Lemma 1. For a multi-echelon supply chain with full logistics flexibility only at stage \( k_f \), let \( Y_{ijn}^{k_f} \) be production flow from plant \( i \) of stage \( K \) to demand node \( n \) of stage \( 0 \), and let \( \mathbf{m}_L(k_f) = (m_{L,in}(k_f)) \) be the matrix where element \( m_{L,in}(k_f) \) denotes the element of the \( i \)th row and \( n \)th column, problem \( P_2(\mathbf{A}_{1full}(k_f), \mathbf{B}, \mathbf{d}, \mathbf{q}) \) can be simplified to the following linear program, \( P_2(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{q}) \):

\[
\pi(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{q}) = \max_{Y_{ijn}} \left\{ \sum_{i=1}^{I} \sum_{n=1}^{N} m_{L,in}(k_f) Y_{ijn}^{k_f} \right\}
\]

subject to:

\[
\sum_{n=1}^{N} Y_{ijn}^{k_f} \leq \min_{1 \leq k \leq N} \{q_i^k\} \quad i = 1, 2, \ldots, I,
\]

\[
\sum_{i=1}^{I} Y_{ijn}^{k_f} \leq \min_{1 \leq k \leq N} \{q_i^k, d_n\} \quad n = 1, 2, \ldots, I.
\]

**Proof of Lemma 1:**

Since a flow \( X_{ijn}^k \) can be non-zero only if \((i,j) \in A^k, (i,n) \in B^k, \) and \((j,n) \in B^{k-1} \), for \( \mathbf{A}_{1full}(k_f) \), we have

\[
X_{ijn}^k \neq 0 \quad \text{only if} \quad i = j = n, \quad i, j, n = 1, 2, \ldots, I, \quad k = 1, 2, \ldots, k_f - 1,
\]

\[
X_{ijn}^k \neq 0 \quad \text{only if} \quad j = n, \quad i, j, n = 1, 2, \ldots, I, \quad k = k_f,
\]

\[
X_{ijn}^k \neq 0 \quad \text{only if} \quad i = j, \quad i, j, n = 1, 2, \ldots, I, \quad k = k_f + 1, k_f + 2, \ldots, K.
\]

Therefore constraint (2) of \( P_2(\mathbf{A}, \mathbf{B}, \mathbf{d}, \mathbf{q}) \) becomes

\[
X_{iin}^k = \cdots = X_{iin}^{k_f - 1} = \sum_{i=1}^{I} X_{iin}^{k_f} \quad n = 1, 2, \ldots, I,
\]

\[
X_{iin}^{k_f} = X_{iin}^{k_f + 1} = \cdots = X_{iin}^{K} \quad i, n = 1, 2, \ldots, I.
\]

And constraint (3) of \( P_2(\mathbf{A}, \mathbf{B}, \mathbf{d}, \mathbf{q}) \) becomes

\[
X_{nii}^k \leq q_i^k \quad n = 1, 2, \ldots, I, \quad \forall \quad 1 \leq k < k_f,
\]

\[
\sum_{n=1}^{N} X_{nii}^{k_f} \leq q_i^{k_f} \quad i = 1, 2, \ldots, I,
\]

\[
\sum_{i=1}^{I} X_{nii}^{k} \leq q_i^k \quad i = 1, 2, \ldots, I, \quad \forall \quad k_f < k \leq K.
\]

From (13) we have \( X_{nii}^k = \sum_{i=1}^{I} X_{iin}^{k_f}, \) \( n = 1, 2, \ldots, I, \quad \forall \quad 1 \leq k < k_f. \) Substituting it into (15) we get

\[
\sum_{i=1}^{I} X_{iin}^{k_f} \leq q_i^k \quad n = 1, 2, \ldots, I, \quad 1 \leq k < k_f.
\]

From (14) we have \( X_{iin}^{k_f} = X_{iin}^{k_f}, \) \( i, j = 1, 2, \ldots, I, \quad \forall \quad k_f < k \leq K. \) Substituting it into (17) we get

\[
\sum_{n=1}^{N} X_{iin}^{k_f} \leq q_i^k \quad i = 1, 2, \ldots, I, \quad k_f < k \leq K.
\]
Combining (19) with (16) we have
\[
\sum_{n=1}^{I} X_{inn}^{k_f} \leq q_i^k \quad i = 1, 2, \ldots, I, \quad k_f \leq k \leq K. \tag{20}
\]
If \( k_f \neq 1 \), constraint (4) of P2(-) becomes
\[
X_{inn}^1 \leq d_n \quad n = 1, 2, \ldots, I,
\]
From (13) we have \( X_{inn}^k = \sum_{i=1}^{I} X_{inn}^{k_f} \), \( n = 1, 2, \ldots, I, \) \( \forall 1 \leq k < k_f \), which by substituting into the above inequality we get
\[
\sum_{i=1}^{I} X_{inn}^{k_f} \leq d_n \quad n = 1, 2, \ldots, I. \tag{21}
\]
And if \( k_f = 1 \), constraint (4) is exactly same as (21). On the other hand, (20) is equivalent to
\[
\sum_{i=1}^{I} X_{inn}^{k_f} \leq \min_{1 \leq k < k_f - 1} \{ q_i^k, d_n \} \quad i = 1, 2, \ldots, I. \tag{22}
\]
And combining (18) and (21) we have
\[
\sum_{i=1}^{I} X_{inn}^{k_f} \leq \min_{1 \leq k < k_f - 1} \{ q_i^k, d_n \} \quad n = 1, 2, \ldots, I. \tag{23}
\]
We define \( Y_{inn}^{k_f} \) as production flow from plant \( i \) at stage \( K \) to demand node \( n \) at stage 0. Since we have logistics flexibility only at stage \( k_f \), the only possible path for flow \( Y_{inn}^{k_f} \) is from plant \( i \) to plant \( i \) at stages \( K, K-1, \ldots, k_f \), then from plant \( i \) at stage \( k_f \) to plant \( n \) at stage \( k_f - 1 \), and then from plant \( n \) to plant \( n \) at stages \( k_f - 2, k_f - 3, \ldots, 1 \). Therefore, flow \( Y_{inn}^{k_f} \) is exactly the same as the flow of product \( n \) that flexible plant \( i \) at stage \( k_f \) sends to plant \( n \) at stage \( k_f - 1 \), i.e., \( Y_{inn}^{k_f} = X_{inn}^{k_f} \). Substituting \( X_{inn}^{k_f} \) with \( Y_{inn}^{k_f} \) in (22) and (23) we get (8) and (9) for \( \text{P2}(\cdot) \).

Now we derive the objective function of \( \text{P2}(\cdot) \). Taking into account (10), (11) and (12), if \( k_f \neq 1 \), objective function (1) can be written as
\[
\sum_{i=1}^{I} r_n X_{inn}^1 - \sum_{k=1}^{k_f-1} \sum_{i=1}^{I} \left( p_n^k + t_{inn}^k \right) X_{inn}^k - \sum_{i=1}^{I} \sum_{n=1}^{2} \left( p^{k_f}_i + t^{k_f}_{inn} \right) X_{inn}^{k_f} - \sum_{k=k_f+1}^{K} \sum_{i=1}^{I} \sum_{n=1}^{I} \left( p^{k}_n + t^{k}_{inn} \right) X_{inn}^{k_f}.
\]
In the above, by substituting \( Y_{inn}^{k_f} = Y_{inn}^k \), \( X_{inn}^k = \sum_{i=1}^{I} X_{inn}^{k_f} = \sum_{i=1}^{I} Y_{inn}^{k_f} \), where \( n = 1, 2, \ldots, I, \) and \( 1 \leq k < k_f \), and also \( X_{inn}^k = Y_{inn}^k \), where \( i, n = 1, 2, \ldots, I, \) and \( k_f < k \leq K \), we will have:
\[
\sum_{n=1}^{I} \left[ r_n \sum_{i=1}^{I} Y_{inn}^k \right] - \sum_{k=1}^{k_f-1} \sum_{i=1}^{I} \left[ \left( p_n^k + t_{inn}^k \right) \sum_{i=1}^{I} Y_{inn}^k \right] - \sum_{i=1}^{I} \sum_{n=1}^{I} \left( p^{k_f}_i + t^{k_f}_{inn} \right) Y_{inn}^{k_f} - \sum_{k=k_f+1}^{K} \sum_{i=1}^{I} \sum_{n=1}^{I} \left( p^{k}_n + t^{k}_{inn} \right) Y_{inn}^{k_f} \tag{24}
\]
\[
= \sum_{i=1}^{I} \sum_{n=1}^{I} r_n Y_{inn}^k - \sum_{i=1}^{I} \sum_{n=1}^{I} \left[ \left( \sum_{k=1}^{k_f-1} \left( p_n^k + t_{inn}^k \right) \right) Y_{inn}^k \right] - \sum_{i=1}^{I} \sum_{n=1}^{I} \left( p^{k_f}_i + t^{k_f}_{inn} \right) Y_{inn}^{k_f} - \sum_{k=k_f+1}^{K} \sum_{i=1}^{I} \sum_{n=1}^{I} \left( \left( \sum_{k=k_f+1}^{K} \left( p_n^k + t_{inn}^k \right) \right) \right) Y_{inn}^{k_f}
\]
\[
= \sum_{i=1}^{I} \sum_{n=1}^{I} \left\{ r_n - \left( \sum_{k=1}^{k_f-1} p_n^k + \sum_{k=k_f}^{K} p_n^k \right) - \left[ \sum_{k=1}^{k_f-1} t_{inn}^k + \sum_{k=k_f+1}^{K} t_{inn}^k \right] \right\} Y_{inn}^{k_f}.
\]
Let
\[ m_{L,in}(k_f) = r_n - \left( \sum_{k=1}^{k_f-1} p_{in}^k + \sum_{k=k_f}^K p_{in}^k \right) - \left[ \left( \sum_{k=1}^{k_f-1} t_{in}^k \right) + t_{in}^{k_f} + \sum_{k=k_f+1}^K t_{in}^k \right], \]
then the objective function (1) can be written as \( \sum_{i=1}^I \sum_{n=1}^m m_{L,in}(k_f) Y_{in}^{k_f} \), which is what we have in (7) for P2.(i).

If \( k_f = 1 \), objective function (1) can be written as
\[ \sum_{i=1}^I \sum_{n=1}^m r_n X_{in}^{k_f} - \sum_{i=1}^I \sum_{n=1}^m \left( p_{in}^k + t_{in}^k \right) X_{in}^{k_f} - \sum_{k=k_f+1}^K \sum_{i=1}^I \sum_{n=1}^m \left( p_{in}^k + t_{in}^k \right) X_{in}^{k_f} \]
substituting \( X_{in}^{k_f} = Y_{in}^{k_f} \) and \( X_{in}^{k_f} = Y_{in}^{k_f} \), where \( i, n = 1, 2, \ldots, I \) and \( k_f < k \leq K \), in the above, we will have:
\[ \sum_{n=1}^m \left[ r_n \sum_{i=1}^I Y_{in}^{k_f} - \sum_{i=1}^I \sum_{n=1}^m \left( p_{in}^k + t_{in}^k \right) Y_{in}^{k_f} - \sum_{k=k_f+1}^K \sum_{i=1}^I \sum_{n=1}^m \left( p_{in}^k + t_{in}^k \right) Y_{in}^{k_f} \right] \]
which is equivalent to (24) with the second term equal to 0. All the argument for \( k_f \neq 1 \) case then follows.
This completes the proof of Lemma 1.

Using Lemma 1, \( P1(A_{1full}(k_f), B, D) \) can be simplified to \( P1(m_L(k_f), D) \), where
\[ V^*(m_L(k_f), D) = \max_{\mu} \left\{ E_{D,Q(\mu)} \left[ \varpi(m_L(k_f), D, Q(\mu)) \right] - \sum_{k=1}^K \sum_{i=1}^I c_i^{k_f} \mu_i^{k_f} \right\} \].

We now introduce Lemma 2.

Lemma 2. For a multi-echelon supply chain with full process flexibility only at stage \( k_f \), let \( Z_{in}^{k_f} \) be the production flow from plant \( n \) at stage \( K \) to demand node \( n \) at stage 0 that is processed in plant \( i \) at stage \( k_f \) and in plant \( n \) at all the other stages, and let \( m_p(k_f) = (m_{P,in}(k_f)) \) be the matrix where element \( m_{P,in}(k_f) \) denotes the element at the \( n^{th} \) row and \( i^{th} \) column, problem
\( P2(A, B_{1full}(k_f), d, q) \) can be simplified to the following linear program, \( P2(m_p(k_f), d, q) \):
\[ \pi(m_p(k_f), d, q) = \max_{Z_{in}^{k_f}} \left\{ \sum_{n=1}^m \sum_{i=1}^I m_{P,in}(k_f) Z_{in}^{k_f} \right\} \]
subject to:
\[ \sum_{i=1}^I Z_{in}^{k_f} \leq \min_{1 \leq k < K, k \neq k_f} \left\{ q_n^k, d_n \right\}, \quad n = 1, 2, \ldots, I, \]
\[ \sum_{n=1}^m Z_{in}^{k_f} \leq q_i^{k_f}, \quad i = 1, 2, \ldots, I. \]

Using Lemma 2, \( P1(A, B_{1full}(k_f), D) \) can be simplified to \( P1(m_p(k_f), D) \), where
\[ V^*(m_p(k_f), D) = \max_{\mu} \left\{ E_{D,Q(\mu)} \left[ \varpi(m_p(k_f), D, Q(\mu)) \right] - \sum_{k=1}^K \sum_{i=1}^I c_i^{k_f} \mu_i^{k_f} \right\}. \]
PROOF OF LEMMA 2:

Since a flow $X_{ijn}^k$ can be non-zero only if $(i, j) \in A^k$, $(i, n) \in B^k$, and $(j, n) \in B^{k-1}$, for $B_{full}(k_f)$, we have

\begin{align}
X_{ijn}^k &\neq 0 \quad \text{only if} \quad i = n = j, \quad i, j, n = 1, 2, \ldots, I, \quad \forall \; k \neq k_f, k \neq k_f + 1, 1 \leq k \leq K, \quad (28) \\
X_{ijn}^k &\neq 0 \quad \text{only if} \quad j = n, \quad i, j, n = 1, 2, \ldots, I, \quad k = k_f, \quad (29) \\
X_{ijn}^k &\neq 0 \quad \text{only if} \quad i = n, \quad i, j, n = 1, 2, \ldots, I, \quad k = k_f + 1. \quad (30)
\end{align}

Therefore constraint (2) of $P_2(A, B, d, q)$ becomes

\begin{align}
X_{1nnn}^1 = \cdots = X_{nnn}^{k_f-1} = \sum_{i=1}^I X_{inn}^{k_f} &\quad n = 1, 2, \ldots, I, \quad (31) \\
X_{inn}^{k_f+1} = X_{inn}^{k_f+2} = \cdots = X_{inn}^K &\quad n, i = 1, 2, \ldots, I, \quad (32) \\
\sum_{i=1}^I X_{inn}^{k_f+1} = X_{nnn}^{k_f+2} = \cdots = X_{nnn}^K &\quad n = 1, 2, \ldots, I. \quad (33)
\end{align}

And constraint (3) of $P_2(A, B, d, q)$ becomes

\begin{align}
X_{nnn}^k &\leq q_n^k \quad n = 1, 2, \ldots, I, \quad \forall \; k \neq k_f, k \neq k_f + 1, 1 \leq k \leq K, \quad (34) \\
\sum_{n=1}^I X_{inn}^{k_f} &\leq q_i^{k_f} \quad i = 1, 2, \ldots, I, \quad (35) \\
\sum_{i=1}^I X_{inn}^{k_f+1} &\leq q_n^{k_f+1} \quad n = 1, 2, \ldots, I. \quad (36)
\end{align}

From (31) we have $X_{nnn}^k = \sum_{i=1}^I X_{inn}^{k_f}$, for $n = 1, 2, \ldots, I$, and $1 \leq k < k_f$. Substituting it into (34) we get

\begin{align}
\sum_{i=1}^I X_{inn}^{k_f} &\leq q_n^k \quad n = 1, 2, \ldots, I, \quad 1 \leq k < k_f. \quad (37)
\end{align}

From (32) and (33) we have $X_{nnn}^k = \sum_{i=1}^I X_{inn}^{k_f+1} = \sum_{i=1}^I X_{inn}^{k_f}$, for $n = 1, 2, \ldots, I$, and $k_f + 2 \leq k \leq K$. Substituting it into (34) we get

\begin{align}
\sum_{i=1}^I X_{inn}^{k_f} &\leq q_n^{k_f} \quad n = 1, 2, \ldots, I, \quad k_f + 2 \leq k \leq K. \quad (38)
\end{align}

And substituting (32) in (36) we have

\begin{align}
\sum_{i=1}^I X_{inn}^{k_f} &\leq q_n^{k_f+1} \quad n = 1, 2, \ldots, I. \quad (39)
\end{align}

If $k_f \neq 1$, constraint (4) of $P_2(\cdot)$ becomes

\begin{align}
X_{nnn}^1 &\leq d_n \quad n = 1, 2, \ldots, I, \quad (40)
\end{align}

From (31) we have $X_{nnn}^1 = \sum_{i=1}^I X_{inn}^{k_f}$, for $n = 1, 2, \ldots, I$, which by substituting into the above inequality we get

\begin{align}
\sum_{i=1}^I X_{inn}^{k_f} &\leq d_n \quad n = 1, 2, \ldots, I.
\end{align}
And if $k_f = 1$, constraint (4) is exactly same as (40). Combining (37), (38), (39) and (40) we have:

$$\sum_{i=1}^{l} X_{inn}^{k_f} \leq \min_{1 \leq k \leq K, k \neq k_f} \left\{ q_n, \ d_n \right\} \quad n = 1, 2, \ldots, I. \quad (41)$$

We define $Z_{in}^{k_f}$ as production flow of product $n$ that is produced in plant $i$ at stage $k_f$. This flow is also equal to the flow in plant $n$ at all the other stages, since we have process flexibility only at stage $k_f$. That is, $Z_{in}^{k_f} = X_{inn}^{k_f} = X_{inn}^{k_f+1}$. Substituting $X_{inn}^{k_f}$ and $X_{inn}^{k_f+1}$ with $Z_{in}^{k_f}$ in (41) and (35) we get constraints (26) and (27) in Lemma 2.

Now consider the objective function (1). Taking into account (28), (29) and (30), if $k_f \neq 1$, objective function (1) can be written as

$$\sum_{n=1}^{I} r_n X_{inn}^{1} - \sum_{k=1}^{K} \sum_{k \neq k_f, k_f+1} \left( \sum_{i=1}^{l} \left( p_{in} + t_{inn} \right) X_{inn}^{k} - \sum_{i=1}^{l} \left( \sum_{i=1}^{j} \left( p_{in} + t_{imm} \right) X_{inn}^{k} - \sum_{i=1}^{l} \sum_{i=1}^{j} \left( p_{in} + t_{inn} \right) X_{inn}^{k+1} \right) \right). \quad (42)$$

In the above function, by substituting $X_{inn}^{k_f} = X_{inn}^{k_f+1} = Z_{in}^{k_f}$, and $X_{inn}^{k} = \sum_{i=1}^{I} X_{inn}^{k_f} = \sum_{i=1}^{I} Z_{in}^{k_f}, \ \forall \ 1 \leq k \leq K, k \neq k_f, k_f + 1$, we will have:

$$\sum_{n=1}^{I} r_n X_{inn}^{1} - \sum_{k=1}^{K} \sum_{k \neq k_f, k_f+1} \left( \sum_{i=1}^{l} \left( p_{in} + t_{inn} \right) Z_{in}^{k_f} - \sum_{i=1}^{l} \sum_{i=1}^{j} \left( p_{in} + t_{inn} \right) Z_{in}^{k_f} - \sum_{i=1}^{l} \sum_{i=1}^{j} \left( p_{in} + t_{inn} \right) Z_{in}^{k_f+1} \right) \quad (42)$$

Let

$$m_{p,in}(k_f) = r_n - \left( \sum_{k=1}^{K} \frac{p_{in}^{k}}{k \neq k_f} + \frac{k_f}{k_f} \right) - \left( \sum_{k=1}^{K} \frac{t_{inn}^{k}}{k \neq k_f, k_f+1} + \frac{k_f}{k_f, k_f+1} \right) \quad (42)$$

then the objective function can be written as $\sum_{i=1}^{I} \sum_{i=1}^{I} m_{p,in}(k_f) Z_{in}^{k_f}$, which is the objective function in Lemma 2.

If $k_f = 1$, objective function (1) can be written as

$$\sum_{i=1}^{I} \sum_{i=1}^{I} r_i X_{inn}^{k_f} - \sum_{k=1}^{K} \sum_{k \neq k_f, k_f+1} \left( \sum_{i=1}^{l} \left( p_{in} + t_{inn} \right) X_{inn}^{k_f} - \sum_{i=1}^{l} \sum_{i=1}^{j} \left( p_{in} + t_{inn} \right) X_{inn}^{k_f+1} \right) \quad (42)$$

Substituting $X_{inn}^{k_f} = X_{inn}^{k_f+1} = Z_{in}^{k_f}$, and $X_{inn}^{k} = \sum_{i=1}^{I} X_{inn}^{k_f} = \sum_{i=1}^{I} Z_{in}^{k_f}, \ \forall \ 1 \leq k \leq K, k \neq k_f, k_f + 1$ in the above function, we get (42). All the argument for $k_f \neq 1$ case then follows. This completes the proof for Lemma 2.
Using Lemma 2, $\mathbf{P}_1(\mathbf{A}, \mathbf{B}_1 \text{full}(k_f), \mathbf{D})$ can be simplified to $\mathbf{P}_1(\mathbf{mp}(k_f), \mathbf{D})$, where

$$V^*(\mathbf{mp}(k_f), \mathbf{D}) = \max_{\mu} \left\{ \mathbb{E}_{\mathbf{D}, \mathbf{Q}(\mu)} \left[ \omega(\mathbf{mp}(k_f), \mathbf{D}, \mathbf{Q}(\mu)) \right] - \sum_{k=1}^{K} \sum_{i=1}^{I} c_k \mu_i \right\}.$$ 

Before we present the proof for Theorem 1, we first need to prove the following lemma. Note that this lemma holds when Assumptions 1 is relaxed.

**Lemma 3.** **CASE 1** (downstream variability): Suppose a supply chain has variability only at stage $k_v$ ($0 \leq k_v \leq K - 2$) and logistics flexibility only at stage $k_f$ ($k_v + 2 \leq k_f \leq K$). Let $\mu = (\mu_i^k)$ be the optimal capacity matrix for the supply chain. Then the optimal capacity matrix $\mu$ satisfies the following:

(a) the optimal capacity configuration is the same for all deterministic stages downstream from the flexible stage (i.e., there exist values $q_i^{\text{down}}$ such that $\mu_1^i = \mu_2^i = \cdots = \mu_k^i = q_i^{\text{down}}$, $i = 1, 2, \ldots, I$);

(b) the optimal capacity configuration is the same for all stages upstream of the flexible stage (i.e., there exist values $q_i^{\text{up}}$ such that $\mu_1^i = \mu_2^i = \cdots = \mu_k^i = q_i^{\text{up}}$, $i = 1, 2, \ldots, I$);

(c) $q_i^{\text{down}} \leq d_i$, for $1 \leq k_v \leq K - 2$, $i = 1, 2, \ldots, I$;

(d) $\sum_{i=1}^{I} q_i^{\text{up}} \leq \sum_{i=1}^{I} q_i^{\text{down}}$.

**CASE 2** (upstream variability): Suppose a supply chain has variability only at stage $k_v$ ($2 \leq k_v \leq K$) and logistics flexibility only at stage $k_f$, which is located downstream of stage $k_v$ (i.e., $1 \leq k_f \leq k_v - 1$). Let $\mu = (\mu_i^k)$ be the optimal capacity matrix for the supply chain. Then for $2 \leq k_f \leq k_v - 1$, the optimal capacity matrix $\mu$ satisfies

(a) the optimal capacity configuration is the same for all stages downstream of the flexible stage (i.e., there exist values $q_i^{\text{down}}$ such that $\mu_1^i = \mu_2^i = \cdots = \mu_k^i = q_i^{\text{down}}$, $i = 1, 2, \ldots, I$);

(b) the optimal capacity configuration is the same for all deterministic stages upstream of the flexible stage (i.e., there exist values $q_i^{\text{up}}$ such that $\mu_1^i = \mu_2^i = \cdots = \mu_k^i = q_i^{\text{up}}$, $i = 1, 2, \ldots, I$);

(c) $q_i^{\text{down}} \leq d_i$, for $i = 1, 2, \ldots, I$;

(d) $\sum_{i=1}^{I} q_i^{\text{up}} \geq \sum_{i=1}^{I} q_i^{\text{down}}$.

For $k_f = 1$, the optimal capacity matrix $\mu$ satisfies:

(a) the optimal capacity configuration is the same for all stages upstream of the flexible stage, except for stage $k_v$, (i.e., there exist values $q_i^{\text{up}}$ such that $\mu_1^i = \mu_2^i = \cdots = \mu_k^i = q_i^{\text{up}}$, $i = 1, 2, \ldots, I$);

(b) $\sum_{i=1}^{I} q_i^{\text{up}} \geq \sum_{i=1}^{I} d_i$.

**PROOF OF LEMMA 3 – Flexibility Upstream of Variability:**

We prove Lemma 3 by the following two cases, namely, $1 \leq k_v \leq K - 2$ and $k_v = 0$.

**CASE 1: $1 \leq k_v \leq K - 2$**

In order to prove the optimal capacity investment matrix $\mu$ satisfies conditions (a) to (d), we show that any $\mu$ that does not satisfy any of the four conditions cannot be optimal, i.e., there exists a matrix $\tilde{\mu}$, which satisfies the condition, and achieves higher total expected profit than $\mu$.  

18
Since the only variability is at stage $k_v$ ($1 \leq k_v \leq K - 2$), for all $k$'s such that $k \neq k_v$, we set $Q^k(\mu^k) = \mu^k$, as described in the paragraph following equation (6). Therefore, we have

$$Q(\mu) = \begin{pmatrix}
\mu_1^1 & \cdots & \mu_1^{k_v-1} & Q_1^1 & \mu_1^{k_v+1} & \cdots & \mu_1^K \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\mu_I^1 & \cdots & \mu_I^{k_v-1} & Q_I^1 & \mu_I^{k_v+1} & \cdots & \mu_I^K \\
\end{pmatrix},$$

where we consider $q$ as the matrix of the realization of capacities as follows:

$$q = \begin{pmatrix}
\mu_1^1 & \cdots & \mu_1^{k_v-1} & q_1^1 & \mu_1^{k_v+1} & \cdots & \mu_1^K \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\mu_I^1 & \cdots & \mu_I^{k_v-1} & q_I^1 & \mu_I^{k_v+1} & \cdots & \mu_I^K \\
\end{pmatrix}.$$

And since demand is deterministic,

$$D = (D_1, D_2, \ldots, D_I) = (d_1, d_2, \ldots, d_I) = d.$$

For full logistics flexibility configuration at stage $k_f$, where $k_v + 2 \leq k_f \leq K$, by Lemma 1, the optimal capacity investment $\mu$ is the solution to $P1(\mu_L(k_f), d)$:

$$V^*(\mu_L(k_f), d) = \max_\mu \left\{ V(\mu_L(k_f), d, Q(\mu)) \right\}$$

$$= \max_\mu \left\{ \mathbb{E}_{Q(\mu)} \left[ \pi(\mu_L(k_f), d, Q(\mu)) \right] - \sum_{k=1}^K \sum_{i=1}^I c_i^k \mu_i^k \right\},$$

where $\pi(\mu_L(k_f), d, Q(\mu))$ is calculated by solving $P2(\mu_L(k_f), d, q)$ and taking expectation over $Q(\mu)$:

$$\pi(\mu_L(k_f), d, q) = \max_{Y_{ij}^{k_f}} \left\{ \sum_{i=1}^I \sum_{j=1}^J [m_{L,ij}(k_f) \cdot Y_{ij}^{k_f}] \right\} \tag{43}$$

subject to:

$$\sum_{j=1}^J Y_{ij}^{k_f} \leq \min_{k_v \leq k \leq K} \{ \mu_i^k \} \quad i = 1, 2, \ldots, I, \tag{44}$$

$$\sum_{j=1}^J Y_{ij}^{k_f} \leq \min_{1 \leq k \leq k_v-1, k \neq k_v} \left\{ \mu_i^k, d_i, q_i^k \right\} \quad i = 1, 2, \ldots, I, \tag{45}$$

where

$$m_{L,ij}(k_f) = r_j - \left( \sum_{k=1}^{k_f} \sum_{j=1}^J p_{ij}^k + \sum_{k=k_f}^K \sum_{j=1}^J p_{ij}^k \right) - \left( \sum_{k=1}^{k_f} \sum_{j=1}^J t_{ij}^k + \sum_{k=k_f+1}^K \sum_{j=1}^J t_{ij}^k \right).$$

**CASE 1 – Part (a) and Part (b):**

Assume that $\mu$ does not satisfy condition (a) or (b). Let, for $i = 1, 2, \ldots, I$,

$$q_{\mu, i}^\text{dun} = \min_{1 \leq k \leq k_v-1, k \neq k_v} \{ \mu_i^k \} \quad \text{and} \quad q_{\mu, i}^\text{up} = \min_{k \leq k_v \leq K} \{ \mu_i^k \}$$

Now consider the capacity matrix $\tilde{\mu}$ as:

$$\tilde{\mu} = \begin{pmatrix}
q_{\mu, i}^\text{dun} & \cdots & q_{\mu, i}^\text{dun} & \mu_1^k & q_{\mu, i}^\text{dun} & \cdots & q_{\mu, i}^\text{dun} & q_{\mu, i}^\text{up} & \cdots & q_{\mu, i}^\text{up} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_{\mu, i}^\text{dun} & \cdots & q_{\mu, i}^\text{dun} & \mu_I^k & q_{\mu, i}^\text{dun} & \cdots & q_{\mu, i}^\text{dun} & q_{\mu, i}^\text{up} & \cdots & q_{\mu, i}^\text{up} \\
\end{pmatrix}.$$
If \( \mathbf{\mu} \) does not satisfy condition (a), then
\[
\mu_i^k > q_{\min,i}^{\text{dwn}} \quad \text{for some} \ i \ (1 \leq i \leq I) \text{ and some} \ k \ (1 \leq k \leq k_f - 1, \ k \neq k_v).
\]

If \( \mathbf{\mu} \) does not satisfy condition (b), then
\[
\mu_i^k > q_{\min,i}^{\text{up}} \quad \text{for some} \ i \ (1 \leq i \leq I) \text{ and some} \ k \ (k_f \leq k \leq K).
\]

In either case we have
\[
\sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k < \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k.
\]
which implies that capacity matrix \( \tilde{\mathbf{\mu}} \) has a lower capacity investment than that of \( \mathbf{\mu} \). On the other hand, since
\[
\min_{1 \leq k \leq k_f - 1, \ k \neq k_v} \{\mu_i^k\} = q_{\min,i}^{\text{dwn}} = \min_{1 \leq k \leq k_f - 1, \ k \neq k_v} \{\mu_i^k\} \quad i = 1, 2, \ldots, I,
\]
and
\[
\min_{k \leq k \leq k} \{\mu_i^k\} = q_{\min,i}^{\text{up}} = \min_{k \leq k \leq k} \{\mu_i^k\} \quad i = 1, 2, \ldots, I,
\]
then, the feasible region for problem \( \mathbf{P}_2 \) (i.e., constraints (44) and (45)) are the same for capacity matrices \( \mathbf{\mu} \) and \( \tilde{\mathbf{\mu}} \). On the other hand, since the objective function of problem \( \mathbf{P}_2(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{q}) \), i.e., (43), is independent of capacity matrices; hence, we have
\[
\varpi(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{Q}(\tilde{\mathbf{\mu}})) = \varpi(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{Q}(\mathbf{\mu})) \quad \text{for all realization of} \ \mathbf{Q}.
\]

Therefore, since \( \tilde{\mathbf{\mu}} \) has a lower capacity investment, it achieves higher expected net profit than \( \mathbf{\mu} \), and hence \( \mathbf{\mu} \) cannot be optimal solution to \( \mathbf{P}_1(\mathbf{m}_L(k_f), \mathbf{d}) \). Thus, the optimal capacity matrix \( \mathbf{\mu} \) must follow both condition (a) and (b).

**CASE 1 – Part (c):**

To prove part (c), suppose
\[
\mathbf{\mu} = 
\begin{pmatrix}
q_{1}^{\text{dwn}} & \cdots & q_{1}^{\text{dwn}} & \mu_{1} & q_{1}^{\text{dwn}} & \cdots & q_{1}^{\text{dwn}} & q_{1}^{\text{up}} & \cdots & q_{1}^{\text{up}} \\
\vdots & \ddots & \ddots & \vdots & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots \\
q_{I}^{\text{dwn}} & \cdots & q_{I}^{\text{dwn}} & \mu_{I} & q_{I}^{\text{dwn}} & \cdots & q_{I}^{\text{dwn}} & q_{I}^{\text{up}} & \cdots & q_{I}^{\text{up}}
\end{pmatrix}
\]
satisfies condition (a) and (b), but does not satisfy (c), i.e., \( q_{i'}^{\text{dwn}} > d_{i'} \) for some \( i' \ (1 \leq i' \leq I) \). Let
\[
\tilde{\mathbf{\mu}} = 
\begin{pmatrix}
q_{1}^{\text{dwn}} & \cdots & q_{1}^{\text{dwn}} & \mu_{1} & q_{1}^{\text{dwn}} & \cdots & q_{1}^{\text{dwn}} & q_{1}^{\text{up}} & \cdots & q_{1}^{\text{up}} \\
\vdots & \ddots & \ddots & \vdots & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots \\
d_{i'} & \cdots & d_{i'} & \mu_{i'} & d_{i'} & \cdots & d_{i'} & q_{i'}^{\text{up}} & \cdots & q_{i'}^{\text{up}} \\
\vdots & \ddots & \ddots & \vdots & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots \\
q_{I}^{\text{dwn}} & \cdots & q_{I}^{\text{dwn}} & \mu_{I} & q_{I}^{\text{dwn}} & \cdots & q_{I}^{\text{dwn}} & q_{I}^{\text{up}} & \cdots & q_{I}^{\text{up}}
\end{pmatrix},
\]
that is, substituting all \( q_{i'}^{\text{dwn}} \) in \( \mathbf{\mu} \) with \( d_{i'} \). Since \( q_{i'}^{\text{dwn}} > d_{i'} \) for \( i' \ (1 \leq i' \leq I) \), we have
\[
\sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k < \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k.
\]
On the other hand, the only difference between capacity matrices \( \mu \) and \( \tilde{\mu} \) is in plant \( i' \) at stages downstream of \( k_f \), except for stage \( k_v \), hence the righthand-side of constraint (45) for capacity matrices \( \mu \) and \( \tilde{\mu} \), respectively, for \( i' \) is

\[
egin{align*}
\min_{1 \leq k \leq k_f - 1, k \neq k_v} \{ \mu^k_i, d_i, q^k_i \} &= \min \{ q^\text{down}_i, d_i, q^k_i \} = \min \{ q^\text{down}_i, q^k_i \}, \\
\min_{1 \leq k \leq k_f - 1, k \neq k_v} \{ \tilde{\mu}^k_i, d_i, q^k_i \} &= \min \{ d_i, q^k_i \} = \min \{ d_i, q^k_i \}.
\end{align*}
\]

Therefore, the feasible region, i.e., constraints (44) and (45), are the same for capacity matrices \( \mu \) and \( \tilde{\mu} \). So, similar to Cases (a) and (b), we have

\[
\varpi(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{Q}(\tilde{\mu})) = \varpi(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{Q}(\mu)) \quad \text{for all realization of } \mathbf{Q}.
\]

Therefore, \( \tilde{\mu} \) achieves higher expected net profit than \( \mu \), and hence \( \mu \) cannot be optimal solution to \( \text{P1}(\mathbf{m}_L(k_f), \mathbf{d}) \). Thus, the optimal capacity matrix \( \mu \) must follow condition (c).

**CASE 1 – Part (d):**

Let

\[
\mu = \begin{pmatrix}
q^\text{down}_1 & \ldots & q^\text{down}_{k-1} & \mu^k_{k_v} & q^\text{down}_{k+1} & \ldots & q^\text{down}_{k_f} & \ldots & q^\text{up}_{k_f} & \ldots & q^\text{up}_K
\end{pmatrix}
\]

\[
\tilde{\mu} = \begin{pmatrix}
q^\text{down}_1 & \ldots & q^\text{down}_{k-1} & \mu^k_{k_v} & q^\text{down}_{k+1} & \ldots & q^\text{down}_{k_f} & \ldots & q^\text{down}_{k_f} & \ldots & q^\text{down}_K
\end{pmatrix}
\]

both of which satisfy condition (a), (b) and (c). Suppose \( \sum_{i=1}^{I} q^\text{up}_i > \sum_{i=1}^{I} q^\text{down}_i \), i.e. \( \mu \) does not satisfy condition (d). It is clear that

\[
\sum_{k=1}^{K} \sum_{i=1}^{I} c_{i} k_{i} \mu_{i} - \sum_{k=1}^{K} \sum_{i=1}^{I} c_{i} k_{i} \tilde{\mu}_{i} = \sum_{k=k_f}^{K} \sum_{i=1}^{I} c_{i} k_{i} (q^\text{up}_i - q^\text{down}_i) > 0,
\]

since \( c_{i} k_{i} = c_{i} \), \( i, j = 1, 2, \ldots, I, \) \( i \neq j \), and \( k = 1, 2, \ldots, K \). Thus, capacity matrix \( \tilde{\mu} \) has a lower capacity investment cost than that of \( \mu \).

For capacity investment vector \( \tilde{\mu} \), \( \varpi(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{Q}(\tilde{\mu})) \) is calculated by solving \( \text{P2}(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{q}) \) and taking expectation over \( \mathbf{Q}(\tilde{\mu}) \):

\[
\pi(\mathbf{m}_L(k_f), \mathbf{d}, \mathbf{q}) = \max_{\mathbf{Y}_{ij}^{k_f}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{J} \left[ \mathbf{m}_L(ij(k_f)) \cdot \mathbf{Y}_{ij}^{k_f} \right] \right\}
\]

subject to:

\[
\sum_{j=1}^{J} Y_{ij}^{k_f} \leq q^\text{down}_i \quad i = 1, 2, \ldots, I, \quad (47)
\]

\[
\sum_{j=1}^{J} Y_{ij}^{k_f} \leq \min \{ q^\text{down}_i, d_i, q^k_i \} = \min \{ q^\text{down}_i, q^k_i \} \quad \text{(since (c) holds)} \quad i = 1, 2, \ldots, I. \quad (48)
\]
On the other hand, since according to Assumption 3,
\[ m_{L,i}(k_f) > m_{L,j}(k_f) \quad \text{and} \quad m_{L,i}(k_f) > m_{L,j}(k_f) \quad \forall \; j \neq i, \quad i = 1, 2, \ldots, I, \]
the most profitable way of allocating flows (Y_{ij}^{k_f}$'s) is to use dedicated ones. Since by comparing right-hand-side of (47) and (48) we have \( q_i^{dwn} \geq \min \{ q_i^{dwn}, \, q_i^{k_v} \} \), we get the maximum profit for \( \tilde{\mu} \) by setting \( Y_{ij} = \min \{ q_i^{dwn}, \, q_i^{k_v} \} \) and \( Y_{ij}^{k_f} = 0 \) \( \forall \, i \neq j \):
\[
\pi(m_{L}(k_f), d, q) = \sum_{i=1}^{I} \min \{ m_{L,i}(k_f) \cdot q_i^{dwn}, \, q_i^{k_v} \}.
\]

On the other hand, for capacity investment vector \( \mu \), \( \varpi(m_{L}(k_f), d, Q(\mu)) \) is calculated by solving \( P2(m_{L}(k_f), d, q) \) and taking expectation over \( Q(\mu) \):
\[
\pi(m_{L}(k_f), d, q) = \max_{Y_{ij}^{k_f}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} \left[ m_{L,i}(k_f) \cdot Y_{ij}^{k_f} \right] \right\}
\]
such that:
\[
\sum_{j=1}^{I} Y_{ij}^{k_f} \leq q_i^{up} \quad \text{and} \quad q_i^{dwn} \leq \min \{ q_i^{dwn}, \, q_i^{k_v} \} \quad \forall \, i = 1, 2, \ldots, I
\]
\[ \sum_{j=1}^{I} Y_{ij}^{k_f} \leq \min \{ q_i^{dwn}, \, q_i^{k_v} \} \quad \forall \, i = 1, 2, \ldots, I. \]

Since \( \sum_{i=1}^{I} q_i^{up} > \sum_{i=1}^{I} q_i^{dwn} \), we have \( \sum_{i=1}^{I} q_i^{up} > \sum_{i=1}^{I} \min \{ q_i^{dwn}, \, q_i^{k_v} \} \) for any \( q \), hence
\[
\sum_{i=1}^{I} \sum_{j=1}^{I} m_{L,i}(k_f) \cdot Y_{ij}^{k_f} = \sum_{i=1}^{I} \sum_{j=1}^{I} m_{L,j}(k_f) \cdot Y_{ji}^{k_f} \leq \sum_{i=1}^{I} \left[ m_{L,i}(k_f) \sum_{j=1}^{I} Y_{ij}^{k_f} \right] \leq \sum_{i=1}^{I} \left[ m_{L,i}(k_f) \cdot \min \{ q_i^{dwn}, \, q_i^{k_v} \} \right].
\]

Therefore
\[
\varpi(m_{L}(k_f), d, Q(\mu)) \leq \varpi(m_{L}(k_f), d, Q(\tilde{\mu})) \quad \text{for all realization of} \; Q \tag{49}
\]
By (46) and (49),
\[ V(m_{L}(k_f), d, Q(\mu)) < V(m_{L}(k_f), d, Q(\tilde{\mu})), \]
and hence \( \mu \) cannot be optimal solution to \( P1(m_{L}(k_f), d) \). Thus, the optimal capacity matrix \( \mu \) must follow condition (d).

**CASE 2: \( k_v = 0 \)**
The proof for the case with \( k_v = 0 \) is similar to that for \( k_v \neq 0 \), except for the following minor modifications, given that now demand is the only source of variability and capacity is deterministic:

(1) \( Q(\mu) = q = \mu \);
(2) Since \( D \) is stochastic and \( Q(\mu) \) is deterministic, to calculate \( V^* \), the expectation is taken over \( D \) instead of \( Q(\mu) \);
(3) Stage \(k_v\) with uncertain capacity is missing in capacity matrices \(Q(\mu), q, \mu\) and \(\tilde{\mu}\), and in constraints such as (45);

(4) In part (d) constraint (48), instead of \(q_i^{k_v}\), we have \(d_i\).

**PROOF OF LEMMA 3 – Flexibility Downstream of Variability:**

**CASE 1:** \(2 \leq k_f \leq k_v - 1\)

In order to prove the optimal capacity vector \(\mu\) satisfies conditions (a) to (d), similar to the proof of the previous case, we show that for any capacity matrix \(\mu\) which does not satisfy any of the four conditions, there exists a capacity matrix \(\tilde{\mu}\), which satisfies the condition and achieves higher total expected profit than \(\mu\); therefore, \(\mu\) cannot be optimal.

Since the only variability is at stage \(k_v\) (\(2 \leq k_v \leq K\)), we have

\[
Q(\mu) = \begin{pmatrix}
\mu_1^{k_v} & \cdots & \mu_{1}^{k_v-1} & Q_1 & \mu_{2}^{k_v+1} & \cdots & \mu_{K}^{k_v+1} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mu_1 & \cdots & \mu_{k_v-1} & Q & \mu_{k_v+1} & \cdots & \mu_{K} \\
\end{pmatrix},
\]

and the matrix of the realization of the above capacity matrix is:

\[
q = \begin{pmatrix}
\mu_1^{k_v} & \cdots & \mu_{1}^{k_v-1} & q_1 & \mu_{2}^{k_v+1} & \cdots & \mu_{K}^{k_v+1} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mu_1 & \cdots & \mu_{k_v-1} & q & \mu_{k_v+1} & \cdots & \mu_{K} \\
\end{pmatrix},
\]

Since \(D\) is deterministic, we have

\[D = (D_1, D_2, \ldots, D_f) = (d_1, d_2, \ldots, d_f) = d.\]

For supply chain with variability at stage \(k_v\) and full logistics flexibility at stage \(k_f\), by Lemma 1, the optimal capacity investment \(\mu\) is solution to \(P_1(m_L(k_f), d)\):

\[
V^*(m_L(k_f), d) = \max_{\mu} \left\{ V(m_L(k_f), d, Q(\mu)) \right\} = \max_{\mu} \left\{ E_{Q(\mu)} \left[ \pi(m_L(k_f), d, Q(\mu)) \right] - \sum_{k=1}^{K} \sum_{i=1}^{I} \mu_i^k \right\},
\]

where \(\pi(m_L(k_f), d, Q(\mu))\) is calculated by solving \(P_2(m_L(k_f), d, q)\) and taking expectation over \(Q(\mu)\):

\[
\pi(m_L(k_f), d, q) = \max_{Y_{ij}^{k_f}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{J} \left[ m_{L,ij}(k_f) \cdot Y_{ij}^{k_f} \right] \right\}
\]

subject to:

\[
\sum_{j=1}^{I} Y_{ij}^{k_f} \leq \min_{k_f \leq k \leq K, k \neq k_v} \left\{ \mu_i^k, q_i^k \right\} \quad i = 1, 2, \ldots, I,
\]

\[
\sum_{j=1}^{I} Y_{ij}^{k_f} \leq \min_{i \leq k \leq k - 1} \left\{ \mu_i^k, d_i \right\} \quad i = 1, 2, \ldots, I.
\]
CASE 1 – Part (a) and Part (b):

Assume that \( \mu \) does not satisfy condition (a) or (b). Let, for \( i = 1, 2, \ldots, I \),

\[
q^{\text{dun},i}_{\min} = \min_{1 \leq k \leq k_f-1} \{ \mu^k_i \} \quad \text{and} \quad q^{\text{up},i}_{\min} = \min_{k_f \leq k \leq K, k \neq k_v} \{ \mu^k_i \} \quad i = 1, 2, \ldots, I,
\]

and

\[
\tilde{\mu} = \begin{pmatrix}
q^{\text{dun},1}_{\min,1} & \cdots & q^{\text{dun},1}_{\min,I} & q^{\text{dun},(k_f-1)}_{\min,1} & \cdots & q^{\text{dun},(k_f-1)}_{\min,I} & q^{\text{dun},k_f}_{\min,1} & \cdots & q^{\text{dun},k_f}_{\min,I} & q^{\text{dun},(k_v+1)}_{\min,1} & \cdots & q^{\text{dun},(k_v+1)}_{\min,I} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q^{\text{dun},I}_{\min,1} & \cdots & q^{\text{dun},I}_{\min,1} & q^{\text{dun},I}_{\min,I} & \cdots & q^{\text{dun},I}_{\min,I} & q^{\text{dun},I}_{\min,1} & \cdots & q^{\text{dun},I}_{\min,I} & q^{\text{dun},I}_{\min,1} & \cdots & q^{\text{dun},I}_{\min,I}
\end{pmatrix}.
\]

If \( \mu \) does not satisfy condition (a), then

\[
\mu^k_i > q^{\text{dun},i}_{\min} \quad \text{for some} \quad i \ (1 \leq i \leq I) \quad \text{and some} \quad k \ (1 \leq k \leq k_f - 1).
\]

If \( \mu \) does not satisfy condition (b), then

\[
\mu^k_i > q^{\text{up},i}_{\min} \quad \text{for some} \quad i \ (1 \leq i \leq I) \quad \text{and some} \quad k \ (k_f \leq k \leq K, k \neq k_v).
\]

In either case we have

\[
\sum_{k=1}^{K} \sum_{i=1}^{I} q^{\text{dun}}_{k,i} \mu^k_i < \sum_{k=1}^{K} \sum_{i=1}^{I} q^{\text{dun}}_{k,i} \tilde{\mu}^k_i,
\]

which implies that \( \tilde{\mu} \) has a lower capacity investment cost than \( \mu \). On the other hand, since

\[
\min_{1 \leq k \leq k_f-1} \{ \tilde{\mu}^k_i \} = q^{\text{dun},i}_{\min} = \min_{1 \leq k \leq k_f-1} \{ \mu^k_i \} \quad i = 1, 2, \ldots, I,
\]

and

\[
\min_{k_f \leq k \leq K, k \neq k_v} \{ \tilde{\mu}^k_i \} = q^{\text{up},i}_{\min} = \min_{k_f \leq k \leq K, k \neq k_v} \{ \mu^k_i \} \quad i = 1, 2, \ldots, I,
\]

constraints (50) and (51) result in the same feasible region for both capacity matrices \( \mu \) and \( \tilde{\mu} \). So we have

\[
\varpi(\text{m}_L(k_f), d, Q(\tilde{\mu})) = \varpi(\text{m}_L(k_f), d, Q(\mu)) \quad \text{for all realization of} \quad Q.
\]

Therefore,

\[
V(\text{m}_L(k_f), d, Q(\tilde{\mu})) > V(\text{m}_L(k_f), d, Q(\mu)),
\]

and \( \tilde{\mu} \) achieves higher expected net profit than \( \mu \), and hence \( \mu \) cannot be optimal solution to \( P1(\text{m}_L(k_f), d) \).

Thus, the optimal capacity matrix \( \mu \) must follow both condition (a) and (b).

CASE 1 – Part (c):

To prove (c), suppose

\[
\mu = \begin{pmatrix}
q^{\text{dun}}_{1,1} & \cdots & q^{\text{dun}}_{1,I} & q^{\text{up}}_{1,1} & \cdots & q^{\text{up}}_{1,I} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q^{\text{dun}}_{I,1} & \cdots & q^{\text{dun}}_{I,I} & q^{\text{up}}_{I,1} & \cdots & q^{\text{up}}_{I,I}
\end{pmatrix},
\]

which satisfies conditions (a) and (b), but not (c), i.e., \( q^{\text{dun}}_{i'} > d_{i'} \) for some \( i' \) \( (1 \leq i' \leq I) \). Now consider capacity matrix \( \tilde{\mu} \) as follows:

\[
\tilde{\mu} = \begin{pmatrix}
q^{\text{dun}}_{1,1} & \cdots & q^{\text{dun}}_{1,I} & q^{\text{up}}_{1,1} & \cdots & q^{\text{up}}_{1,I} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
d_{i'} & \cdots & q^{\text{up}}_{i',1} & \cdots & q^{\text{up}}_{i',I} & \cdots & q^{\text{up}}_{i',I} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
q^{\text{dun}}_{I,1} & \cdots & q^{\text{dun}}_{I,I} & q^{\text{up}}_{I,1} & \cdots & q^{\text{up}}_{I,I}
\end{pmatrix},
\]

24
that is, substituting all \( q_i^{dun} \) in \( \mu \) with \( d_i' \) and get \( \tilde{\mu} \). Since \( q_i^{dun} > d_i' \) for \( i' \leq i \leq I \), we have

\[
\sum_{k=1}^{K} \sum_{l=1}^{I} c_k^l \tilde{\mu}_i^l < \sum_{k=1}^{K} \sum_{l=1}^{I} c_k^l \mu_i^l.
\]

On the other hand, the only difference between capacity matrices \( \mu \) and \( \tilde{\mu} \) is in plant \( i' \) at stages downstream of \( k_f \); hence, the right-hand-side of constraint (51) for capacity matrix \( \mu \) and \( \tilde{\mu} \), respectively, for \( i' \) is

\[
\begin{align*}
\min_{1 \leq k \leq k_f-1} \{ \mu_k^i, d_i' \} &= \min \{ q_i^{dun}, d_i' \} = d_i', \\
\min_{1 \leq k \leq k_f-1} \{ \tilde{\mu}_k^i, d_i' \} &= \min \{ \tilde{d}_i, d_i' \} = d_i'.
\end{align*}
\]

Hence constraints (50) and (51) are the same for capacity matrices \( \mu \) and \( \tilde{\mu} \). So we have

\[
\varpi(\mathbf{m}_f(k_f), \mathbf{d}, \mathbf{Q} ) = \varpi(\mathbf{m}_f(k_f), \mathbf{d}, \mathbf{Q} ) \quad \text{for all realization of } \mathbf{Q}.
\]

Therefore,

\[
V(\mathbf{m}_f(k_f), \mathbf{d}, \mathbf{Q} ) > V(\mathbf{m}_f(k_f), \mathbf{d}, \mathbf{Q} )
\]

which implies that \( \mu \) cannot be optimal solution to P1(\( \mathbf{m}_f(k_f), \mathbf{d} \)). Thus, the optimal capacity matrix \( \mu \) must follow condition (c).

**CASE 1 – Part (d):**

For this part, we establish a new unit margin matrix \( \mathbf{m}_{min} \) based on the smallest margin value of the original unit margin matrix \( \mathbf{m} \). We first prove that a capacity matrix which satisfies condition (d) achieves higher expected profit than a capacity matrix that does not under this new margin matrix. And then we show that this is also true under original unit margin matrix \( \mathbf{m} \).

For proof in this part, we model the uncertain capacity of plant \( i \) at stage \( k_v \) as \( Q_i^{k_v}(\mu_i^{k_v}) = \mu_i^{k_v} + \epsilon_i^{k_v} \), where \( \mu_i^{k_v} \) is investment level decided by the firm and \( \epsilon_i^{k_v} \) is an additive random term to account for variability in the capacity of plant \( i \) at stage \( k_v \). Let \( \epsilon \) be the corresponding random vector. Let \( \varepsilon \) be observation of \( \epsilon \).

Assume that capacity matrix \( \mu \) is:

\[
\mu = 
\begin{pmatrix}
q_1^{dun} & \cdots & q_1^{dun} & q_I^{up} & \cdots & q_I^{up} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_I^{dun} & \cdots & q_I^{dun} & q_I^{up} & \cdots & q_I^{up}
\end{pmatrix},
\]

which satisfies conditions (a) to (c), but not (d). That is, \( \sum_{i=1}^{I} q_i^{up} < \sum_{i=1}^{I} q_i^{dun} \). Now consider another capacity matrix

\[
\tilde{\mu} = 
\begin{pmatrix}
q_1^{dun} & \cdots & q_1^{dun} & \tilde{\mu}_I & \cdots & \tilde{\mu}_I \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_I^{dun} & \cdots & q_I^{dun} & \tilde{\mu}_I & \cdots & \tilde{\mu}_I
\end{pmatrix}
\]

that satisfies conditions (a) to (d), with \( \sum_{i=1}^{I} \tilde{\mu}_i = \sum_{i=1}^{I} q_i^{dun} \). In the following paragraphs, we refer to the change from \( \mu \) to \( \tilde{\mu} \), i.e., increasing the sum of capacities at stages upstream of flexibility, as increasing capacity from \( \mu \) to \( \tilde{\mu} \).

Let \( m_{L, min} = \min_{i,j \leq I} \{ m_{L, ij}(k_f) \} \) and

\[
\mathbf{m}_{min} = 
\begin{pmatrix}
m_{L, min} & \cdots & m_{L, min} \\
\vdots & \ddots & \vdots \\
m_{L, min} & \cdots & m_{L, min}
\end{pmatrix}
\]

25
Under the new unit margin matrix $m_{l_{min}}$, for capacity investment $\mu$,

$$\pi(m_{l_{min}}, d, q) = \max_{Y_{ij}^{k}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{J} (m_{L_{min}} \cdot Y_{ij}^{k}) \right\} = \max_{Y_{ij}^{k}} \left\{ m_{L_{min}} \cdot \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}^{k} \right\}.$$ 

Since $\mu$ follows condition (b), $\min_{k \leq k \leq k} \{ \mu_{k}^{b}, \mu_{k}^{c} \} = \min \{ q_{i}^{up}, q_{i}^{down} \}$, from (50) we have

$$\sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}^{k} \leq \sum_{i=1}^{I} \min \{ q_{i}^{up}, q_{i}^{down} \}. \quad (52)$$

Similarly, since $\mu$ follows condition (a) and (c), $\min_{i \leq k \leq i-1} \{ \mu_{i}^{k}, \mu_{i}^{c} \} = \min \{ q_{i}^{down}, d_{i} \} = q_{i}^{down}$, from (51) we have

$$\sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}^{k} \leq \sum_{i=1}^{I} q_{i}^{down}. \quad (53)$$

Since $\sum_{i=1}^{I} q_{i}^{up} < \sum_{i=1}^{I} q_{i}^{down}$, then $\sum_{i=1}^{I} \min \{ q_{i}^{up}, q_{i}^{down} \} < \sum_{i=1}^{I} q_{i}^{down}$, combining with (52) and (53), the maximum value that $\sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}^{k}$ can achieve is $\sum_{i=1}^{I} \min \{ q_{i}^{up}, q_{i}^{down} \}$. Thus, the optimal solution to problem $P2(m_{min}, d, q)$ is

$$\pi(m_{l_{min}}, d, q) = \max_{Y_{ij}^{k}} \left\{ m_{L_{min}} \cdot \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij}^{k} \right\} = m_{L_{min}} \cdot \sum_{i=1}^{I} \min \{ q_{i}^{up}, q_{i}^{down} \}.$$ 

Hence

$$V(m_{l_{min}}, d, Q(\mu)) = E_{Q(\mu)} \left[ \pi(m_{l_{min}}, d, Q(\mu)) \right] = m_{L_{min}} \cdot \sum_{i=1}^{I} \sum_{j=1}^{J} \mu_{i}^{k} \cdot \sum_{i=1}^{I} \min \{ q_{i}^{up}, q_{i}^{down} \} = m_{L_{min}} \cdot \sum_{i=1}^{I} \min \{ q_{i}^{up}, q_{i}^{down} \}.$$ 

where $f(q_{i}^{k})$ is probability density function (pdf) of $Q_{i}^{k}(\mu_{i}^{k})$.

Since $Q_{i}^{k}(\mu_{i}^{k}) = \mu_{i}^{k} + \epsilon_{i}^{k} = r(\epsilon_{i}^{k})$, we have $\epsilon_{i}^{k} = Q_{i}^{k}(\mu_{i}^{k}) - \mu_{i}^{k} \approx s(Q_{i}^{k}(\mu_{i}^{k})).$ Let $g(\epsilon_{i}^{k})$ be pdf of $\epsilon_{i}^{k}$, then

$$g(\epsilon_{i}^{k}) = f(r(\epsilon_{i}^{k})) \left| \frac{dr(\epsilon_{i}^{k})}{d\epsilon_{i}^{k}} \right| = f(\mu_{i}^{k} + \epsilon_{i}^{k}),$$

and

$$f(q_{i}^{k}) = g(s(q_{i}^{k})) \left| \frac{ds(q_{i}^{k})}{dq_{i}^{k}} \right| = g(q_{i}^{k} - \mu_{i}^{k}).$$

So

1^Note that for the sake of calculation, we assume the integration is from $-\infty$ to $\infty$. The approach can be easily revised for cases where the random capacity is bounded from below and from above. For those cases one can set the probability of the random capacity being less than the lower bound and larger than the upper bound to zero in our calculation.

26
\[ V(\cdot) = m_{L,\min} \sum_{i=1}^{I} \left( \int_{-\infty}^{q_{i}^{up} - \mu_{i}^{k_i}} (\mu_{i}^{k_i} + \varepsilon_{i}^{k_i}) f(\mu_{i}^{k_i} + \varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} + \int_{q_{i}^{up} - \mu_{i}^{k_i}}^{\infty} q_{i}^{up} f(\mu_{i}^{k_i} + \varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} \right) \right. \\
\left. - \sum_{k=1}^{K} \sum_{i=1}^{I} c_{i}^{k} \mu_{i}^{k_i} - \sum_{k=1}^{K} \sum_{i=1}^{I} c_{i}^{k} q_{i}^{up} \right)

= m_{L,\min} \sum_{i=1}^{I} \left( \int_{-\infty}^{q_{i}^{up} - \mu_{i}^{k_i}} (\mu_{i}^{k_i} + \varepsilon_{i}^{k_i}) g(\varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} + \int_{q_{i}^{up} - \mu_{i}^{k_i}}^{\infty} q_{i}^{up} g(\varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} \right) \\
- \sum_{k=1}^{K} \sum_{i=1}^{I} c_{i}^{k} q_{i}^{up}.

The optimal \( \mu_{i}^{k_i}^{*} \) satisfies

\[
\frac{\partial V}{\partial \mu_{i}^{k_i}} \bigg|_{\mu_{i}^{k_i} = \mu_{i}^{k_i}^{*}} = \left[ m_{L,\min} \int_{-\infty}^{q_{i}^{up} - \mu_{i}^{k_i}} \frac{d(q_{i}^{up} - \mu_{i}^{k_i})}{d\mu_{i}^{k_i}} (\mu_{i}^{k_i} + q_{i}^{up} - \mu_{i}^{k_i}) g(\varepsilon_{i}^{k_i}) + m_{L,\min} \int_{q_{i}^{up} - \mu_{i}^{k_i}}^{\infty} \frac{d(q_{i}^{up} - \mu_{i}^{k_i})}{d\mu_{i}^{k_i}} g(\varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} \right] \\
- m_{L,\min} \sum_{k=1}^{K} \sum_{i=1}^{I} c_{i}^{k} q_{i}^{up} g(\varepsilon_{i}^{k_i}) - c_{i}^{k} \left] \mid_{\mu_{i}^{k_i} = \mu_{i}^{k_i}^{*}} \right.

= m_{L,\min} \int_{-\infty}^{q_{i}^{up} - \mu_{i}^{k_i}} g(\varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} - c_{i}^{k}.

(54)

On the other hand,

\[
\frac{\partial V}{\partial q_{i}^{up}} = m_{L,\min} \int_{-\infty}^{q_{i}^{up} - \mu_{i}^{k_i}} \frac{d(q_{i}^{up} - \mu_{i}^{k_i})}{dq_{i}^{up}} (\mu_{i}^{k_i} + q_{i}^{up} - \mu_{i}^{k_i}) g(\varepsilon_{i}^{k_i}) \\
- m_{L,\min} \int_{q_{i}^{up} - \mu_{i}^{k_i}}^{\infty} q_{i}^{up} g(\varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} - \sum_{k=k_f, k\neq k_v}^{K} c_{i}^{k}.

Therefore, by envelope theory,

\[
\frac{\partial V^{*}}{\partial q_{i}^{up}} = \frac{\partial V}{\partial q_{i}^{up}} \bigg|_{\mu_{i}^{k_i} = \mu_{i}^{k_i}^{*}} = m_{L,\min} \int_{q_{i}^{up} - \mu_{i}^{k_i}}^{\infty} g(\varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} - \sum_{k=k_f, k\neq k_v}^{K} c_{i}^{k},

but by (54) we know that

\[ m_{L,\min} \int_{q_{i}^{up} - \mu_{i}^{k_i}}^{\infty} g(\varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} = m_{L,\min} - m_{L,\min} \int_{-\infty}^{q_{i}^{up} - \mu_{i}^{k_i}} g(\varepsilon_{i}^{k_i}) d\varepsilon_{i}^{k_i} = m_{L,\min} - \varepsilon_{i}^{k_i}. \]
hence,

\[
\frac{\partial V^*}{\partial q^*_{i+k}} = m_{L,\text{min}} - c^i_k - \sum_{k=k_f, k \neq k_f}^K c^k_k
\]

\[
= m_{L,\text{min}} - \sum_{k=k_f}^K c^k_k
\]

\[
> m_{L,\text{min}} - \sum_{k=1}^K c^k_k
\]

\[
> 0,
\]

since we assume the supply chain makes profit even through the least profitable arc.

Therefore, for a capacity matrix \( \mu \) with \( \sum_{i=1}^I q^i_{d, \text{up}} > \sum_{i=1}^I q^i_{d, \text{down}} \), profit can be improved by increasing \( q^i_{d, \text{up}} \) for some \( i \) until the sum of capacities for the upstream and downstream stages equate. Hence

\[
V(m_{L,\text{min}}, d, Q(\tilde{\mu})) > V(m_{L,\text{min}}, d, Q(\mu)).
\]

That is, under the profit margin matrix \( m_{L,\text{min}} \), increasing capacity from \( \mu \) to \( \tilde{\mu} \) increases profit \( V(\cdot) \).

Now consider the original profit margin matrix \( m_{L,\text{min}} \). We know that the profit \( V(m_{L,\text{d}}, Q(\mu)) \) consists of two parts, revenue \( E[Q(\mu)] \cdot \mathbb{E}(m_{L,\text{d}}, Q(\mu)) \) and cost \( \sum_{k=1}^K \sum_{i=1}^I c^k_i \mu^k_i \). Compared with the case under margin matrix \( m_{L,\text{min}} \), increasing capacity from \( \mu \) to \( \tilde{\mu} \) under profit margin matrix \( m_{L,\text{min}} \) increases \( E[\pi(\cdot)] \) even more because of higher profit margin. On the other hand, the extra cost of increasing capacity from \( \mu \) to \( \tilde{\mu} \), \( \sum_{k=1}^K \sum_{i=1}^I c^k_i \mu^k_i - \sum_{k=1}^K \sum_{i=1}^I c^k_i \mu^k_i \), is independent of the margin. Therefore, under the profit margin matrix \( m_{L,\text{min}} \), increasing capacity from \( \mu \) to \( \tilde{\mu} \) increases profit \( V(\cdot) \) more than the case under profit margin matrix \( m_{L,\text{min}} \), i.e.,

\[
V(m_{L,\text{d}}, Q(\tilde{\mu})) - V(m_{L,\text{d}}, Q(\mu)) \geq V(m_{L,\text{min}}, d, Q(\tilde{\mu})) - V(m_{L,\text{min}}, d, Q(\mu))
\]

\[
> 0.
\]

and so \( \mu \) is not optimal solution to \( P1(m_{L,\text{d}}(k_f), d) \).

\section*{CASE 2: \( k_f = 1 \)}

The proof for this case is similar to the \( k_f > 1 \) case, with the following minor modifications:

\begin{enumerate}
\item Since \( k_f = 1 \), there is no plant stages downstream of \( k_f \), i.e., stages \( 1, 2, \ldots, k_f - 1 \) are missing in capacity matrices \( Q(\mu), q, \mu \) and \( \tilde{\mu} \), and in constraints such as (51);
\item Proof of Part (b) is similar to proof of Part (d) in Case 1, except that all \( q^i_{d, \text{down}} \) are replaced by \( d_i \).
\end{enumerate}

This completes the proof of Lemma 3.

With respect to the statements in Lemma 3, we would like to add the following:

- Limits of \( k_f \) for fixed \( k_v \):
  - Lemma 3 CASE 1 is for flexibility upstream of variability \( (k_v + 1 \leq k_f \leq K) \). As stated in Theorem 1 (1), when flexibility is upstream of variability, the optimal location of flexibility is at the stage closest to variability, i.e., \( k_f = k_v + 1 \), and all the other locations for flexibility are suboptimal. Lemma 3 describes property of structures that are suboptimal, i.e., \( k_f : k_v + 2 \leq k_f \leq K \).
  - Lemma 3 CASE 2 is for flexibility downstream of variability \( (1 \leq k_f \leq k_v) \). As stated in Theorem 1 (2), when flexibility is downstream of variability, the optimal location of flexibility is at the stage closest to variability, i.e., \( k_f = k_v \), and all the other locations for flexibility are suboptimal. Lemma 3 describes property of structures that are suboptimal, i.e., \( k_f : 1 \leq k_f \leq k_v - 1 \).
• The two optimal structures, namely $k_f = k_v + 1$ for flexibility upstream of variability, and $k_f = k_v$ for flexibility downstream of variability, are not included in the Lemma, because of the following:

- The proof of Theorem 1 does not need property of optimal structures, namely $k_f = k_v + 1$ for the case with flexibility upstream of variability, and $k_f = k_v$ for the case with flexibility downstream of variability.
- The optimal structures, namely $k_f = k_v + 1$ for the case with flexibility upstream of variability, and $k_f = k_v$ for the case with flexibility downstream of variability, have the similar property stated in Lemma 3. But including these structures in the lemma, while not needed, will lead to more subcases and thus complicate the statement of the lemma.

• Limits of $k_v$:

- Limits of $k_v$ is such that the corresponding $k_f$ does not exceed limits of the system, i.e., $1 \leq k_f \leq K$. For CASE 1, $k_v + 2 \leq k_f \leq K$ requires $k_v \leq K - 2$. And $k_v \geq 0$ is to make sure $k_v$ does not exceed the supply chain. For CASE 2, $1 \leq k_f \leq k_v - 1$ requires $k_v \geq 2$. And $k_v \leq K$ is to make sure $k_v$ does not exceed the supply chain.

**PROOF OF THEOREM 1, PART (1):**

In order to prove

$$V^*(m_L(k_f), D) \geq V^*(m_L(k_f + 1), D) \quad \forall k_f : k_v + 1 \leq k_f \leq K - 1,$$

by sample path method, it suffices to show that for each possible optimal $\mu$ for system with flexibility at stage $k_f + 1$, there exists at least one corresponding $\tilde{\mu}$ for system with flexibility at stage $k_f$, such that

$$V(m_L(k_f), D, Q(\tilde{\mu})) \geq V(m_L(k_f + 1), D, Q(\mu)),$$

which in turn requires

$$\pi(m_L(k_f), d, q) - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k \geq \pi(m_L(k_f + 1), d, q) - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k \quad \text{for all } d, q. \quad (55)$$

By “possible optimal” capacity we refer to capacity matrices that satisfy conditions in Lemma 3.

From Assumption 1, we have

$$m_{L,ij}(k) = m_{L,ij}(k + 1), \quad i, j = 1, 2, \ldots, I, \quad \forall k : 1 \leq k < K.$$

Therefore, to simplify notation we use $m_{L,ij} = m_{L,ij}(k) = m_{L,ij}(k + 1)$, and hence we have

$$m_L(k) = m_L \quad \forall k : 1 \leq k \leq K.$$

By Assumption 3, we write

$$c_i^k = c^k, \quad i = 1, 2, \ldots, I, \quad \forall k : 1 \leq k \leq K.$$

**CASE 1: $1 \leq k_v \leq K - 2$**

In this case, the capacity of plants at stage $k_v$ are sources of variability. By Lemma 3, in order for $\mu$ to be optimal when flexibility is located at stage $k_f + 1$, $\mu$ must have the form

$$\mu = \begin{pmatrix}
\begin{array}{cccccccc}
q_{1}^\text{down} & \cdots & q_{1}^\text{down} & \mu_{1}^{K_v} & q_{1}^\text{up} & \cdots & q_{1}^\text{up} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
q_{I}^\text{down} & \cdots & q_{I}^\text{down} & \mu_{I}^{K_v} & q_{I}^\text{up} & \cdots & q_{I}^\text{up}
\end{array}
\end{pmatrix} \quad (56)$$

29
with
\[ q_{i\text{down}} \leq d_i \quad \text{for} \quad i = 1, 2, \ldots, I, \quad \text{and} \quad \sum_{i=1}^{I} q_{i\text{up}} \leq \sum_{i=1}^{I} q_{i\text{down}}. \]

We claim that for all such \( \mu \) (i.e., certain \( q_{i\text{down}}, q_{i\text{up}} \) and \( \mu_{i\text{up}}^{k_v} \) for all \( i \)) in the structure with flexibility at stage \( k_f + 1 \), there exists a capacity matrix \( \tilde{\mu} \) (with element \( \tilde{\mu}_i^k \)) in the structure with flexibility at stage \( k_f \), where
\[
\tilde{\mu} = \begin{pmatrix}
    q_{1\text{down}}^{\text{stage 1}} & \cdots & q_{1\text{down}}^{\text{stage (k_f-1)}} & \mu_1 & q_{1\text{down}}^{\text{stage (k_f+1)}} & \cdots & q_{1\text{down}}^{\text{stage (k_f-1)}} & q_{1\text{up}} & \cdots & q_{1\text{up}}^{\text{stage K}} \\
    \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots \\
    q_{I\text{down}}^{\text{stage 1}} & \cdots & q_{I\text{down}}^{\text{stage (k_f-1)}} & \mu_I & q_{I\text{down}}^{\text{stage (k_f+1)}} & \cdots & q_{I\text{down}}^{\text{stage (k_f-1)}} & q_{I\text{up}} & \cdots & q_{I\text{up}}^{\text{stage K}}
\end{pmatrix}, \tag{57}
\]
and \( \tilde{\mu} \) results in a higher profit for any realization of capacity \( \mu_{i\text{up}}^{k_v} \) for all \( i \).

When flexibility is at stage \( k_f \), for any realization of capacity matrix \( Q(\mu) \), we have from Lemma 1,
\[
\pi(m_L, d, q) = \max_{Y_{ij}^{k_f}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} [m_{L,ij} \cdot Y_{ij}^{k_f}] \right\} \tag{58}
\]
subject to:
\[
\begin{align*}
    \sum_{j=1}^{I} Y_{ij}^{k_f} & \leq \min_{k_f \leq k \leq K} \left\{ \tilde{\mu}_i^k \right\} \quad i = 1, 2, \ldots, I, \tag{59} \\
    \sum_{j=1}^{I} Y_{ji}^{k_f} & \leq \min_{1 \leq k \leq k_f-1, k \neq k_v} \left\{ \mu_i^k, d_i, \tilde{q}_i^{k_v} \right\} \quad i = 1, 2, \ldots, I. \tag{60}
\end{align*}
\]

Using (57) for \( i = 1, 2, \ldots, I \), we have:
\[
\tilde{\mu}_i^k = \begin{cases} 
    q_{i\text{up}}^{k_f} & : k_f \leq k \leq K, \\
    q_{i\text{down}}^{k_f} & : 1 \leq k \leq k_f - 1, k \neq k_v,
\end{cases}
\]
then, constraints (59) and (60) becomes
\[
\begin{align*}
    \sum_{j=1}^{I} Y_{ij}^{k_f} & \leq q_{i\text{up}}^{k_f} \quad i = 1, 2, \ldots, I, \tag{61} \\
    \sum_{j=1}^{I} Y_{ji}^{k_f} & \leq \begin{cases} 
    \min \left\{ d_i, \tilde{q}_i^{k_v} \right\} & : \text{if } k_v = 1, k_f = 2, \\
    \min \left\{ q_{i\text{down}}^{k_f}, d_i, \tilde{q}_i^{k_v} \right\} & : \text{otherwise},
\end{cases} \quad i = 1, 2, \ldots, I. \tag{62}
\end{align*}
\]

When flexibility is at stage \( k_f + 1 \), for any realization of capacity matrix \( Q(\mu) \), we have
\[
\pi(m_L, d, q) = \max_{Y_{ij}^{k_f}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} [m_{L,ij} \cdot Y_{ij}^{k_f}] \right\} \tag{63}
\]
subject to:
\[
\begin{align*}
    \sum_{j=1}^{I} Y_{ij}^{k_f} & \leq \min_{k_f+1 \leq k \leq K} \left\{ \mu_i^k \right\} \quad i = 1, 2, \ldots, I, \tag{64} \\
    \sum_{j=1}^{I} Y_{ji}^{k_f} & \leq \min_{1 \leq k \leq k_f+1, k \neq k_v} \left\{ \mu_i^k, d_i, \tilde{q}_i^{k_v} \right\} \quad i = 1, 2, \ldots, I. \tag{65}
\end{align*}
\]
CASE 2: \( k_v = 0 \)

The proof for the \( k_v = 0 \) case is similar to the \( k_v \neq 0 \) case, except for the following minor modifications, given that now demand is the only source of variability and capacity is deterministic:

1. \( Q(\mu) = q = \mu; \)
2. Since \( D \) is stochastic and \( Q(\mu) \) is deterministic, to calculate \( V^* \), the expectation is taken over \( D \) instead of \( Q(\mu) \);
3. Stage \( k_v \) with uncertain capacity is missing in capacity matrices \( \mu \) and \( \tilde{\mu} \). And \( q_i^{k_v} \) is omitted in all the constraints regarding capacities, since demand is the only source of variability.

**PROOF OF THEOREM 1, PART (2):**

In order to prove

\[
V^*(m_L(k_f), D) \geq V^*(m_L(k_f - 1), D) \quad \forall \ k_f : 2 \leq k_f \leq k_v,
\]

Using (56), for \( i = 1, 2, \ldots, I \), we have:

\[
\mu_i^k = \begin{cases} 
q_i^{up} & : k_f + 1 \leq k \leq K, \\
q_i^{down} & : 1 \leq k \leq k_f, k \neq k_v,
\end{cases}
\]

then, constraints (64) and (65) becomes

\[
\sum_{j=1}^I Y_{ij}^{k_f} \leq q_i^{up} \quad i = 1, 2, \ldots, I, \tag{66}
\]

\[
\sum_{j=1}^I Y_{ij}^{k_f} \leq \min \{ d_i, q_i^{k_v} \} \quad i = 1, 2, \ldots, I. \tag{67}
\]

Compare constraints (61) and (62) with (66) and (67), the feasible region defined by constraints (61) and (62) is either larger than (when \( k_v = 1, k_f = 2 \)) or equivalent to (when \( k_v \neq 1, k_f \neq 2 \)) the feasible region defined by (66) and (67). Since the objective functions (58) and (63) have the same structure, therefore

\[
\varpi(m_L, d, Q(\tilde{\mu})) \geq \varpi(m_L, d, Q(\mu)) \quad \text{for all realization of } Q. \tag{68}
\]

With respect to the capacity investment cost,

\[
\sum_{k=1}^K \sum_{i=1}^I c_i^k \mu_i^k = \sum_{k=1, k \neq k_v}^{k_f-1} \sum_{i=1}^I c_i^k q_i^{down} + \sum_{k=k_f}^K \sum_{i=1}^I c_i^k q_i^{up} + \sum_{i=1}^I c_i^{k_v} \mu_i^{k_v},
\]

\[
K \sum_{k=1}^I c_i^k \mu_i^k = \sum_{k=1, k \neq k_v}^{k_f-1} \sum_{i=1}^I c_i^k q_i^{down} + \sum_{k=k_f+1}^K \sum_{i=1}^I c_i^k q_i^{up} + \sum_{i=1}^I c_i^{k_v} \mu_i^{k_v}.
\]

And therefore

\[
\sum_{k=1}^K \sum_{i=1}^I c_i^k \mu_i^k - \sum_{k=1}^K \sum_{i=1}^I c_i^k \mu_i^k = \sum_{i=1}^I c_i^{k_f} q_i^{up} - \sum_{i=1}^I c_i^{k_f} q_i^{down} = e^{k_f} \cdot \left( \sum_{i=1}^I q_i^{up} - \sum_{i=1}^I q_i^{down} \right) \leq 0. \tag{69}
\]

Considering (68) and (69), it becomes clear that capacity matrix \( \tilde{\mu} \) has a lower capacity investment and results in a higher profit than \( \mu \) for any realization of \( Q_i^{k_v} \). Therefore, we have proven (55) and hence

\[
V^*(m_L(k_f), D) \geq V^*(m_L(k_f + 1), D) \quad \forall \ k_f : k_v + 1 \leq k_f \leq K - 1. \tag{70}
\]
by sample path method, it suffices to show that for each possible optimal \( \mu \) for system with flexibility at stage \( k_f - 1 \), there exists at least one corresponding \( \tilde{\mu} \), such that

\[
V(m_L(k_f), D, Q(\tilde{\mu})) \geq V(m_L(k_f - 1), D, Q(\mu)),
\]

which in turn requires

\[
\pi(m_L(k_f), d, q) - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \tilde{\mu}_i^k \geq \pi(m_L(k_f - 1), d, q) - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k \quad \text{for all } d, q.
\]

By “possible optimal” capacity we refer to capacity matrices that satisfy conditions in Lemma 3. From Assumption 1, we have

\[
m_{L, ij}(k) = m_{L, ij}(k - 1), \quad i, j = 1, 2, \ldots, I, \quad \forall k : 1 < k \leq K.
\]

Therefore, to simplify notation we use \( m_{L, ij} = m_{L, ij}(k) = m_{L, ij}(k - 1) \), and hence we have

\[
m_L(k) = m_L \quad \forall k : 1 \leq k \leq K.
\]

By Assumption 3, we write

\[
c_i^k = c_i, \quad i = 1, 2, \ldots, I, \quad \forall k : 1 \leq k \leq K.
\]

**CASE 1: 2 < k_f \leq k_u**

In this case, capacity of plants at stage \( k_u \) are sources of variability. By Lemma 3, in order for \( \mu \) to be optimal when logistics flexibility is located at stage \( k_f - 1 \), \( \mu \) must have the form

\[
\mu = \begin{pmatrix}
q_{1, 1}^{\text{down}} & \ldots & q_{1, I}^{\text{down}} & q_{1, 1}^{\text{up}} & \ldots & q_{1, I}^{\text{up}} & \mu_1^k & \mu_2^k & \ldots & \mu_I^k & q_{I, 1}^{\text{up}} & \ldots & q_{I, I}^{\text{up}}
\end{pmatrix},
\]

with

\[
q_{i, I}^{\text{down}} \leq d_i \quad i = 1, 2, \ldots, I, \quad \text{and} \quad \sum_{i=1}^{I} q_{i, I}^{\text{down}} \geq \sum_{i=1}^{I} q_{i, I}^{\text{up}}.
\]

We claim that for all such \( \mu \) in the structure with flexibility at stage \( k_f - 1 \), there exists a capacity matrix \( \tilde{\mu} \) in the structure with flexibility at stage \( k_f \), where

\[
\tilde{\mu} = \begin{pmatrix}
\tilde{q}_{1, 1}^{\text{down}} & \ldots & \tilde{q}_{1, I}^{\text{down}} & \tilde{q}_{1, 1}^{\text{up}} & \ldots & \tilde{q}_{1, I}^{\text{up}} & \tilde{\mu}_1^k & \tilde{\mu}_2^k & \ldots & \tilde{\mu}_I^k & \tilde{q}_{I, 1}^{\text{up}} & \ldots & \tilde{q}_{I, I}^{\text{up}}
\end{pmatrix},
\]

and \( \tilde{\mu} \) results in a higher profit for any realization of capacity \( Q_i^{k_r} \) for all \( i \).

When flexibility is at stage \( k_f \), for any realization of capacity matrix \( Q(\tilde{\mu}) \), we have

\[
\pi(m_L, d, q) = \max_{Y_{ij}^{k_f}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} \left[m_{L, ij} \cdot Y_{ij}^{k_f} \right] \right\}
\]

subject to:

\[
\sum_{j=1}^{I} Y_{ij}^{k_f} \leq \min_{k_f \leq k \leq K, k \neq k_u} \{\tilde{\mu}_i^k, q_i^{k_r}\} \quad i = 1, 2, \ldots, I, \quad (74)
\]

\[
\sum_{j=1}^{I} Y_{ji}^{k_f} \leq \min_{1 \leq k \leq k_f - 1} \{\tilde{\mu}_i^k, d_i\} \quad i = 1, 2, \ldots, I. \quad (75)
\]
Based on (73), we have for \( i = 1, 2, \ldots, I \):

\[
\tilde{\mu}^k_i = \begin{cases} 
q_{i}^{up} & : k_f \leq k \leq K, k \neq k_v, \\
q_{i}^{down} & : 1 \leq k \leq k_f - 1, \\
\mu^k_i & : k = k_v,
\end{cases}
\]

and constraints (74) and (75) become

\[
\sum_{j=1}^{I} Y_{ij}^{k_f} \leq \begin{cases} 
q_{i}^{up} \cdot k_v & : \text{if } k_v = k_f = K \\
\min \{q_{i}^{up}, q_{i}^{down}\} & : \text{otherwise}
\end{cases}
\]

\( i = 1, 2, \ldots, I, \) (76)

\[
\sum_{j=1}^{I} Y_{ji}^{k_f} \leq q_{i}^{down} \quad i = 1, 2, \ldots, I.
\]

(77)

When flexibility is at stage \( k_f - 1 \), for any realization of capacity matrix \( \mathbf{Q}(\mu) \), we have

\[
\pi(\mathbf{m}_L, \mathbf{d}, \mathbf{q}) = \max_y \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} \left[ m_{L,ij} \cdot Y_{ij}^{k_f} \right] \right\}
\]

subject to:

\[
\sum_{j=1}^{I} Y_{ij}^{k_f} \leq \min_{k_f - 1 \leq k \leq K, k \neq k_v} \{ \mu^k_i, q_{i}^{k_v} \} \quad i = 1, 2, \ldots, I,
\]

(78)

\[
\sum_{j=1}^{I} Y_{ji}^{k_f} \leq \min_{1 \leq k \leq k_f - 2} \{ \mu^k_i, d_i \} \quad i = 1, 2, \ldots, I.
\]

(79)

Using (72) for \( i = 1, 2, \ldots, I \) we have:

\[
\mu^k_i = \begin{cases} 
q_{i}^{up} & : k_f - 1 \leq k \leq K, k \neq k_v, \\
q_{i}^{down} & : 1 \leq k \leq k_f - 2,
\end{cases}
\]

therefore, constraints (78) and (79) become

\[
\sum_{j=1}^{I} Y_{ij}^{k_f} \leq \min \{q_{i}^{up}, q_{i}^{k_v}\} \quad i = 1, 2, \ldots, I,
\]

(80)

\[
\sum_{j=1}^{I} Y_{ji}^{k_f} \leq q_{i}^{down} \quad i = 1, 2, \ldots, I.
\]

(81)

Comparing constraints (76) and (77) with (80) and (81), we see that the feasible region defined by constraints (76) and (77) is either larger than or equivalent to the feasible region defined by (80) and (81). Therefore

\[
\varpi(\mathbf{m}_L, \mathbf{d}, \mathbf{Q}(\tilde{\mu})) \geq \varpi(\mathbf{m}_L, \mathbf{d}, \mathbf{Q}(\mu)) \quad \text{for all realization of } \mathbf{Q}.
\]

(82)

With respect to the capacity investment cost,

\[
K \sum_{k=1}^{k_f} \sum_{i=1}^{I} c^k_{i} \mu^k_i = \sum_{k=k_f-1}^{k_f-2} \sum_{i=1}^{I} c^k_{i} \mu^k_i + \sum_{k=k_f}^{K} \sum_{i=1}^{I} c^k_{i} \mu^k_i + \sum_{i=1}^{I} c^k_{i} \mu^k_i,
\]

\[
K \sum_{k=1}^{k_f} \sum_{i=1}^{I} c^k_{i} \mu^k_i = \sum_{k=k_f-1}^{k_f-2} \sum_{i=1}^{I} c^k_{i} \mu^k_i + \sum_{k=k_f}^{K} \sum_{i=1}^{I} c^k_{i} \mu^k_i + \sum_{i=1}^{I} c^k_{i} \mu^k_i.
\]
And therefore
\[
\sum_{k=1}^{K} \sum_{i=1}^{l} c_i^k \mu_i^k - \sum_{k=1}^{K} \sum_{i=1}^{l} c_i^k \mu_i^k = \sum_{i=1}^{l} c_i^{k-1} q_i^{den} - \sum_{i=1}^{l} c_i^{k-1} q_i^{up} = c^{k-1} \left[ \sum_{i=1}^{l} q_i^{den} - \sum_{i=1}^{l} q_i^{up} \right] \leq 0. \tag{83}
\]

Considering (82) and (83), it becomes clear that capacity matrix \( \tilde{\mu} \) has a lower capacity investment and results in a higher profit than \( \mu \) for any realization of \( Q_i^{k_v} \). Therefore, we have proven (71) and hence
\[
V^* (m_L(k_f), D) \geq V^* (m_L(k_f - 1), D) \quad \forall \ k_f : 2 < k_f \leq k_v. \tag{84}
\]

**CASE 2: \( k_f = 2 \)**

Proof for this case is similar to the \( 2 < k_f \leq k_v \) case. But now in order for \( \mu \) to be optimal when flexibility is at stage 1, by Lemma 3 Case(2): \( k_f = 1 \), it must have the form
\[
\begin{pmatrix}
\begin{array}{ccccccc}
\text{stage 1} & \cdots & \text{stage } (k_v - 1) & \text{stage } k_v & \text{stage } (k_v + 1) & \cdots & \text{stage } K \\
q_{1}^{up} & \cdots & q_{1}^{up} & \mu_i^{k_v} & q_{1}^{up} & \cdots & q_{1}^{up} \\
\vdots & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots \\
q_{l}^{up} & \cdots & q_{l}^{up} & \mu_i^{k_v} & q_{l}^{up} & \cdots & q_{l}^{up}
\end{array}
\end{pmatrix}
\tag{85}
\]

with
\[
\sum_{i=1}^{l} q_i^{up} \geq \sum_{i=1}^{l} d_i.
\]

We claim that for all such \( \mu \), there exists a capacity matrix \( \tilde{\mu} \) in the structure with flexibility at stage 2, where
\[
\begin{pmatrix}
\begin{array}{ccccccc}
\text{stage 1} & \text{stage 2} & \cdots & \text{stage } (k_v - 1) & \text{stage } k_v & \text{stage } (k_v + 1) & \cdots & \text{stage } K \\
d_1 & q_{1}^{up} & \cdots & q_{1}^{up} & \mu_i^{k_v} & q_{1}^{up} & \cdots & q_{1}^{up} \\
\vdots & \ddots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
d_l & q_{l}^{up} & \cdots & q_{l}^{up} & \mu_i^{k_v} & q_{l}^{up} & \cdots & q_{l}^{up}
\end{array}
\end{pmatrix}
\tag{86}
\]

and \( \tilde{\mu} \) results in a higher profit for any realization of capacity \( Q_i^{k_v} \) for all \( i \).

All arguments for the \( 2 < k_f \leq k_v \) case follows by replacing \( q_i^{den} \) with \( d_i \).

Before we present the proof for Theorem 2, we first need to present Lemma 4. Note that Lemma 4 holds when Assumptions 1 is relaxed.

**Lemma 4.** Suppose a supply chain has variability only at stage \( k_v \) (\( 0 \leq k_v \leq K \)) and process flexibility only at stage \( k_f \) (\( 1 \leq k_f \leq K \)). Let \( \mu = \left( \mu_i^k \right) \) be the optimal capacity investment matrix for the supply chain. Then the optimal capacity matrix \( \mu \) satisfies the following:

(a) the optimal capacity configuration is the same for all stages except for the flexible and variable stage (i.e., there exist values \( q_i \) such that \( q_i = \mu_i^k \), \( \forall k \neq k_f, k_v, \ k = 1, 2, \ldots, K, \ i = 1, 2, \ldots, I \);

(b) \( q_i \leq d_i \), for \( 1 \leq k_v \leq K, \ i = 1, 2, \ldots, I \);

(c) \( \mu_i^{k_f} \leq q_i \), for \( i = 1, 2, \ldots, I \).
PROOF OF LEMMA 4:

We prove the lemma by proving the following two cases, namely, $1 \leq k_v \leq K$ and $k_v = 0$.

CASE 1: $1 \leq k_v \leq K$

In order to prove the optimal capacity matrix $\boldsymbol{\mu}$ satisfies conditions (a) to (c), we show that any $\boldsymbol{\mu}$ that does not satisfy any of the three conditions cannot be optimal, i.e., there exists a vector $\tilde{\boldsymbol{\mu}}$, which satisfies these conditions, and achieves higher total expected profit than $\boldsymbol{\mu}$.

Since the only variability is at stage $k_v$ ($1 \leq k_v \leq K$), for all $k$’s such that $k \neq k_v$, we set $Q^k_i(\mu^k_v) = \mu^k_v$. Therefore, we have

$$Q(\mu) = \begin{pmatrix} \mu_1 & \cdots & \mu_{k_v-1}^{k_v} & Q_1^{k_v} & \mu_{k_v+1}^{k_v} & \cdots & \mu_K^{k_v} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_1 & \cdots & \mu_{k_v-1}^{k_v} & Q_1^{k_v} & \mu_{k_v+1}^{k_v} & \cdots & \mu_K^{k_v} \end{pmatrix}.$$ 

$q = \begin{pmatrix} \mu_1 & \cdots & \mu_{k_v-1}^{k_v} & q_1^{k_v} & \mu_{k_v+1}^{k_v} & \cdots & \mu_K^{k_v} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_1 & \cdots & \mu_{k_v-1}^{k_v} & q_1^{k_v} & \mu_{k_v+1}^{k_v} & \cdots & \mu_K^{k_v} \end{pmatrix}.$

And since demand is deterministic,

$$\mathbf{D} = (D_1, D_2, \ldots, D_I) = (d_1, d_2, \ldots, d_I) = \mathbf{d}.$$ 

For full process flexibility configuration at stage $k_f$, by Lemma 2, the optimal capacity matrix $\boldsymbol{\mu}$ is the solution to $\mathbf{P1}(\mathbf{m}_P(k_f), \mathbf{d})$:

$$V^*(\mathbf{m}_P(k_f), \mathbf{d}) = \max_{\boldsymbol{\mu}} \left\{ V(\mathbf{m}_P(k_f), \mathbf{d}, Q(\boldsymbol{\mu})) \right\}$$

$$= \max_{\boldsymbol{\mu}} \left\{ \mathbb{E}_{Q(\boldsymbol{\mu})} [\pi(\mathbf{m}_P(k_f), \mathbf{d}, Q(\boldsymbol{\mu}))] - \sum_{k=1}^{K} \sum_{i=1}^{I} c^k_i \mu^k_i \right\},$$

where $\pi(\mathbf{m}_P(k_f), \mathbf{d}, Q(\boldsymbol{\mu}))$ is calculated by solving $\mathbf{P2}(\mathbf{m}_P(k_f), \mathbf{d}, \mathbf{q})$ and taking expectation over $Q(\boldsymbol{\mu})$:

$$\pi(\mathbf{m}_P(k_f), \mathbf{d}, \mathbf{q}) = \max_{Z_{ij}^k} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{J} \left[ m_{P,ij}(k_f) \cdot Z_{ij}^k \right] \right\}$$

subject to:

$$\sum_{j=1}^{J} Z_{ij}^k \leq \min_{1 \leq k \leq K} \left\{ \mu^k_i, q^k_i, d_i \right\} \quad i = 1, 2, \ldots, I,$$

$$\sum_{j=1}^{J} Z_{ij}^k \leq \mu^k_i \quad i = 1, 2, \ldots, I,$$

where

$$m_{P,ij}(k_f) = r_i - \left[ \left( \sum_{k=1}^{K} p^k_{ij} \right) + p^k_{ij} \right] - \left[ \left( \sum_{k=1}^{K} t^k_{ij} \right) + t^k_{ij} + t^{k+1}_{ij} \right].$$
CASE 1 – Part (a):

Assume that \( \mu \) does not satisfy condition (a). Let, for \( i = 1, 2, \ldots, I \),

\[
q_{\min, i} = \min_{k \neq k_f, k_v} \{ \mu_k^i \}.
\]

Now consider the capacity matrix \( \tilde{\mu} \) as:

\[
\tilde{\mu} = \begin{pmatrix}
q_{\min,1} & \cdots & q_{\min,1} & \mu_1^k & q_{\min,1} & \cdots & q_{\min,1} & \mu_1^k & \cdots & q_{\min,1} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_{\min,I} & \cdots & q_{\min,I} & \mu_I^k & q_{\min,I} & \cdots & q_{\min,I} & \mu_I^k & \cdots & q_{\min,I}
\end{pmatrix}.
\]

If \( \mu \) does not satisfy condition (a), then

\[
\mu_i^k > q_{\min,i} \quad \text{for some } i \ (1 \leq i \leq I) \text{ and some } k \ (1 \leq k \leq K, k \neq k_f, k_v).
\]

So we have

\[
\sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k < \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \tilde{\mu}_i^k.
\]

which implies that capacity matrix \( \tilde{\mu} \) has a lower capacity investment cost than that of \( \mu \). On the other hand, since

\[
\min_{1 \leq k \leq K} \{ \mu_i^k \} = q_{\min,i} = \min_{1 \leq k \leq K} \{ \mu_i^k \} \quad i = 1, 2, \ldots, I,
\]

Then, the feasible region for problem P2 (i.e., constraints (88) and (89)) are the same for capacity matrices \( \mu \) and \( \tilde{\mu} \). On the other hand, since the objective function of problem P2, i.e., (87), is independent of capacity matrices; hence, we have

\[
\varpi(\mathbf{m}_P(k_f), \mathbf{d}, Q(\tilde{\mu})) = \varpi(\mathbf{m}_P(k_f), \mathbf{d}, Q(\mu)) \quad \text{for all realization of } Q.
\]

Therefore, since \( \tilde{\mu} \) has a lower capacity investment, it achieves higher expected net profit than \( \mu \), and hence \( \mu \) cannot be optimal solution to P1(\( \mathbf{m}_P(k_f), \mathbf{d} \)). Thus, the optimal capacity matrix \( \mu \) must follow condition (a).

CASE 1 – Part (b):

To prove part (b), consider

\[
\mu = \begin{pmatrix}
q_1 & \cdots & q_1 & \mu_1^k & q_1 & \cdots & q_1 & \mu_1^k & \cdots & q_1 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_I & \cdots & q_I & \mu_I^k & q_I & \cdots & q_I & \mu_I^k & \cdots & q_I
\end{pmatrix},
\]

which satisfies condition (a), but does not satisfy (b), i.e., \( q_{i'} > d_{i'} \) for some \( i' \ (1 \leq i' \leq I) \). Let

\[
\tilde{\mu} = \begin{pmatrix}
q_1 & \cdots & q_1 & \mu_1^k & q_1 & \cdots & q_1 & \mu_1^k & \cdots & q_1 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
d_{i'} & \cdots & d_{i'} & \mu_{i'}^k & d_{i'} & \cdots & d_{i'} & \mu_{i'}^k & \cdots & d_{i'} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_I & \cdots & q_I & \mu_I^k & q_I & \cdots & q_I & \mu_I^k & \cdots & q_I
\end{pmatrix}.
\]

36
that is, substituting all \( q_i \) in \( \mu \) with \( d_i \). Since \( q_i > d_i \) for \( i' \) \((1 \leq i' \leq I)\), we have
\[
\sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k < \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \tilde{\mu}_i^k.
\]

On the other hand, the only difference between capacity matrices \( \mu \) and \( \tilde{\mu} \) is in plant \( i' \) at all stages except for stages \( k_f \) and \( k_v \), and for these capacities, we have
\[
\min_{1 \leq k \leq K} \{ \mu_i^k, d_i \} = d_i = \min_{1 \leq k \leq K} \{ \tilde{\mu}_i^k, d_i \}.
\]

Hence, the feasible region, i.e., constraints (88) and (89), are the same for capacity matrices \( \mu \) and \( \tilde{\mu} \). So, similar to Part (a), we have
\[
\varpi(\mathbf{m}_P(k_f), \mathbf{d}, Q(\tilde{\mu})) = \varpi(\mathbf{m}_P(k_f), \mathbf{d}, Q(\mu)) \quad \text{for all realization of } Q.
\]

Therefore, \( \tilde{\mu} \) achieves higher expected net profit than \( \mu \), and hence \( \mu \) cannot be optimal solution to \( \mathbf{P1}(\mathbf{m}_P(k_f), \mathbf{d}) \). Thus, the capacity matrix that does not satisfy condition (b) cannot be optimal.

**CASE 1 – Part (c):**

First we need to show that \( \sum_{i=1}^{I} k_i^{k_f} \leq \sum_{i=1}^{I} q_i \). Suppose

\[
\mu = \begin{pmatrix}
\begin{array}{cccccccc}
\text{stage 1} & \cdots & \text{stage } (k_v - 1) & \text{stage } k_v & \text{stage } (k_v + 1) & \cdots & \text{stage } (k_f - 1) & \text{stage } k_f & \cdots & \text{stage } K \\
q_1 & \cdots & q_1 & \mu_1^{k_v} & q_1 & \cdots & q_1 & \mu_1^{k_f} & \cdots & q_1 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_I & \cdots & q_I & \mu_I^{k_v} & q_I & \cdots & q_I & \mu_I^{k_f} & \cdots & q_I 
\end{array}
\end{pmatrix}
\]

satisfies (a) and (b), but with \( \sum_{i=1}^{I} k_i^{k_f} > \sum_{i=1}^{I} q_i \). Then

\[
\tilde{\mu} = \begin{pmatrix}
\begin{array}{cccccccc}
\text{stage 1} & \cdots & \text{stage } (k_v - 1) & \text{stage } k_v & \text{stage } (k_v + 1) & \cdots & \text{stage } (k_f - 1) & \text{stage } k_f & \cdots & \text{stage } K \\
q_1 & \cdots & q_1 & \mu_1^{k_v} & q_1 & \cdots & q_1 & \mu_1^{k_f} & \cdots & q_1 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_I & \cdots & q_I & \mu_I^{k_v} & q_I & \cdots & q_I & \mu_I^{k_f} & \cdots & q_I 
\end{array}
\end{pmatrix}
\]

has a lower capacity investment cost, as
\[
\sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \mu_i^k - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \tilde{\mu}_i^k = \sum_{k=1}^{K} \sum_{i=1}^{I} (c_i^k (\mu_i^k - q_i) > 0,
\]

since \( c_i^k = c_i^{k_f} \), for \( i, j = 1, 2, \ldots, I, i \neq j \), and \( k = 1, 2, \ldots, K \).

For capacity matrix \( \tilde{\mu} \), \( \varpi(\mathbf{m}_P(k_f), \mathbf{d}, Q(\tilde{\mu})) \) is calculated by solving \( \mathbf{P2}(\mathbf{m}_P(k_f), \mathbf{d}, \mathbf{q}) \) and taking expectation over \( Q(\tilde{\mu}) \):

\[
\pi(\mathbf{m}_P(k_f), \mathbf{d}, \mathbf{q}) = \max_{Z^{k_f}_{ij}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} m_{P,ij}^{(k_f)} \cdot Z^{k_f}_{ij} \right\}
\]

subject to:
\[(91)\]
\[
\sum_{j=1}^{I} Z^{k_f}_{ij} \leq \min_{1 \leq k \leq K} \{ \mu_i^k, q_k^{k_v} \} = \min \{ q_i, q_i^{k_v} \} \quad \text{(since (a) and (b) hold) } i = 1, 2, \ldots, I.
\]

\[
\sum_{j=1}^{I} Z^{k_f}_{ij} \leq \mu_i^{k_f} = q_i \quad i = 1, 2, \ldots, I.
\]
Since according to Assumption 3,  
\[ m_{P,i}(k_f) > m_{P,i}(k_f) \quad \text{and} \quad m_{P,i}(k_f) > m_{P,ji}(k_f) \quad \forall \ j \neq i, \quad i = 1, 2, \ldots, I, \]
the most profitable way of allocating flows (\(Z_{ij}^k\)'s) is to use dedicated plants. By comparing right-hand-side of (91) and (92) we have  
\[ q_i \geq \min \{ \tilde{q}_i, q_i^{k_v} \}. \]
Therefore, we get the maximum gross profit for \( \tilde{\mu} \) by setting  
\[ Z_{ii} = \min \{ q_i, q_i^{k_v} \} \] and hence  
\[ \sum Z_{ij}^k = 0, \quad \forall \ i \neq j; \]
By (90) and (93),  
\[ \pi(m_P(k_f), d, q) = \sum_{i=1}^{I} \left[ m_{P,i}(k_f) \cdot \min \{ q_i, q_i^{k_v} \} \right]. \]
On the other hand, for capacity investment vector \( \mu, \varpi(m_P(k_f), d, Q(\mu)) \) is calculated by solving \( P2(m_P(k_f), d, q) \) and taking expectation over \( Q(\mu) \):
\[
\pi(m_P(k_f), d, q) = \max_{Z_{ij}^k} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} \left[ m_{P,ij}(k_f) \cdot Z_{ij}^k \right] \right\}
\]
subject to:
\[
\sum_{j=1}^{I} Z_{ij}^k \leq \min \{ q_i, q_i^{k_v} \} \quad i = 1, 2, \ldots, I,
\]
\[
\sum_{j=1}^{I} Z_{ji}^k \leq \mu_i^{k_f} \quad i = 1, 2, \ldots, I.
\]
Hence
\[
\sum_{i=1}^{I} \sum_{j=1}^{I} m_{P,ij}(k_f) \cdot Z_{ij}^k \leq \sum_{i=1}^{I} \sum_{j=1}^{I} m_{P,ji}(k_f) \cdot Z_{ji}^k \leq \sum_{i=1}^{I} \left[ m_{P,ii}(k_f) \sum_{j=1}^{I} Z_{ji}^k \right] \leq \sum_{i=1}^{I} \left[ m_{P,ii}(k_f) \cdot \min \{ q_i, q_i^{k_v} \} \right]
\]
Therefore
\[
\varpi(m_P(k_f), d, Q(\mu)) \leq \varpi(m_P(k_f), d, Q(\tilde{\mu})) \quad \text{for all realization of } Q \quad \text{(93)}
\]
By (90) and (93),  
\[ V(m_P(k_f), d, Q(\mu)) < V(m_P(k_f), d, Q(\tilde{\mu})), \]
and hence \( \mu \) cannot be optimal solution to \( P1(m_P(k_f), d) \). Thus, the optimal capacity matrix \( \mu \) must follow  
\[ \sum_{i=1}^{I} \mu_i^{k_f} \leq \sum_{i=1}^{I} q_i. \]
Now we show that \( \mu_i^{k_f} \leq q_i \) for all \( i = 1, 2, \ldots, I \). Suppose \( \mu_i^{k_f} > q_i \) for some \( i \), since \( \sum_{i=1}^{I} \mu_i^{k_f} \leq \sum_{i=1}^{I} q_i \), there has to be some \( i' \), such that \( \mu_{i'}^{k_f} < q_{i'} \). Without loss of generality, we assume that plant \( i'' \) is the only plant at stage \( k_f \) with such property. Let us design \( \tilde{\mu} \) such that \( \tilde{\mu}_{i'}^{k_f} = q_{i'} \) and \( \tilde{\mu}_{i''}^{k_f} = \mu_{i''}^{k_f} + \mu_{i''}^{k_f} - q_{i'} \), that is, at stage \( k_f, \mu_i^{k_f} - q_i \) units of capacity is moved to plant \( i'' \). From Assumption 4, we have  
\[ c_i^{k_f} = c_i^{k_f}, \forall i, j = 1, 2, \ldots, I, \quad i \neq j, \forall k = 1, 2, \ldots, K. \] So \( \tilde{\mu} \) and \( \mu \) have same capacity investment cost. Only now \( \mu_i^{k_f} - q_i \) units of capacity can be used through dedicated arc from plant \( i'' \), rather than flexible arcs from plant \( i' \). Since Assumption 3 assures that dedicated arc is more profitable than flexible ones, \( \tilde{\mu} \) gives higher revenue than
\( \mu \). Hence \( \mu \) cannot be optimal. Therefore the optimal capacity investment matrix has property \( \mu_{i}^{k} \leq q_{i} \) for all \( i = 1, 2, \ldots, I \).

**CASE 2: \( k_{v} = 0 \)**

The proof for the \( k_{v} = 0 \) case is similar to that for the \( k_{v} \neq 0 \) case, except for the following minor modifications. Given that now demand is the only source of variability and capacity is deterministic:

1. \( Q(\mu) = q = \mu \);
2. Since \( D \) is stochastic and \( Q(\mu) \) is deterministic, to calculate \( V^{*} \), the expectation is taken over \( D \) instead of \( Q(\mu) \);
3. Stage \( k_{v} \) with uncertain capacity is missing in capacity matrices \( Q(\mu), q, \mu \) and \( \tilde{\mu} \), and in constraints such as (88);
4. We have \( d_{i} \) instead of \( q_{i}^{k_{v}} \) in constraint (91) of Part (c) and the arguments following the constraint.

**PROOF OF THEOREM 2:**

From Assumption 1, we have \( m_{P,ij}(k) = m_{P,ij}(k') \), \( \forall i, j, \forall k \neq k' \). To simplify notation we use \( m_{P,ij} = m_{P,ij}(k) \), and hence we have

\[
m_{P}(k) = m_{P} \quad \forall k : 1 \leq k \leq K.
\]

**PART (1) - Variability in Demand:**

To prove Part (1), we show that when demand is variable, the total expected profit of all configurations in which only one stage has process flexibility are the same. Consider two flexibility configurations \( (k_{f} \neq k'_{f}) \):

(I) Full process flexibility at stage \( k_{f} \),

(II) Full process flexibility at stage \( k'_{f} \),

we show

\[
V^{*}(m_{P}(k_{f}),D) = V^{*}(m_{P}(k'_{f}),D) \quad \forall 1 \leq k_{f}, k'_{f} \leq K.
\]

Let \( \mu \) be a candidate for the optimal capacity investment matrix for configuration (I). By Lemma 4,

\[
\mu = \begin{pmatrix}
q_{1} & \cdots & q_{1} & \mu_{1}^{k_{f}} & q_{1} & \cdots & q_{1} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
q_{I} & \cdots & q_{I} & \mu_{I}^{k_{f}} & q_{I} & \cdots & q_{I}
\end{pmatrix},
\]

and

\[
V(m_{P}(k_{f}),D,Q(\mu)) = \mathbb{E}_{D,Q(\mu)}[\pi(m_{P}(k_{f}),D,Q(\mu))] - K \sum_{k=1}^{I} \sum_{i=1}^{I} c_{i}^{k} \mu_{i}^{k}
\]

\[
= \mathbb{E}_{D}[\pi(m_{P},D,q(\mu))] - \sum_{k=1}^{I} \sum_{i=1}^{I} c_{i}^{k} \mu_{i}^{k}
\]

\[
= \mathbb{E}_{D}[\pi(m_{P},D,q(\mu))] - \sum_{i=1}^{I} (K-1)q_{i} + \mu_{i}^{k_{f}}, \quad (94)
\]
where \( \mathbb{E}_D \left[ \varpi(m_P, D, q(\mu)) \right] \) is calculated by solving the following problem \( \text{P2}(m_P, d, q) \) and taking expectation over \( D \):

\[
\pi(m_P, d, q) = \max_{Z_{ij}^{k_j}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} \left[ m_{P,ij} \cdot Z_{ij}^{k_j} \right] \right\}
\]

subject to:

\[
\sum_{j=1}^{I} Z_{ij}^{k_j} \leq \min_{1 \leq k \leq K} \{ \mu_i^k, d_i \} = \min \{ q_i, d_i \} \quad i = 1, 2, \ldots, I,
\]

\[
\sum_{j=1}^{I} Z_{ji}^{k_j} \leq \tilde{\mu}_i^{k_j} \quad i = 1, 2, \ldots, I.
\]

Note that by Assumption 5, we write \( c_i^k = c_i \), \( i = 1, 2, \ldots, I \), \( \forall k : 1 \leq k \leq K \).

For configuration (II), build a capacity matrix \( \tilde{\mu} \) using elements of \( \mu \) as follows:

\[
\tilde{\mu} = \begin{pmatrix}
q_1 & \cdots & q_1 & \mu_{1}^{k_f} & q_1 & \cdots & q_1 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
q_I & \cdots & q_I & \mu_{I}^{k_f} & q_I & \cdots & q_I
\end{pmatrix}
\]

we have

\[
V(m_P(k_f'), D, Q(\tilde{\mu})) = \mathbb{E}_{D, Q}[\varpi(m_P(k_f'), D, Q(\tilde{\mu}))] = - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i^k \tilde{\mu}_i^k
\]

\[
= \mathbb{E}_D \left[ \varpi(m_P, D, q(\tilde{\mu})) \right] - \sum_{i=1}^{I} \sum_{k=1}^{K} c_i^k \tilde{\mu}_i^k
\]

\[
= \mathbb{E}_D \left[ \varpi(m_P, D, q(\tilde{\mu})) \right] - \sum_{i=1}^{I} c_i [(K - 1)q_i + \mu_i^{k_f'}],
\]

(96)

where \( \mathbb{E}_D \left[ \varpi(m_P, D, q(\tilde{\mu})) \right] \) is calculated by solving \( \text{P2}(m_P, d, q) \) and taking expectation over \( D \):

\[
\pi(m_P, d, q) = \max_{Z_{ij}^{k_j}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} \left[ m_{P,ij} \cdot Z_{ij}^{k_j} \right] \right\}
\]

(97)

subject to:

\[
\sum_{j=1}^{I} Z_{ij}^{k_j} \leq \min_{1 \leq k \leq K} \{ \mu_i^k, d_i \} = \min \{ q_i, d_i \} \quad i = 1, 2, \ldots, I,
\]

\[
\sum_{j=1}^{I} Z_{ji}^{k_j} \leq \tilde{\mu}_i^{k_j} \quad i = 1, 2, \ldots, I.
\]

Comparing (94) and (95) with (96) and (97), we conclude that for all candidate optimal capacity investment matrix \( \mu \) for configuration (I), there exists a corresponding capacity matrix \( \tilde{\mu} \) for configuration (II), such that \( V(m_P(k_f), D, Q(\mu)) = V(m_P(k_f'), D, Q(\tilde{\mu})) \) for all realization of \( D \). And this is true for the other direction as well. Therefore we have

\[
V^*(m_P(k_f), D) = V^*(m_P(k_f'), D) \quad \forall 1 \leq k_f, k_f' \leq K.
\]

(98)
In other words, configurations (I) and (II) result in the same total expected profit. This completes the proof.

**PART (2) - Variability at Stage $k_f$:**

First we prove that investing in process flexibility at stage $k_f$ is optimal. We consider the following two configurations:

(I) Full process flexibility at stage $k_f$;

(II) Full process flexibility at stage $k_v \neq k_f$.

Let $\mu$ be a candidate for the optimal capacity investment matrix for configuration (II), assuming $k_v < k_f$ without loss of generality. By Lemma 4,

$$
\begin{align*}
\mu = \begin{pmatrix}
\begin{array}{cccccccc}
\text{stage 1} & \cdots & \text{stage (k_v-1)} & \text{stage k_v} & \text{stage (k_v+1)} & \cdots & \text{stage (k_f-1)} & \text{stage k_f} & \cdots & \text{stage K} \\
q_1 & \cdots & q_1 & \mu_1^{k_v} & q_1 & \cdots & q_1 & \mu_1^{k_f} & \cdots & q_1 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_I & \cdots & q_I & \mu_I^{k_v} & q_I & \cdots & q_I & \mu_I^{k_f} & \cdots & q_I \\
\end{array}
\end{pmatrix}
\end{align*}
$$

with $q_i \leq d_i$ and $\mu_i^{k_f} \leq q_i$, $\forall i$, which gives

$$V(m_P(k_f), D, Q(\mu)) = E_{D,Q(\mu)}[\mathbb{E}(m_P(k_f), D, Q(\mu))] - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i \mu_i^{k}
$$

where $E_{Q(\mu)}[\mathbb{E}(m_P, d, Q(\mu))]$ is calculated by solving the following $P2(m_P, d, q)$ and taking expectation over $Q(\mu)$:

$$
\pi(m_P, d, q) = \max_{Z_{ij}} \left\{ \sum_{i=1}^{l} \sum_{j=1}^{l} m_{P,ij} : Z_{ij}^{k_f} \right\}
$$

subject to:

$$\begin{align*}
\sum_{i=1}^{l} Z_{ij}^{k_f} & \leq \min_{1 \leq k \leq K, k \neq k_f, k_v} \{ \mu_i^{k_v}, q_i^{k_v}, d_i \} = \min \{ q_i, q_i^{k_v} \} & i = 1, 2, \ldots, I, \quad (100) \\
\sum_{j=1}^{l} Z_{ij}^{k_f} & \leq \mu_i^{k_f} & i = 1, 2, \ldots, I. \quad (101)
\end{align*}
$$

For configuration (I), consider capacity matrix $\tilde{\mu}$ which is exactly the same as $\mu$:

$$
\tilde{\mu} = \mu = \begin{pmatrix}
\begin{array}{cccccccc}
\text{stage 1} & \cdots & \text{stage (k_v-1)} & \text{stage k_v} & \text{stage (k_v+1)} & \cdots & \text{stage (k_f-1)} & \text{stage k_f} & \cdots & \text{stage K} \\
q_1 & \cdots & q_1 & \mu_1^{k_v} & q_1 & \cdots & q_1 & \mu_1^{k_f} & \cdots & q_1 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_I & \cdots & q_I & \mu_I^{k_v} & q_I & \cdots & q_I & \mu_I^{k_f} & \cdots & q_I \\
\end{array}
\end{pmatrix}
$$

Since process flexibility is only at stage $k_v$, we have

$$
\begin{align*}
V(m_P(k_v), D, Q(\tilde{\mu})) &= E_{D,Q(\tilde{\mu})}[\mathbb{E}(m_P(k_v), D, Q(\tilde{\mu}))] - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i \tilde{\mu}_i^{k}
\end{align*}
$$

$$
\begin{align*}
&= E_{Q(\tilde{\mu})}[\mathbb{E}(m_P, d, Q(\tilde{\mu}))] - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i \tilde{\mu}_i^{k}

&= E_{Q(\mu)}[\mathbb{E}(m_P, d, Q(\mu))] - \sum_{k=1}^{K} \sum_{i=1}^{I} c_i \mu_i^{k},
\end{align*}
$$

41
where \( E_{Q(\tilde{\mu})} \left[ \pi(\mathbf{m}_P, \mathbf{d}, Q(\tilde{\mu})) \right] \) is calculated by solving the following problem \( \text{P2}(\mathbf{m}_P, \mathbf{d}, \mathbf{q}) \) and taking expectation over \( Q(\tilde{\mu}) \):

\[
\pi(\mathbf{m}_P, \mathbf{d}, \mathbf{q}) = \max_{Z_{ij}^{k_f}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{I} \left[ m_{P,ij} \cdot Z_{ij}^{k_f} \right] \right\} \tag{102}
\]

subject to:

\[
\sum_{j=1}^{I} Z_{ij}^{k_f} \leq \min_{1 \leq k \leq K \neq k_v} \{ \tilde{\mu}_i^k, d_i \} = \min \{ q_i, \mu_i^{k_f}, d_i \} = \mu_i^{k_f} \text{ (by Lemma 4)} \quad i = 1, 2, \ldots, I, \tag{103}
\]

\[
\sum_{j=1}^{I} \tilde{Z}_{ji}^{k_f} \leq q_i^{k_v} \text{ (since flexibility and variability are both at stage } k_v) \quad i = 1, 2, \ldots, I. \tag{104}
\]

Comparing problems (99) and (102), it is clear that constraint (101) is the same as constrain (103), and constraint (100) is more strict than constraint (104). Therefore, for each realization of \( Q \), problem (102) achieves higher value than (99), i.e., \( \pi(\mathbf{m}_P, \mathbf{d}, \mathbf{q}(\tilde{\mu})) \geq \pi(\mathbf{m}_P, \mathbf{d}, \mathbf{q}(\mu)) \). Consequently, for all candidate optimal capacity investment matrix \( \mu \) for configuration (II), there exists a corresponding capacity matrix \( \tilde{\mu} \) for configuration (I), such that \( V(\mathbf{m}_P(k_v), \mathbf{d}, Q(\tilde{\mu})) \geq V(\mathbf{m}_P(k_f), \mathbf{d}, Q(\mu)) \) for all realization of \( Q \). Therefore, we have

\[
V^*(\mathbf{m}_P(k_v), \mathbf{d}) \geq V^*(\mathbf{m}_P(k_f), \mathbf{d}) \quad \forall k_f \neq k_v. \tag{105}
\]

In other words, configurations (I) outperforms configuration (II). Since \( k_f \) and \( k_v \) are arbitrarily chosen, this completes the proof.

Now we prove that investing in process flexibility at all the stages other than stage \( k_f \) gives the same profit. The proof is similar to PART (1), with the following minor modifications:

1. Since \( Q \) is stochastic and \( D \) is deterministic, to calculate \( V^* \), the expectations are taken over \( Q \) instead of \( D \);

2. In capacity matrices \( \mu \) and \( \tilde{\mu} \), there is a stage \( k_v \) with capacity investment \( \mu_i^{k_v} \);

3. In constraints with \( d_i \), \( d_i \) is replaced by \( q_i^{k_v} \), which is one realization of random capacity \( Q_i^{k_v} \).
ON-LINE APPENDIX II
Multi-Stage Flexibility Configurations: The Impact of Cost and Variability

Many facilities (e.g., automotive assembly plants, chemical plants, etc.) require investment to make their processes flexible. When the fixed cost of process flexibility is non-zero, it may not be feasible to have process or logistics flexibility at all stages. On the other hand, there are cases where flexibility can be located in more than one stage in a supply chain. To study the impact of cost and location of variability on the optimal location of process and logistics flexibility, in this Appendix, we focus on the optimal location of process and logistics flexibility in a two-echelon supply chain with two products when flexibility can be located in more than one stage, and the flexibility costs are non-zero.

Taking into account the interdependency of logistics and process flexibility, for a two-product two-echelon system, we depict the possible five flexibility configurations in Figure 2. Note that these are the only potentially optimal configurations for a two-product, two-plant system. (There are other configurations in which a certain stage of logistics (process) flexibility has been installed but is not in use because it is not accompanied by the necessary process (logistics) flexibility. Those configurations are suboptimal, since the fixed cost of logistics (process) flexibility is non-zero, and therefore not included in our analysis.) In these configurations we also show the raw material inventory and we include its cost of logistics flexibility (if established) in our analysis. Note that in order to make use of full process flexibility at stage \( K = 2 \), all plants at that stage need to be supplied with all types of raw material and therefore logistics flexibility is required from raw material to stage \( K \), as depicted in the figure. Structure (1) has no flexibility, structure (5) has full flexibility, while structures (2), (3) and (4) have partial flexibility (i.e., at least one stage does not have flexibility). Comparing the five structures we observe that if we vary costs of logistics and process flexibility and keep all the other parameters fixed, the optimal flexibility configuration follows a threshold structure. Depending on different parameter values, for demand side variability case, the optimal threshold structure results in a 2-regain, or a 3-regain, or a 4-region policy. Figure 3(Left) presents the case where the threshold structure results in a 4-region policy.

For the supply side variability case, the optimal threshold structure results in a 2- or 3-region policy. Figure 3(Right) presents the case where the optimal threshold results in a 3-region policy. To save space, in Theorem 3 we only present the results for the structures in Figure 3. The proof of this theorem along with the proofs for other threshold structures are presented at the end of this Appendix.

Theorem 3. For a two-product two-stage supply chain, consider \( V_1^* \) to \( V_5^* \) to be the optimal expected profit for configurations 1 to 5, respectively, and let \( \Lambda^k = \Lambda \neq 0 \), and \( \Psi^k = \Psi \neq 0 \) for \( k = 1, 2 \).

(1) (Variability only in demand): If \( V_3^* - V_1^* > V_2^* - V_1^* \), \( \frac{V_5^* - V_1^*}{2} > V_3^* - V_4^* \) and \( V_3^* - \frac{V_5^* + V_2^*}{2} > V_5^* - V_4^* \), then optimal flexibility configuration follows a 4-region policy as depicted in Figure 3(Left), which can be described as follows:

(a) \( \Psi + \Lambda < V_5^* - V_4^* \), and \( \Lambda < V_5^* - V_3^* \), then configuration (5) is optimal;

(b) \( \Psi + \Lambda < \frac{V_5^* - V_2^*}{2} \), \( \Lambda > V_5^* - V_4^* \), and \( \Psi < V_4^* - V_3^* \), then configuration (4) is optimal;

(c) \( \Psi + \Lambda > V_5^* - V_3^* \), \( 2\Lambda + \Psi < V_3^* - V_1^* \), and \( \Psi > V_4^* - V_3^* \), then configurations (3) and (2) are optimal;

(d) \( \Psi + \Lambda > \frac{V_5^* - V_2^*}{2} \), and \( 2\Lambda + \Psi > V_3^* - V_1^* \), then configuration (1) is optimal;

(2) (Variability only in Supply): If \( V_3^* - V_1^* > V_2^* - V_1^* \) and \( \frac{V_5^* - V_2^*}{2} < V_5^* - V_3^* \), then optimal flexibility configuration follows a 3-region policy as depicted in Figure 3(Right), which can be described as follows:

(a) \( \Psi + \Lambda < V_5^* - V_3^* \), and \( 2\Psi + 3\Lambda < V_5^* - V_1^* \), then configuration (5) is optimal;

(b) \( \Psi + \Lambda > V_5^* - V_3^* \) and \( 2\Lambda + \Psi < V_3^* - V_1^* \), then configuration (3) is optimal;

(c) \( 2\Lambda + \Psi > V_3^* - V_1^* \), and \( 2\Psi + 3\Lambda > V_5^* - V_1^* \), then configuration (1) is optimal;
Figure 2: Flexibility Configurations for Two-Product Two-Plant System: (1) No flexibility; (2) Full process flexibility at stage 1 with minimally required logistics flexibility; (3) Full process flexibility at stage 2 with minimally required logistics flexibility; (4) Full process flexibility at stages 1 and 2 with minimally required logistics flexibility; (5) Full process flexibility and full logistics flexibility at stages 1 and 2.

As both threshold structures in Figure 3 show, when the cost of flexibility is low, both types of flexibility are used at both stages. When the cost of flexibility is sufficiently high, no flexibility is used. For intermediate costs of process and logistics flexibility, thresholds define boundaries between use and nonuse of flexibility. Notice, however, that supply side variability still favors configurations with upstream logistics and process flexibility (Configuration 3) over those with downstream process and/or logistics flexibility (Configurations 2 and 4). In contrast, demand variability favors configurations with downstream logistics flexibility (Configuration 4) over configurations with upstream logistics flexibility (Configuration 3), unless the cost of process flexibility is high, which increases the cost of downstream logistics flexibility and offsets its natural advantage. Furthermore, for the case of demand variability, the location of process flexibility does not matter (i.e., Configurations 2 and 3 are equally good). These observations are consistent with our analytical results in Sections 3 and 4.

As the number of products and stages in the supply chain increases, the number of possible flexibility configurations also increases. While we cannot depict the optimal policy with a two-dimensional figure, we can still use the approach of Theorem 3 to demonstrate that the optimal policy consists of structured regions divided by thresholds.

PROOF OF THEOREM 3:

To evaluate a supply chain configuration, we introduce the total expected profit for a configuration $i$, $TP(i)$, which is defined as expected profit minus fixed cost associated with the investment in process and/or logistics flexibility. We list this total expected profit for all the five configurations as follows:

- $TP(1) = V_1^*$;
- $TP(2) = V_2^* - 2\Lambda - \Psi$;
- $TP(3) = V_3^* - 2\Lambda - \Psi$;
- $TP(4) = V_4^* - 2\Lambda - 2\Psi$;
- $TP(5) = V_5^* - 3\Lambda - 2\Psi$. 

44
From the figures, we observe the following relationship between the values of shown in Figure 5.

We use Figure 4 to show the above relationship in the space of Λ and Ψ. Note that Lines D, E, and F cross at point (V̄₁ − V̄₃, V̄₅ − V̄₇), Line A, E, and C cross at (2V̄₃ + V̄₅ − 3V̄₁, 2V̄₅ − V̄₇ − V̄₅), and Lines B, F, C cross at (3V̄₃ − V̄₅ − 2V̄₁, V̄₅ − V̄₇). Depending on the value of V̄₁ to V̄₅, we have the six threshold structures shown in Figure 5.

From the figures, we observe the following relationship between the values of V̄₁ to V̄₅, and the threshold structures in Figure 5.

1. If V̄₃ > V̄₅ and V̄₃ ≤ V̄₅ ≤ V̄₃, then threshold structure is shown in Figure 5 [6];
2. If V̄₃ > V̄₅ and V̄₃ ≤ V̄₅ ≤ V̄₃, then threshold structure is shown in Figure 5 [5];
3. If V̄₃ > V̄₅ and V̄₃ ≤ V̄₅ ≤ V̄₃, then threshold structure is shown in Figure 5 [4];
4. If V̄₃ > V̄₅ and V̄₃ ≤ V̄₅ ≤ V̄₃, then threshold structure is shown in Figure 5 [3];
5. If V̄₃ > V̄₅ and V̄₃ ≤ V̄₅ ≤ V̄₃, then threshold structure is shown in Figure 5 [2];
6. If V̄₃ > V̄₅ and V̄₃ ≤ V̄₅ ≤ V̄₃, then threshold structure is shown in Figure 5 [1].
PART (2):
When variability is in supply, it can be shown using Theorem 1 that configuration 3 outperforms configuration 4. Theorem 2, on the other hand, implies that configuration 3 outperforms configuration 2. Therefore, we only compare configurations 1, 3 and 5 in the following paragraphs.

Since configuration 5 has full logistics and process flexibility at both stages, configuration 1 has no flexibility, and configurations 3 has partial flexibility, it is clear that $V_1^* \leq V_3^* \leq V_5^*$. Therefore, we have

$$TP(1) > TP(3) \quad \text{if} \quad 2\Lambda + \Psi > V_3^* - V_1^*,$$

$$TP(1) > TP(5) \quad \text{if} \quad 3\Lambda + 2\Psi > V_5^* - V_1^*,$$

$$TP(3) > TP(5) \quad \text{if} \quad \Lambda + \Psi > V_5^* - V_3^*.$$

Diagrams (a), (c) and (e) in Figure 4 show the above relation. Depending on the values of $V_1^*$, $V_3^*$ and $V_5^*$, it can be concluded from diagrams (a), (c) and (e) in Figure 4 that there exists four threshold structures for optimal policy, which are shown in Figure 6. From the figures, we observe the following relationship between the values of $V_1^*$ to $V_5^*$ and the thresholds:

1. If $V_3^* - V_1^* \leq V_5^* - V_3^*$, then threshold structure is shown in Figure 6 [4];
2. If $V_3^* - V_1^* > V_5^* - V_3^*$ and $\frac{V_3^* - V_1^*}{2} < V_5^* - V_3^*$, then threshold structure is shown in Figure 6 [3];
3. If $\frac{V_3^* - V_1^*}{2} = V_5^* - V_3^*$, then threshold structure is shown in Figure 6 [2];
4. If $\frac{V_3^* - V_1^*}{2} > V_5^* - V_3^*$, then threshold structure is shown in Figure 6 [1].
Figure 5: Optimal Threshold Structures for Demand Side Variability.

Figure 6: Optimal Threshold Structures for Supply Side Variability.