u') \leq f(v'(\xi))$ since $v'(\xi) \leq u' \leq \hat{u}$ and $f(\cdot)$ is convex. Hence, $f(u' \wedge \xi) \leq f(v'(\xi))$ for any $\xi \in \mathcal{X}$.

Case (2): if $u' > \hat{u}$, then

$$l(u^*) + E[f(u^* \wedge \Xi)] \leq l(\hat{u}) + E[f(\hat{u} \wedge \Xi)] \leq l(u') + E[f(v'(\Xi))].$$

The first inequality is due to the optimality of u^*. Since $v'(\xi) \leq \xi$, it is easy to see that $f(\hat{u} \wedge \xi) \leq f(v'(\xi))$ for any $\xi \in \mathcal{X}$, and thus $E[f(\hat{u} \wedge \Xi)] \leq E[f(v'(\Xi))]$. Then the second inequality follows from the fact that $l(\cdot)$ is increasing.

Combing the above two cases we have $\tau^* \leq \pi^*$. This completes the proof when $n=1$.

We now consider the general case with $n > 1$. We start from the first component. Let $(u', v') = (u'_1, ..., u'_n, (v'_1(\xi_1), ..., v'_n(\xi_n), \xi \in \mathcal{X}))$ denote an optimal solution of problem (4). Given $(u'_2, ..., u'_n, (v'_2(\xi_2), ..., v'_n(\xi_n), \xi \in \mathcal{X}))$, π^* is equal to the optimal objective value of the following problem

$$\begin{aligned}
&\inf \; l(u_1, u'_2, ..., u'_n) + E[f(v_1(\Xi_1), v'_2(\Xi_2), \ldots, v'_n(\Xi_n))] \\
&\text{s.t. } v_1(\xi_1) \in \mathcal{F}, \\
&\quad\;\; v_1(\xi_1) \leq \xi_1 \;\; \forall \xi_1 \in \mathcal{X}_1, \\
&\quad\;\; v_1(\xi_1) \leq u_1 \;\; \forall \xi_1 \in \mathcal{X}_1, \\
&\quad\;\; v_1(\xi_1) \text{ is increasing } \forall \xi_1 \in \mathcal{X}_1.
\end{aligned} \tag{EC.1}$$

Our objective is to show that there exists a u_1^* such that

$$l(u_1^*, u'_2, ..., u'_n) + E[f(u_1^* \wedge \Xi_1, v'_2(\Xi_2), \ldots, v'_n(\Xi_n))] = \pi^*.$$

By the law of iterative expectation, we can rewrite the second term in the objective function of problem (EC.1) as

$$\begin{aligned}
&E_{\Xi_1}[E_{\Xi_2,\ldots,\Xi_n|\Xi_1}[f(v_1(\Xi_1), v'_2(\Xi_2), \ldots, v'_n(\Xi_n))]] \\
&= E_{\Xi_1}[g(v_1(\Xi_1), \Xi_1)],
\end{aligned}$$

where

$$g(v_1, \xi_1) = E_{\Xi_2,\ldots,\Xi_n|\Xi_1=\xi_1}[f(v_1, v'_2(\Xi_2), \ldots, v'_n(\Xi_n))].$$

Since f is componentwise convex, $g(v_1, \xi_1)$ is convex in v_1 for any ξ_1. For reasons which will be clear later, in the following we show that $g(v_1, \xi_1)$ is supermodular in (v_1, ξ_1). Define $w = (w_2, ...w_n)$ and

$$\tilde{f}(v_1, w) = f(v_1, v'_2(w_2), ..., v'_n(w_n)).$$

Then

$$g(v_1, \xi_1) = \int \tilde{f}(v_1, w) d\tilde{F}_{\xi_1}(w),$$

where $\tilde{F}_{\xi_1}(w)$ is the joint distribution of $\Xi_2, ..., \Xi_n$ conditional on $\Xi_1 = \xi_1$. By assumption (T1.c), we have that $\{\tilde{F}_{\xi_1}(w) | \xi_1 \in \mathcal{X}_1\}$ is a stochastically increasing collection of distribution functions. Since $v'_j(\xi_j)$ are increasing for $j = 2, ..., n$ and f is supermodular, $\tilde{f}$ is supermodular in (v_1, w) (See section 9.A.4 of Shaked and Shanthikumar 2006). It then follows from Theorem 3.10.1 of Topkis (1998) that $g(v_1, \xi_1)$ is supermodular. This implies that there exists an optimal solution $\hat{v}_1(\xi_1)$ of the unconstrained optimization problem $\inf_{v_1} g(v_1, \xi_1)$ which is decreasing in ξ_1 (see Theorem 2.2.8 of Simchi-Levi et al. 2014). Notice that $(\hat{v}_1(\xi_1) | \xi_1 \in \mathcal{X}_1)$ is an optimal solution of the unconstrained optimization problem $\inf E_{\Xi_1}[g(v_1(\Xi_1), \Xi_1)]$. And we will use $(\hat{v}_1(\xi_1) | \xi_1 \in \mathcal{X}_1)$ for the remainder of our proof.

Let $l_1(u_1)=l(u_1,u_2',...,u_n')$, which is clearly increasing. Now we can reformulate problem (EC.1) as follows:

$$\begin{aligned}
\inf\ & l_1(u_1)+E_{\Xi_1}[g(v_1(\Xi_1),\Xi_1)]\\
\text{s.t. } & v_1(\xi_1)\in\mathcal{F},\\
& v_1(\xi_1)\le \xi_1\ \forall\xi_1\in\mathcal{X}_1,\\
& v_1(\xi_1)\le u_1\ \forall\xi\in\mathcal{X}_1,\\
& v_1(\xi_1) \text{ is increasing } \forall\xi_1\in\mathcal{X}_1.
\end{aligned} \tag{EC.2}$$

Let $(u_1',(v_1'(\xi_1),\xi_1\in\mathcal{X}_1))$ be an optimal solution of problem (EC.2). We are going to show that there exists a u_1^* such that $l_1(u_1^*)+E[g(u_1^*\wedge\Xi_1,\Xi_1)]\le l_1(u_1')+E[g(v_1'(\Xi_1),\Xi_1)]$. We divide our discussion into the following three cases and construct u_1^* for each case respectively. These cases are divided based on the relative position of $(\hat v_1(\xi_1)|\xi_1\in\mathcal{X}_1)$ and $(v_1'(\xi_1),\xi_1\in\mathcal{X}_1)$. (1) When there exists some $\hat\xi_1$ such that $\hat v_1(\xi_1)\ge v_1'(\xi_1)$ for $\xi_1<\hat\xi_1$ and $\hat v_1(\xi_1)\le v_1'(\xi_1)$ for $\xi_1\ge\hat\xi_1$, choose $u_1^*=\sup_{\xi_1<\hat\xi_1}v_1'(\xi_1)$; (2) When $v_1'(\xi_1)\le\hat v_1(\xi_1)\ \forall\xi_1\in\mathcal{X}_1$, choose $u_1^*=\sup v_1'(\xi_1)$; (3) When $v_1'(\xi_1)\ge\hat v_1(\xi_1)\ \forall\xi_1\in\mathcal{X}_1$, let $u_1^*=\inf v_1'(\xi_1)$. Recall that $(\hat v_1(\xi_1)|\xi_1\in\mathcal{X}_1)$ is an optimal solution of the unconstrained optimization problem $\inf E_{\Xi_1}[g(v_1(\Xi_1),\Xi_1)]$, and $\hat v_1(\xi_1)$ is decreasing in $\xi_1\in\mathcal{X}_1$.

Case (1): If $\xi_1\ge\hat\xi_1$, then $u_1^*\wedge\xi_1=u_1^*$ since $u_1^*\le v_1'(\hat\xi_1)\le\hat\xi_1$. We also have $\hat v_1(\xi_1)\le u_1^*$ since $\hat v_1(\xi_1)$ is decreasing, and $v_1'(\xi_1)\ge u_1^*$ since $v_1'(\xi_1)$ is increasing. Since the function $g(v_1,\xi_1)$ is convex in v_1 for any ξ_1, and $\hat v_1(\xi_1)\le u_1^*\wedge\xi_1\le v_1'(\xi_1)$, we have $g(u_1^*\wedge\xi_1,\xi_1)\le g(v_1'(\xi_1),\xi_1)$. If $\xi_1\le\hat\xi_1$, we have $v_1'(\xi_1)\le u_1^*\wedge\xi_1$ since $v_1'(\xi_1)\le\xi_1$, $v_1'(\hat\xi_1)\le u_1^*$, and $v_1'(\xi_1)$ is increasing. Because $v_1'(\xi_1)\le u_1^*\wedge\xi_1\le\hat v_1(\xi_1)$ and $g(v_1,\xi_1)$ is convex in v_1 for any $\xi_1\in\mathcal{X}_1$, we obtain that $g(u_1^*\wedge\xi_1,\xi_1)\le g(v_1'(\xi_1),\xi_1)\ \forall\xi_1\in\mathcal{X}_1$. Since $u_1^*\le u_1'$ and $l_1(u_1)$ is increasing, we have $l_1(u_1^*)+E[g(u_1^*\wedge\Xi_1,\Xi_1)]\le l_1(u_1')+E[g(v_1'(\Xi_1),\Xi_1)]$.

Case (2): When $v_1'(\xi_1)\le\hat v_1(\xi_1)\ \forall\xi_1\in\mathcal{X}_1$, we have $v_1'(\xi_1)\le u_1^*\wedge\xi_1\le\hat v_1(\xi_1)\ \forall\xi_1\in\mathcal{X}_1$. Since $g(v_1,\xi_1)$ is convex in v_1 for any ξ_1, we have $g(u_1^*\wedge\xi_1,\xi_1)\le g(v_1'(\xi_1),\xi_1)\ \forall\xi_1\in\mathcal{X}_1$. Notice that in this case we must have $u_1'=u_1^*$, hence $l_1(u_1^*)+E[g(u_1^*\wedge\Xi_1,\Xi_1)]\le l_1(u_1')+E[g(v_1'(\Xi_1),\Xi_1)]$.

Case (3): When $v_1'(\xi_1)\ge\hat v_1(\xi_1)\ \forall\xi_1\in\mathcal{X}_1$, we have $u_1^*\wedge\xi_1=u_1^*$. Hence, $\hat v_1(\xi_1)\le u_1^*\wedge\xi_1\le v_1'(\xi_1)\ \forall\xi_1\in\mathcal{X}_1$ and $g(u_1^*\wedge\xi_1,\xi_1)\le g(v_1'(\xi_1),\xi_1)\ \forall\xi_1\in\mathcal{X}_1$. Since $u_1^*\le u_1'$, we obtain $l_1(u_1^*)+E[g(u_1^*\wedge\Xi_1,\Xi_1)]\le l_1(u_1')+E[g(v_1'(\Xi_1),\Xi_1)]$.

Therefore, we have shown that there exists a u_1^* such that

$$l(u_1^*,u_2',...,u_n')+E[f(u_1^*\wedge\Xi_1,v_2'(\Xi_2),\dots,v_n'(\Xi_n))]=\pi^*.$$

Next move on to the second component. Similarly to (EC.1), π^* is equal to the optimal objective value of the following problem

$$\begin{aligned}
\inf\ & l(u_1^*,u_2,u_3',...,u_n')+E[f(u_1^*\wedge\Xi_1,v_2(\Xi_2),v_3'(\Xi_3),\dots,v_n'(\Xi_n))]\\
\text{s.t. } & v_2(\xi_2)\in\mathcal{F},\\
& v_2(\xi_2)\le\xi_2\ \forall\xi_2\in\mathcal{X}_2,\\
& v_2(\xi_2)\le u_2\ \forall\xi_2\in\mathcal{X}_2,\\
& v_2(\xi_2) \text{ is increasing } \forall\xi_2\in\mathcal{X}_2.
\end{aligned} \tag{EC.3}$$

Following the proceeding analysis, there exists a u_2^* such that

$$\pi^* = l(u_1^*, u_2^*, u_3', ..., u_n') + E[f(u_1^* \wedge \Xi_1, u_2^* \wedge \Xi_2, v_3'(\Xi_3), ..., v_n'(\Xi_n))].$$

Continue this process while applying the same approach we can find $u_j^*, j = 1, ..., n$ such that $\pi^* = l(u_1^*, ..., u_n^*) + E[f(u_1^* \wedge \Xi_1, ..., u_n^* \wedge \Xi_n)]$. Therefore, we have $\pi^* = \tau^*$. This completes the proof of part (ii).

For part (i), the proof for $n = 1$ is the same as that from part (ii). When $n > 1$, since Ξ has independent components, we can follow the same arguments as the proof of Theorem 1 in Chen et al. (2018) to obtain the desired results. □

Proof of Lemma 2. The indicator function of a set $\mathcal{V} \subseteq \mathcal{F}^n$, denoted by $\delta_{\mathcal{V}}$, is defined as $\delta_{\mathcal{V}}(x) = 0$ for $x \in \mathcal{V}$ and $+\infty$ otherwise. The indicator function is a supermodular function if for any $x, y \in \mathcal{F}^n$, we have

$$\delta_{\mathcal{V}}(x) + \delta_{\mathcal{V}}(y) \le \delta_{\mathcal{V}}(x \wedge y) + \delta_{\mathcal{V}}(x \vee y). \tag{EC.4}$$

We prove by contradiction. Suppose that (EC.4) does not hold for some x and y. Then we must have $\delta_{\mathcal{V}}(x \wedge y) = \delta_{\mathcal{V}}(x \vee y) = 0$ while at least one of $\delta_{\mathcal{V}}(x)$ or $\delta_{\mathcal{V}}(y)$ is equal to $+\infty$. Without loss of generality suppose $\delta_{\mathcal{V}}(x) = +\infty$. Hence, there exists some i such that $\psi_i(x) > 0$. If $\psi_i(\cdot)$ is increasing, then $\psi_i(x \vee y) > 0$, and thus $\delta_{\mathcal{V}}(x \vee y) = +\infty$; If $\psi_i(\cdot)$ is decreasing, then $\psi_i(x \wedge y) > 0$, and thus $\delta_{\mathcal{V}}(x \wedge y) = +\infty$. For either case there is a contradiction. □

Proof of Theorem 2. In the following we provide the proof of part (ii). The proof of part (i) is similar and therefore omitted.

Problem (5) is equivalent to the following unconstrained optimization problem.

$$\inf_{u \in \mathcal{F}^n} l(u) + E[f(u \wedge \Xi)] + \delta_{\mathcal{U}}(u), \tag{EC.5}$$

where $\delta_{\mathcal{U}}(u)$ is an indicator function such that $\delta_{\mathcal{U}}(u) = 0$ for $u \in \mathcal{U}$ and $\delta_{\mathcal{U}}(u) = +\infty$ otherwise. For any $v \in \mathcal{F}^n$, let $\hat{f}(v) = f(v) + \delta_{\mathcal{V}}(v)$, where $\mathcal{V}$ is defined in (6). Since the indicator function of $\mathcal{V}$ is supermodular, together with the supermodularity of f we obtain that $\hat{f}$ is supermodular. Then the optimal objective value of problem (EC.5) is equivalent to that of the following problem

$$\inf_{u \in \mathcal{F}^n} l(u) + E[\hat{f}(u \wedge \Xi)]. \tag{EC.6}$$

To see this, note that for any $u \in \mathcal{U}$, we have $u \wedge \xi \in \mathcal{V}\ \forall \xi \in \mathcal{X}$. Hence, given any feasible solution u of $\inf_{u \in \mathcal{U}}\{l(u) + E[f(u \wedge \Xi)]\}$, we can always find a $\hat{u} \in \mathcal{F}^n$ (simply let $\hat{u} = u$) such that $l(\hat{u}) + E[\hat{f}(\hat{u} \wedge \Xi)] = l(u) + E[f(u \wedge \Xi)]$. Therefore, $\inf_{u \in \mathcal{F}^n}\{l(u) + E[\hat{f}(u \wedge \Xi)]\} \le \inf_{u \in \mathcal{U}}\{l(u) + E[f(u \wedge \Xi)]\}$. On the other hand, due to Assumption 1 (A1.a), given any feasible solution $\hat{u}$ of $\inf_{u \in \mathcal{F}^n}\{l(u) + E[\hat{f}(u \wedge \Xi)]\}$, we can always find a $u \in \mathcal{U}$ such that $l(\hat{u}) + E[\hat{f}(\hat{u} \wedge \xi)] \ge l(u) + E[f(u \wedge \Xi)]$. Therefore, $\inf_{u \in \mathcal{F}^n}\{l(u) + E[\hat{f}(u \wedge \Xi)]\} \ge \inf_{u \in \mathcal{U}}\{l(u) + E[f(u \wedge \Xi)]\}$. Applying Theorem 1 part (ii) to problem (EC.6), we can obtain the transformed problem (8). □

Proof of Lemma 3. Recall that we define $\bar{\xi}_j = \text{ess}\sup\{\xi_j | \xi \in \mathcal{X}\}$. We first assume that $\bar{\xi}_j < \infty$ for all j. Similarly to the proof in lemma 2 of Chen et al. (2018), we demonstrate that $\mathcal{V}$ is equivalent to the following set $\{w | Aw \le b, \underline{u}_j \le w_j \le \bar{\xi}_j, j = 1, ..., n\}$, denoted by $\mathcal{V}_w$. For any $w = u \wedge \xi \in \mathcal{V}$, we have $Aw = A(u \wedge \xi) \le b$ since $a_{ij} \ge 0$ for all i, j and $Au \le b$; $\underline{u}_j \le u_j \wedge \xi_j = w_j \le \bar{\xi}_j$ since $\underline{u}_j \le \xi_j$ $\forall \xi_j \in \mathcal{X}_j$. For any $w \in \mathcal{V}_w$, let $u = w$, $\xi_j = \bar{\xi}_j$ for all j. Then $w = u \wedge \xi$ since $w_j \le \bar{\xi}_j$ for all j, and $u \in \mathcal{U}$. Hence, $\mathcal{V} = \mathcal{V}_w$. Clearly $\mathcal{V}$ is a convex set. Since all entries of A are nonnegative, the indicator function of $\mathcal{V}$ is supermodular due to Lemma 2. Given any u satisfying $u \wedge \xi \in \mathcal{V}$ $\forall \xi \in \mathcal{X}$, we define $u' = u \wedge \bar{\xi}$, One can easily check that $u' \le u$ and $u' \wedge \xi = u \wedge \xi$ $\forall \xi \in \mathcal{X}$. We are left to show that $u' \in \mathcal{U}$. Clearly $u' \ge \underline{u}$, hence we are left to show that $Au' \le b$.

For any row i of A, if there is only one positive entry, denoted by j, then obviously $a_{ij} u'_j \le b_i$ since $A(u \wedge \xi) \le b$ $\forall \xi \in \mathcal{X}$. If there are more than one positive entries, then we show that $\sum_{j=1}^n a_{ij} u'_j \le b_i$ by contradiction. Suppose that $\sum_{j=1}^n a_{ij} u'_j > b_i$, we have $\sum_{j=1}^n a_{ij}(u_j \wedge \bar{\xi}_j) > b_i$. Since $A(u \wedge \xi) \le b$ $\forall \xi \in \mathcal{X}$, there exists $\epsilon > 0$, $k, l \in \{1, ..., n\}$ such that $Pr(\Xi_k > \bar{\xi}_k - \epsilon, \Xi_l > \bar{\xi}_l - \epsilon) = 0$, $Pr(\Xi_k > \bar{\xi}_k - \epsilon) > 0$, $Pr(\Xi_l > \bar{\xi}_l - \epsilon) > 0$. Hence, $Pr(\Xi_k > \bar{\xi}_k - \epsilon | \Xi_j > \bar{\xi}_j - \epsilon) = 0$. Since the components of Ξ are positively dependent, we have $Pr(\Xi_k > \bar{\xi}_k - \epsilon | \Xi_l > \xi_l)$ is increasing in ξ_l. Thus, $Pr(\Xi_k > \bar{\xi}_k - \epsilon | \Xi_l > \xi_l) = 0$ $\forall \xi_l \in \mathcal{X}_l$. This means that $Pr(\Xi_k > \bar{\xi}_k - \epsilon) = 0$, which is a contradiction.

If $\bar{\xi}_j = \infty$ for any j, then $u'_j = u_j$ and following similar arguments we can obtain the desired results. $\square$

Proof of Theorem 3. The proof is similar to that of Theorem 3 in Chen et al. (2018). In the following we provide the proof of part (ii).

According to Theorem 2, problem (9) can be converted to the following problem:

$$
\begin{aligned}
\inf\ & l(x, z, u) + E[f(x, v_1(\Xi_1), v_2(\Xi_2))] \\
\text{s.t. } & v_1(\xi_1) \le z_1 + \xi_1 \quad \forall \xi_1 \in \mathcal{X}_1, \\
& v_1(\xi_1) \le u_1 \quad \forall \xi_1 \in \mathcal{X}_1, \\
& v_2(\xi_2) \le z_2 + \xi_2 \quad \forall \xi_2 \in \mathcal{X}_2, \\
& v_2(\xi_2) \le u_2 \quad \forall \xi_2 \in \mathcal{X}_2, \\
& v_1(\xi_1) \text{ is increasing } \forall \xi_1 \in \mathcal{X}_1, \\
& v_2(\xi_2) \text{ is increasing } \forall \xi_2 \in \mathcal{X}_2, \\
& (x, z_1, z_2, v_1(\xi_1), v_2(\xi_2)) \in \mathcal{A}^{\Xi} \quad \forall \xi \in \mathcal{X}.
\end{aligned}
\tag{EC.7}
$$

If functions l and f are convex, the objective function is a convex function. The constraint set is a convex set when $\mathcal{A}^{\Xi}$ is convex. It follows from Proposition 2.1.15 of Simchi-Levi et al. (2014) that g is also convex.

Define $(\tilde{v}_1, \tilde{v}_2) = (v_1, -v_2)$. Then problem (EC.7) is equivalent to

$$\begin{aligned}
\inf\ & \tilde{l}(z, \tilde{z}, \tilde{u}) + E[\tilde{f}(x, \tilde{v}_1(\tilde{\Xi}_1), \tilde{v}_2(\tilde{\Xi}_2))] \\
\text{s.t. } & \tilde{v}_1(\tilde{\xi}_1) \le \tilde{z}_1 + \tilde{\xi}_1 \quad \forall \tilde{\xi}_1 \in \tilde{\mathcal{X}}_1, \\
& \tilde{v}_1(\tilde{\xi}_1) \le \tilde{u}_1 \quad \forall \tilde{\xi}_1 \in \tilde{\mathcal{X}}_1, \\
& \tilde{v}_2(\tilde{\xi}_2) \ge \tilde{z}_2 + \tilde{\xi}_2 \quad \forall \tilde{\xi}_2 \in \tilde{\mathcal{X}}_2, \\
& \tilde{v}_2(\tilde{\xi}_2) \ge \tilde{u}_2 \quad \forall \tilde{\xi}_2 \in \tilde{\mathcal{X}}_2, \\
& \tilde{v}_1(\tilde{\xi}_1) \text{ is increasing } \forall \tilde{\xi}_1 \in \tilde{\mathcal{X}}_1, \\
& \tilde{v}_2(\tilde{\xi}_2) \text{ is decreasing } \forall \tilde{\xi}_2 \in \tilde{\mathcal{X}}_2, \\
& (x, \tilde{z}_1, \tilde{z}_2, \tilde{v}_1(\tilde{\xi}_1), \tilde{v}_2(\tilde{\xi}_2)) \in \tilde{\mathcal{A}}^{\Xi} \quad \forall \tilde{\xi} \in \tilde{\mathcal{X}}.
\end{aligned} \tag{EC.8}$$

Clearly the optimal objective function of (EC.8) is $\tilde{g}(x, \tilde{z})$. Since $\tilde{f}$ and l are submodular, the objective function of (EC.8) is submodular. The constraint set is a lattice. It then follows from Theorem 2.7.6 of Topkis (1998) that $\tilde{g}$ is submodular. □

Proof of Theorem 5. (i) Based on the definition of the distortion risk measure, problem (12) can be formulated as

$$\inf_{u \in \mathcal{U}} \int_0^1 \inf_{\hat{\lambda}(\alpha) \in \Re,\ \forall \alpha \in [0,1)} [\hat{\lambda}(a) + \frac{1}{1-\alpha} E[l(u) + f(u \wedge \Xi) - \hat{\lambda}]] d\mu(\alpha),$$

which is equivalent to

$$\inf_{\hat{\lambda}(\alpha) \in \Re, \forall \alpha \in [0,1), u \in \mathcal{U}} \int_0^1 \hat{\lambda}(a) + \frac{1}{1-\alpha} E[l(u) + f(u \wedge \Xi) - \hat{\lambda}] d\mu(\alpha).$$

Define $\lambda(a) = \hat{\lambda}(a) - l(u)\ \forall a \in [0,1)$ and $g(u, \lambda, \alpha) = \lambda + \frac{1}{1-\alpha}(f(u) - \lambda)^+$, the above problem can be reformulated as

$$\inf_{\lambda(a) \in \Re,\ \forall \alpha \in [0,1), u \in \mathcal{U}} l(u) + E\left[\int_0^1 g(u \wedge \Xi, \lambda(a), \alpha) d\mu(\alpha)\right].$$

We have that $\int_0^1 g(u, \lambda(\alpha), \alpha) d\mu(\alpha)$ is componentwise convex in u since f is componentwise convex. By Theorem 2 part (i), given any $\lambda(\alpha)$, $\alpha \in [0,1)$, the problem $\inf_{u \in \mathcal{U}} l(u) + E\left[\int_0^1 g(u \wedge \Xi, \lambda(a), \alpha) d\mu(\alpha)\right]$ can be transformed to

$$\begin{aligned}
\inf\ & l(u) + E[\textstyle\int_0^1 g(v(\Xi), \lambda(\alpha), \alpha) d\mu(\alpha)] \\
\text{s.t. } & v(\xi) = (v_1(\xi_1), \ldots, v_n(\xi_n)) \in \mathcal{V}\ \forall \xi \in \mathcal{X}, \\
& v(\xi) \le \xi\ \forall \xi \in \mathcal{X}, \\
& v(\xi) \le u\ \forall \xi \in \mathcal{X}.
\end{aligned}$$

Therefore, problem (12) is equivalent to problem (14).

(ii) When Ξ has positively dependent components, according to Theorem 2 part (ii) we need to show that $g(u, \lambda, \alpha)$ is supermodular in u for any fixed $\lambda, \alpha \in [0,1)$. We can rewrite $g(u, \lambda, \alpha)$ as a composition function $h(f(u))$ where $h(x) = \lambda + \frac{1}{1-\alpha}(x - \lambda)^+$. Note that for notational simplicity, we drop the dependence of λ and α in the definition of h. Since $h(x)$ is not continuously differentiable, we smoothen h by $h_\epsilon(x) = E[h(x-\epsilon Z)]$, where Z follows the standard gaussian distribution. We then need to show that $\lim_{\epsilon\to 0} h_\epsilon(f(u))$ is supermodular. We first assume that f is twice continuously differentiable. Then we have

$$\frac{\partial^2 h_\epsilon(f(u))}{\partial u_i \partial u_j} = \frac{\partial[h'_\epsilon(f(u))\frac{\partial f}{\partial u_i}]}{\partial u_j} = h''_\epsilon(f(u))\frac{\partial f}{\partial u_i}\frac{\partial f}{\partial u_j} + h'_\epsilon(f(u))\frac{\partial^2 f}{\partial u_i \partial u_j}.$$

Since $h(\cdot)$ is increasing and convex, we have that $h_\epsilon(\cdot)$ is also increasing and convex (see Proposition 2.1.3(e) of Simchi-Levi et al. 2014), which means that $h''_\epsilon(f(u)) \geq 0, h'_\epsilon(f(u)) \geq 0$. Since f is supermodular, we have $\frac{\partial^2 f}{\partial u_i \partial u_j} \geq 0$, which, together with the assumption $\frac{\partial f}{\partial u_i}\frac{\partial f}{\partial u_j} \geq 0 \ \forall u \in \mathcal{U}$, gives the desired results.

If f is not twice continuously differentiable, then we can also smoothen f by $f_{\hat\epsilon} = E[f(u - \hat\epsilon \hat Z)]$, where $\hat Z = (\hat Z_1, ..., \hat Z_n)$, and $\hat Z_1, ..., \hat Z_n$ are independent and follow the standard normal distribution. Then we need that $\lim_{\epsilon,\hat\epsilon\to 0} h_\epsilon(f_{\hat\epsilon}(u))$ is supermodular. Similarly to the previous arguments we have that

$$\lim_{\epsilon,\hat\epsilon\to 0} h''_\epsilon(f_{\hat\epsilon}(u))\frac{\partial f_{\hat\epsilon}}{\partial u_i}\frac{\partial f_{\hat\epsilon}}{\partial u_j} + h'_\epsilon(f_{\hat\epsilon}(u))\frac{\partial^2 f_{\hat\epsilon}}{\partial u_i \partial u_j} \geq 0.$$

Therefore, the desired results can be obtained. □

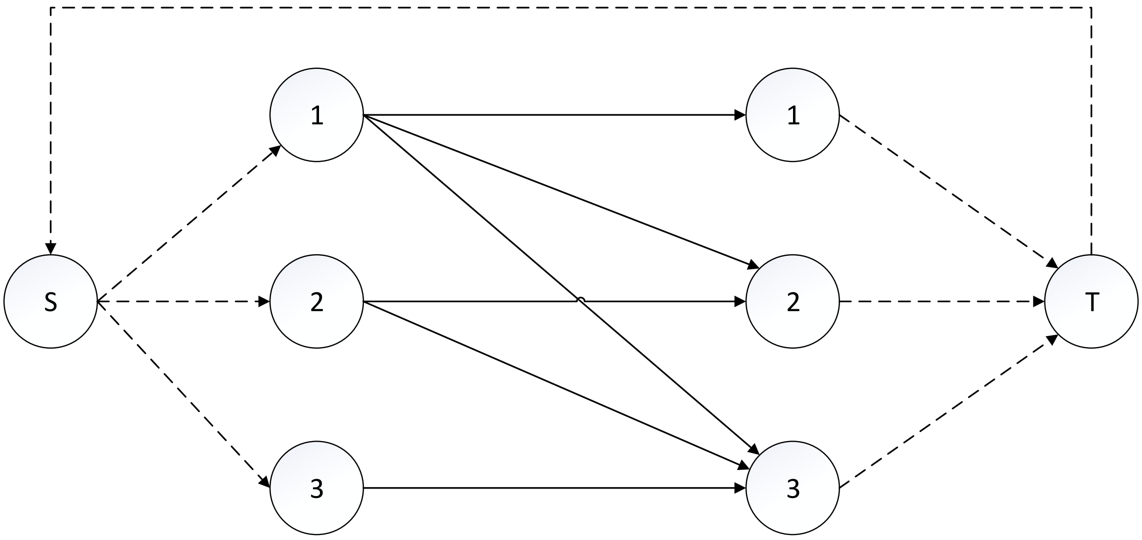
Proof of Theorem 6. We first show that the function $L(\cdot|d)$ is convex and supermodular for any d. For this purpose, we construct the following maximum weight circulation problem:

$$\begin{aligned} F(y|d) = \max\ & -\textstyle\sum_{i=1}^N \sum_{j=1}^N s_{ij} w_{ij} + \sum_{i=1}^N h_i f_{Si} + \sum_{j=1}^N p_j f_{jT} \\ \text{s.t.}\ & \textstyle\sum_{j=i}^N w_{ij} = f_{Si},\ \forall i = 1, ..., N, \\ & \textstyle\sum_{i=1}^j w_{ij} = f_{jT},\ \forall j = 1, ..., N, \\ & 0 \leq f_{Si} \leq y_i, 0 \leq f_{jT} \leq d_j, w_{ij} \geq 0,\ \forall i = 1, ..., N, j = 1, ..., N. \end{aligned} \tag{EC.9}$$

In the above maximum weight circulation problem, we add two more nodes S and T to the original product-demand graph (see Figure EC.1 for an example with $n = 3$). Let w_{ij} be the flow from product i to demand j, f_{Si} be the flow from node S to product i, and f_{jT} be the flow from demand j to node T. Flows f_{Si} and f_{jT} are bounded above by y_i and d_j respectively. The weights for w_{ij}, f_{Si}, f_{jT} are $-s_{ij}, h_i, p_j$ respectively. Since the arcs associated with flow f_{Si} are parallel, by Theorem 1.3 and 1.4 of Murota (2005), F is concave and submodular. Notice that $L(y|d) = -F(y|d) + \sum_{i=1}^N h_i y_i + \sum_{j=1}^N p_j d_j$, which is convex and supermodular in y. Therefore, the function $L(\cdot|d)$ is convex and supermodular for any d.

The constraint set of problem (17) is $y \geq x$, and clearly $x + K_j \geq x \ \forall K \in \mathcal{K}$ since the support $\mathcal{K}$ is nonnegative. According to Lemma 3 we have that Assumption 1 is satisfied. Since $L(\cdot|d)$ is convex and supermodular for any d, we have that the function g is convex and supermodular (see Proposition 2.1.2(e) and Proposition 2.2.5(d) of Simchi-Levi et al. 2014). Therefore, after applying Theorem 2, the original problem can be transformed to (18). □

Figure EC.1 The equivalent maximum weight circulation problem



EC.2. Numerical Studies

Let $\tilde{c}_i = c_{e,i} + h_i^+, \ \forall i = 1, ..., N, \tilde{s}_{ij} = s_{ij} - h_i^+ - p_j, \ \forall i = 1, ..., N, j = i, ..., N$. Problem (18) can be reformulated as the following three-stage stochastic program:

$$\begin{aligned}
\min \ & \{\textstyle\sum_{i=1}^{n} c_{o,i} y_i + E_K[\min \sum_{i=1}^{N} \tilde{c}_i v_i(K_i) + E_D[\min \sum_{i=1}^{N}\sum_{j=i}^{N} \tilde{s}_{ij} w_{ij}(K, D)]]\} \\
\text{s.t.} \ & v(k) = (v_1(k_1), ..., v_n(k_n)) \ \forall k \in \mathcal{K}, \\
& \textstyle\sum_{j=i}^{N} w_{ij}(k, d) \le v_i(k), \ \forall i = 1, ..., N, \forall k \in \mathcal{K}, d \in \mathcal{D}, \\
& \textstyle\sum_{i=1}^{j} w_{ij}(k, d) \le d_j, \ \forall j = 1, ..., N, \forall k \in \mathcal{K}, d \in \mathcal{D}, \\
& w_{ij}(k, d) \ge 0, \ \forall i = 1, ..., n, j = 1, ..., N, \forall k \in \mathcal{K}, d \in \mathcal{D}, \\
& x_i \le v_i(k_i) \le x_i + k_i, \ \forall i = 1, ..., N, \forall k_i \in \mathcal{K}_i, \\
& v_i(k_i) \le y_i, \ \forall i = 1, ..., N, \ \forall k_i \in \mathcal{K}_i, \\
& v_i(k_i) \text{ is increasing}, \ \forall i = 1, ..., N.
\end{aligned} \tag{EC.10}$$

Many approximation methods can be applied to solve the above stochastic programming problem, such as the scenario approximation using Monte Carlo sampling or the piecewise linear decision rule approximation proposed in Georghiou et al. (2011).

In the following we conduct numerical experiments to study the price of ignoring the capacity dependence. In particular, we compare the optimal total cost when the firm correctly considers dependent capacities with that obtained by incorrectly assuming that capacities are independent. The number of products is chosen as $N = 3$ or $N = 5$. We let $c_{o,i} = 0, c_{e,i} = 1 + 0.5(N - i)$. Holding costs and substitution costs are equal to 0. The penalty costs are chosen as $p_i = p_0 c_i$, where $p_0 = 2, 5$, or 10.

Table EC.1 Percentage Cost Gap (in %), $N = 3$

		CV		
	p	0.3	0.5	1
Mean of capacity=100	2	3.7	4.9	8.7
	5	4.5	7.8	6.8
	10	4.2	4.3	5.9
Mean of capacity=150	2	<0.1	5.8	25.3
	5	2.6	19.9	18.3
	10	12.4	18.8	17.0

Table EC.2 Percentage Cost Gap (in %), $N = 5$

		CV		
	p	0.3	0.5	1
Mean of capacity=100	2	4.5	6.3	9.3
	5	5.3	7.9	8.4
	10	4.7	5.8	7.3
Mean of capacity=150	2	<0.1	4.5	23.9
	5	2.1	17.9	20.2
	10	17.4	16.7	14.9

The demand of each product is 100. The capacities of all products are perfectly positively correlated, so that product i's capacity is $K_i = Z$, where Z is a truncated Gaussian variable. The coefficient of correlation (CV) can be $0.3, 0.5$ or 1. The support of Z is either $[0, 200]$ (mean of Z is 100) or $[0, 300]$ (mean of Z is 150). The approximate optimal cost $\mathcal{C}^*$ is obtained by solving (EC.10) with 2000 scenarios.

When the firm ignores the capacity dependence, she incorrectly assumes that products' capacities are independent Gaussian variables. We generate another 2000 scenarios with this wrong assumption and solve (EC.10) again to obtain the approximate optimal ordering quantities. We then use simulation to calculate the resulting cost $\mathcal{C}'$ with the true random capacities. We define the percentage cost gap as $(\frac{\mathcal{C}'}{\mathcal{C}^*} - 1) \times 100\%$. Table 1 and Table 2 summarize the percentage cost gaps obtained from our numerical experiments.

It is clear that ignoring capacity dependence can be very harmful to the firm. Table 1 and 2 demonstrate that the percentage cost gap is on average 9.6% for the cases we tested, and the worst case is over 25%. Higher variability in capacity in general leads to higher percentage cost gap. In our numerical experiments, we find out that when she ignores the capacity dependence, the firm often orders too many products. This is because the firm overestimates the chance of substitution when she incorrectly assume that the random capacities are independent. This overestimation is most severe when the supply capacity is on average higher than the demand and has a large variability.

References

Chen, Xin, Xiangyu Gao, Zhan Pang. 2018. Preservation of structural properties in optimization with decisions truncated by random variables and its applications. *Operations Research* **66**(2) 340–357.

Efron, Bradley. 1965. Inreasing properties of Polya frequency function. *The Annals of Mathematical Statistics* **36**(1) 272–279.

Georghiou, Angelos, Wolfram Wiesemann, Daniel Kuhn. 2011. The decision rule approach to optimisation under uncertainty: Methodology and applications in operations management .

Henig, Mordechai, Yigal Gerchak. 1990. The structure of periodic review policies in the presence of random yield. *Operations Research* **38**(4) 634–643.

Murota, K. 2005. Note on multimodularity and l-convexity. *Math. Oper. Res* **30**(3) 658–661.

Shaked, M., G. Shanthikumar. 2006. *Stochastic Orders*. Springer-Verlag, New York.

Shumsky, Robert A, Fuqiang Zhang. 2009. Dynamic capacity management with substitution. *Operations research* **57**(3) 671–684.

Simchi-Levi, David, Xin Chen, Julien Bramel. 2014. *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics Management*. 3rd ed. Springer-Verlag, New York.

Topkis, D. M. 1998. *Supermodularity and Complementarity*. Princeton University Press, Princeton, NJ.