Electronic Companion to "Tailored Base-Surge Policies in Dual-Sourcing Inventory Systems with Demand Learning"

by Boxiao Chen and Cong Shi

EC.1. Structural Properties for the Lost Sales Model

Since this part is not the focus of this paper, to avoid unnecessary repetition, we only briefly point out the differences in the proofs compared with Janakiraman et al. (2014).

Proof of Proposition 1. In Theorem 1, Lemma 2, Theorem 3, and Theorem 4 of Janakiraman et al. (2014), replace their function $C^{\mathfrak{D},Q,*}(h,b,c,l^E,l^R,F)$ (optimal cost under given Q for the backlogging system) by $\mathbb{E}[V(Q,S^*(Q))]$ (optimal cost under given Q for the lost sales system) in this paper and keep definitions of all the other notation unchanged, Theorem 1, Lemma 2, Theorem 3, and Theorem 4 continue to hold. In Theorem 5, Lemma 6, and Corollary 7, the definition of $C^{\mathfrak{D},*}(h,b,c,l^E,l^R,F)$ changes from the "optimal cost for the backlogging system" to the "optimal cost for the lost sales system", and Theorem 5, Lemma 6, and Corollary 7 continue to hold.

(a) Replace b by $b-c$, and Theorem 1 in Janakiraman et al. (2014) carries over directly.

(b) Lemma 2 in Janakiraman et al. (2014) carries over to the lost sales system because for any given Q, $\mathbb{E}[V(Q,S^*(Q))] \le C^{\mathfrak{D},Q,*}(h,b,c,l^E,l^R,F)$ since $b-c \le b$.

(c) Theorem 3 is a direct application of Lemma 2.

(d) Theorem 4 holds true because for any given Q, $\mathbb{E}[V(Q,S^*(Q))] \le C^{\mathfrak{D},Q,*}(h,b,c,l^E,l^R,F)$, thus $\min_Q \mathbb{E}[V(Q,S^*(Q))] \le \min_Q C^{\mathfrak{D},Q,*}(h,b,c,l^E,l^R,F)$.

(e) Theorem 5 can be proved in the same way for the lost sales case.

(f) Lemma 6 can be proved in the same way for the lost sales case for $A \le Q^l$, where A is a parameter in Janakiraman et al. (2014).

(g) Corollary 7 is a direct application of Lemma 6.

(h) Theorem 8 can be carried over for the following reasoning. In the lost sales system, any reasonable policy should satisfy $d_{\text{low}} \le Q \le S$. Given $Q = d_{\text{low}}$, because $S^*(Q) = d_{\text{low}}$ in the backlogging system, then it is also true that $S^*(Q) = d_{\text{low}}$ in the lost sales system. Because $S^*(Q)$ is decreasing in Q, then for any Q, one has $S^*(Q) = d_{\text{low}}$. Setting $S = d_{\text{low}}$ in (6), one sees that the optimal Q^* decreases in the penalty cost. Because in the backlogging system $Q^* = d_{\text{low}}$, in the lost sales system one also has $Q^* = d_{\text{low}}$.

The proof of Proposition 1 is thus complete. **Q.E.D.**

EC.2. Technical Proofs

Proof of Lemma 1. Construct another stochastic process as

$$\breve{O}_1 = O_\infty(Q) \text{ and } \breve{O}_{t+1} = (\breve{O}_t + Q - D_t)^+, \tag{EC.1}$$

and we see that $\breve{O}_t$ follows the same distribution as $O_\infty(Q)$ for any $t \geq 1$. For any $t \geq L$ and any measurable set $\Xi \subset [0,\infty)$, one has

$$|\mathbb{P}(\hat{O}_t \in \Xi) - \mathbb{P}(O_\infty(Q) \in \Xi)| = |\mathbb{P}(\hat{O}_t \in \Xi) - \mathbb{P}(\breve{O}_t \in \Xi)| \leq \mathbb{P}(\hat{O}_L = \breve{O}_L) \times 0 + \mathbb{P}(\hat{O}_L \neq \breve{O}_L) \times 2, \tag{EC.2}$$

where the inequality holds because once $\hat{O}_L = \breve{O}_L$, then the two processes couple, and $|\mathbb{P}(\hat{O}_t \in \Xi) - \mathbb{P}(O_\infty(Q) \in \Xi)| = 0$ for any $t \geq L$ and any measurable set $\Xi \subset [0,\infty)$.

$$\begin{aligned}
&\mathbb{P}(\hat{O}_L = \breve{O}_L)\\
\geq\ & \mathbb{P}\Big(\hat{O}_L = \breve{O}_L \big| \breve{O}_1 \leq \frac{1}{\kappa}\log T \log\log T\Big)\mathbb{P}\Big(\breve{O}_1 \leq \frac{1}{\kappa}\log T \log\log T\Big)\\
\geq\ & \mathbb{P}\Big(\sum_{t=1}^{L}(D_t - Q) \geq \max\big\{1, \frac{1}{\kappa}\big\}\log T \log\log T\Big)\mathbb{P}\Big(\breve{O}_1 \leq \frac{1}{\kappa}\log T \log\log T\Big)\\
\geq\ & \mathbb{P}\Big(\sum_{t=1}^{L}(D_t - Q) \geq \max\big\{1, \frac{1}{\kappa}\big\}\log T \log\log T\Big)(1 - e^{-\log T \log\log T}) \qquad \text{(EC.3)}\\
\geq\ & \mathbb{P}\Big(\sum_{t=1}^{L}(D_t - Q) - L(\mu - Q) \geq \max\big\{1, \frac{1}{\kappa}\big\}\log T \log\log T - L(\mu - Q^h)\Big)(1 - e^{-\log T \log\log T})\\
\geq\ & \mathbb{P}\Big(\sum_{t=1}^{L}(D_t - Q) - L(\mu - Q) \geq -2\sigma \log T(\log\log T)^{\frac{3}{2}}\Big)(1 - e^{-\log T \log\log T})\\
\geq\ & 1 - \frac{1}{T^3}, \qquad \text{(EC.4)}
\end{aligned}$$

where σ is the standard deviation of D_t, and (EC.3) follows from Lemmas EC.2 and EC.3, and (EC.4) follows from Lemma EC.4. Therefore, by (EC.2),

$$|\mathbb{P}(\hat{O}_t \in \Xi) - \mathbb{P}(O_\infty(Q) \in \Xi)| \leq \frac{2}{T^3}.$$

The proof of Lemma 1 is thus complete. **Q.E.D.**

Proof of Lemma 2. We proceed as the following, for any $n, m, j \in \{l, c, r\}$,

$$\begin{aligned}
\mathbb{P}(\hat{\mathcal{B}}_{nmj}) &= \mathbb{P}(\cap_{i\in\{1,\ldots,I_m\}}\hat{\mathcal{A}}^i_{nmj})\\
&\geq \mathbb{P}(\hat{\mathcal{A}}^{I_m}_{nmj} | \cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^i_{nmj})\mathbb{P}(\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^i_{nmj}).
\end{aligned} \tag{EC.5}$$

Next we are going to develop a lower bound for $\mathbb{P}(\hat{\mathcal{A}}^{I_m}_{nmj} | \cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^i_{nmj})$.

One has

$$\begin{aligned}&\mathbb{P}(\hat{\mathcal{A}}^{I_m}_{nmj}|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj})\\ &=\mathbb{P}\Big(\hat{O}_{\tau^j_{nm}+I_mL}\le\frac{1}{\kappa}\log T\log\log T,\ O_{\tau^j_{nm}+I_mL}=\hat{O}_{\tau^j_{nm}+I_mL}|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}\Big).\end{aligned}\tag{EC.6}$$

By Lemmas EC.2 and EC.3, one has

$$\mathbb{P}\Big(O_\infty(Q^j_n)>\frac{1}{\kappa}\log T\log\log T\Big)\le e^{-\kappa\frac{1}{\kappa}\log T\log\log T}\le\frac{1}{T^3}.\tag{EC.7}$$

Under $\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}$, $\hat{O}_{\tau^j_{nm}+(I_m-1)L}\le 1/\kappa\log T\log\log T$, then by Lemma 1 and (EC.7), one has that

$$\mathbb{P}\Big(\hat{O}_{\tau^j_{nm}+I_mL}>\frac{1}{\kappa}\log T\log\log T\Big|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}\Big)\le\frac{3}{T^3},$$

which leads to

$$\mathbb{P}\Big(\hat{O}_{\tau^j_{nm}+I_mL}\le\frac{1}{\kappa}\log T\log\log T\Big|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}\Big)>1-\frac{3}{T^3}.\tag{EC.8}$$

Now we are going to study $\mathbb{P}\Big(O_{\tau^j_{nm}+I_mL}=\hat{O}_{\tau^j_{nm}+I_mL}|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}\Big)$. Under $\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}$, one has that

$$|O_{\tau^j_{nm}+(I_m-1)L+1}-\hat{O}_{\tau^j_{nm}+(I_m-1)L+1}|\le|S^{I_m}_{nmj}-S^{I_m-1}_{nmj}|\le\eta_{I_m}\max\{h,c-b\},$$

and

$$O_{\tau^j_{nm}+(I_m-1)L+1}\le K_8\log T\log\log T,\quad \hat{O}_{\tau^j_{nm}+(I_m-1)L+1}\le K_8\log T\log\log T.$$

In order for $\hat{O}_t$ and O_t to couple during period $t\in\{\tau^j_{nm}+(I_m-1)L+2,\ldots,\tau^j_{nm}+I_mL\}$, a sufficient condition is

$$\sum_{t=\tau^j_{nm}+(I_m-1)L+1}^{\tau^j_{nm}+I_mL-1}(D_t-Q^j_n)\ge K_8\log T\log\log T.\tag{EC.9}$$

Therefore, one has

$$\begin{aligned}&\mathbb{P}\Big(O_{\tau^j_{nm}+I_mL}=\hat{O}_{\tau^j_{nm}+I_mL}\Big|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}\Big)\\ &\ge\mathbb{P}\Big(\sum_{t=\tau^j_{nm}+(I_m-1)L+1}^{\tau^j_{nm}+I_mL-1}(D_t-Q^j_n)\ge K_8\log T\log\log T\Big|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}\Big)\\ &\ge\mathbb{P}\Big(\sum_{t=\tau^j_{nm}+(I_m-1)L+1}^{\tau^j_{nm}+I_mL-1}(D_t-Q^j_n)-(L-1)(\mu-Q^j_n)\ge K_8\log T\log\log T-(L-1)(\mu-Q^h)\Big|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}\Big)\\ &\ge\mathbb{P}\Big(\sum_{t=\tau^j_{nm}+(I_m-1)L+1}^{\tau^j_{nm}+I_mL-1}(D_t-Q^j_n)-(L-1)(\mu-Q^j_n)\ge-2\sigma\log T(\log\log T)^{\frac{3}{2}}\Big|\cap_{i\in\{1,\ldots,I_m-1\}}\hat{\mathcal{A}}^{i}_{nmj}\Big)\\ &\ge 1-\frac{1}{T^3},\end{aligned}\tag{EC.10}$$

where the last inequality is implied by Lemma EC.4.

Combining (EC.8) and (EC.10), one obtains a lower bound for (EC.6) as

$$\mathbb{P}(\hat{\mathcal{A}}^{I_m}_{nmj} | \cap_{i\in\{1,\dots,I_m-1\}} \hat{\mathcal{A}}^i_{nmj}) \geq 1 - \frac{4}{T^3}. \tag{EC.11}$$

Plugging (EC.11) into (EC.5), one has that for any $n, m, j \in \{l, c, r\}$,

$$\mathbb{P}(\cap_{i\in\{1,\dots,I_m\}} \hat{\mathcal{A}}^i_{nmj}) \geq (1 - \frac{4}{T^3})\mathbb{P}(\cap_{i\in\{1,\dots,I_m-1\}} \hat{\mathcal{A}}^i_{nmj}),$$

and continue the same procedure, one will have

$$\begin{aligned}\mathbb{P}(\hat{\mathcal{B}}_{nmj}) &= \mathbb{P}(\cap_{i\in\{1,\dots,I_m\}} \hat{\mathcal{A}}^i_{nmj}) \\ &\geq (1 - \frac{4(I_m-1)}{T^3})\mathbb{P}(\hat{\mathcal{A}}^1_{nmj}) \\ &\geq (1 - \frac{4(I_m-1)}{T^3})\mathbb{P}(\hat{\mathcal{A}}^1_{nmj} | \hat{\mathcal{B}}_{\Gamma(nmj)})\mathbb{P}(\hat{\mathcal{B}}_{\Gamma(nmj)}),\end{aligned} \tag{EC.12}$$

where $\Gamma(nmj)$ denotes the tuple right before n, m, j. $\mathbb{P}(\hat{\mathcal{A}}^1_{nmj} | \hat{\mathcal{B}}_{\Gamma(nmj)})$ in (EC.12) can be bounded as

$$\mathbb{P}(\hat{\mathcal{A}}^1_{nmj} \big| \hat{\mathcal{B}}_{\Gamma(nmj)}) = \mathbb{P}\Big(\hat{O}_{\tau^j_{nm}+L} \leq \frac{1}{\kappa}\log T \log\log T, O_{\tau^j_{nm}+L} = \hat{O}_{\tau^j_{nm}+L} \Big| \hat{\mathcal{B}}_{\Gamma(nmj)}\Big). \tag{EC.13}$$

Again, by Lemmas EC.2 and EC.3, one has

$$\mathbb{P}\Big(O_\infty(Q^j_n) > \frac{1}{\kappa}\log T \log\log T\Big) \leq \frac{1}{T^3}.$$

By definition, $\hat{O}_{\tau^j_{nm}+1} = 0$, and by Lemma 1, one has

$$\mathbb{P}\Big(\hat{O}_{\tau^j_{nm}+L} > \frac{1}{\kappa}\log T \log\log T \Big| \hat{\mathcal{B}}_{\Gamma(nmj)}\Big) \leq \frac{3}{T^3},$$

which leads to

$$\mathbb{P}\Big(\hat{O}_{\tau^j_{nm}+L} \leq \frac{1}{\kappa}\log T \log\log T \Big| \hat{\mathcal{B}}_{\Gamma(nmj)}\Big) > 1 - \frac{3}{T^3}. \tag{EC.14}$$

In order to bound $\mathbb{P}\Big(O_{\tau^j_{nm}+L} = \hat{O}_{\tau^j_{nm}+L} | \hat{\mathcal{B}}_{\Gamma(nmj)}\Big)$, we start by analyzing the behaviors of stochastic processes $\{O_t\}^{\tau^j_{nm}+L}_{\tau^j_{nm}+1}$ and $\{\hat{O}_t\}^{\tau^j_{nm}+L}_{\tau^j_{nm}+1}$.

During periods $t = \tau^j_{nm}+1, \dots, \tau^j_{nm}+l-1$, O_t and $\hat{O}_t$ follow different recursions as

$$O_{t+1} = ((O_t + S^{j1}_{nm} - d_t)^+ + q^R_{t+1-l} - S^{j1}_{nm})^+$$

as opposed to

$$\hat{O}_{t+1} = (\hat{O}_t + Q^j_n - D_t)^+ \text{ with } \hat{O}_{\tau^j_{nm}+1} = 0.$$

Note that $q^R_{t+1-l} \neq Q^j_n$ due to the l periods' delay from the regular source.

During periods $t = \tau^j_{nm} + l, \ldots, \tau^j_{nm} + L$, they follow the same recursion, since the system of O_t starts to receive Q^j_n every period from the regular source, and because $Q^j_n \leq S^{j1}_{nm}$, we have

$$O_{t+1} = (O_t + Q^j_n - D_t)^+,$$

and by definition, one has

$$\hat{O}_{t+1} = (\hat{O}_t + Q^j_n - D_t)^+.$$

Under $\hat{\mathcal{B}}_{\Gamma(nmj)}$, it holds that $O_{\tau^j_{nm}+1} \leq (\frac{1}{\kappa} + 1)\log T \log\log T$, then $O_{\tau^j_{nm}+l} \leq O_{\tau^j_{nm}+1} + K_9 \leq K_{10} \log T \log\log T$. On the other hand, $\hat{O}_{\tau^j_{nm}+1} = 0$, and then $\hat{O}_{\tau^j_{nm}+l} \leq \hat{O}_{\tau^j_{nm}+1} + K_{11} \leq K_{12}$.

The stochastic processes O_t and $\hat{O}_t$ will couple during a period $t \in \{\tau^j_{nm} + l + 1, \ldots, \tau^j_{nm} + L\}$, if there exists $t_0 \in \{\tau^j_{nm} + l, \ldots, \tau^j_{nm} + L - 1\}$ such that

$$\sum_{t=\tau^j_{nm}+l}^{t_0} (D_t - Q^j_n) \geq \max\{O_{\tau^j_{nm}+l}, \hat{O}_{\tau^j_{nm}+l}\}.$$

Therefore, one has

$$\begin{aligned}
&\mathbb{P}\Big(O_{\tau^j_{nm}+L} = \hat{O}_{\tau^j_{nm}+L} \Big| \hat{\mathcal{B}}_{\Gamma(nmj)}\Big) \\
&\geq \mathbb{P}\Big(\sum_{t=\tau^j_{nm}+l}^{\tau^j_{nm}+L-1} (D_t - Q^j_n) \geq K_{13} \log T \log\log T \Big| \hat{\mathcal{B}}_{\Gamma(nmj)}\Big) \\
&\geq \mathbb{P}\Big(\sum_{t=\tau^j_{nm}+l}^{\tau^j_{nm}+L-1} (D_t - Q^j_n) - (L-l)(\mu - Q^j_n) \geq K_{13} \log T \log\log T - (L-l)(\mu - Q^h) \Big| \hat{\mathcal{B}}_{\Gamma(nmj)}\Big) \\
&\geq \mathbb{P}\Big(\sum_{t=\tau^j_{nm}+l}^{\tau^j_{nm}+L-1} (D_t - Q^j_n) - (L-l)(\mu - Q^j_n) \geq -2\sigma \log T (\log\log T)^{\frac{3}{2}} \Big| \hat{\mathcal{B}}_{\Gamma(nmj)}\Big) \\
&\geq 1 - \frac{1}{T^3},
\end{aligned} \tag{EC.15}$$

where the last inequality is implied by Lemma EC.4.

Combining (EC.14) and (EC.15), one obtains a lower bound for (EC.13) as

$$\mathbb{P}(\hat{\mathcal{A}}^1_{nmj} | \hat{\mathcal{B}}_{\Gamma(nmj)}) \geq 1 - \frac{4}{T^3}. \tag{EC.16}$$

Plugging (EC.16) into (EC.12), one has that

$$\mathbb{P}(\hat{\mathcal{B}}_{nmj}) \geq \big(1 - \frac{4I_m}{T^3}\big)\mathbb{P}(\hat{\mathcal{B}}_{\Gamma(nmj)}). \tag{EC.17}$$

Set $O_1=\hat{O}_1=0$ (or to achieve the same theoretical result, set $\mathbb{P}(\hat{\mathcal{B}}_{\Gamma(111)})=1$), and repeat the above procedure, one will obtain that

$$\mathbb{P}(\hat{\mathcal{B}}_{111})\geq 1-\frac{4I_1}{T^3},$$

then by (EC.17), one has that for any $n,m,j\in\{l,c,r\}$,

$$\mathbb{P}(\hat{\mathcal{B}}_{nmj})\geq 1-\frac{4}{T^3}\sum_n\sum_m\sum_j I_m=1-\frac{4}{T^2}.$$

The proof of Lemma 2 is thus complete. **Q.E.D.**

Proof of Proposition 3. One has

$$\begin{aligned}
&\left|\mathbb{E}\left[\sum_{t=1}^{T}C_t-C^{B1}\right]\right|\\
=&\left|\mathbb{E}\Bigg[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\sum_{i=1}^{I_m}\sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL}\big((-cQ_n^j+h(S_{nm}^{ji}+O_t-D_t)^+\right.\\
&\qquad\left.-(b-c)\min\{S_{nm}^{ji}+O_t,D_t\})-G(Q_n^j,S_{nm}^{ji})\big)\Bigg]\right|+c\mathbb{E}[x_{T+1}]\\
\leq&\left|\mathbb{E}\Bigg[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\sum_{i=1}^{I_m}\sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL}\big((-cQ_n^j+h(S_{nm}^{ji}+O_t-D_t)^+-(b-c)\min\{S_{nm}^{ji}+O_t,D_t\})\right.\\
&\qquad\left.-(-cQ_n^j+h(S_{nm}^{ji}+\hat{O}_t-D_t)^+-(b-c)\min\{S_{nm}^{ji}+\hat{O}_t,D_t\})\big)\Bigg]\right|\\
&+\left|\mathbb{E}\Bigg[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\sum_{i=1}^{I_m}\sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL}\big((-cQ_n^j+h(S_{nm}^{ji}+\hat{O}_t-D_t)^+\right.\\
&\qquad\left.-(b-c)\min\{S_{nm}^{ji}+\hat{O}_t,D_t\})-G(Q_n^j,S_{nm}^{ji})\big)\right|+c\mathbb{E}[x_{T+1}].
\end{aligned}\tag{EC.18}$$

To bound the first term in (EC.18), for any given $n,m,j\in\{l,c,r\}$ and any realizations of τ_{nm}^j, we proceed as the following.

For $i=1$ and periods $t=\tau_{nm}^j+(i-1)L+1,\dots,\tau_{nm}^j+iL$, one has that, on event $\hat{\mathcal{B}}_{\Gamma(nmj)}$,

$$|O_t-\hat{O}_t|\leq K_{14}\log T\log\log T;$$

on event $(\hat{\mathcal{B}}_{\Gamma(nmj)})^C$,

$$S_{nm}^{ji}+O_t\leq K_{15}t \text{ and } S_{nm}^{ji}+\hat{O}_t\leq K_{16}(t-\tau_{nm}^j).\tag{EC.19}$$

For $i=2,\dots,I_m$ and periods $t=\tau_{nm}^j+(i-1)L+1,\dots,\tau_{nm}^j+iL$, one has that, on event $\hat{\mathcal{B}}_{nmj}$,

$$|O_t-\hat{O}_t|\leq|S_{nm}^{ji}-S_{nm}^{j(i-1)}|\leq\eta_i\max\{h,c-b\};$$

on event $(\hat{\mathcal{B}}_{nmj})^C$,

$$S_{nm}^{ji} + O_t \le K_{17} t \text{ and } S_{nm}^{ji} + \hat{O}_t \le K_{18}(t - \tau_{nm}^j). \tag{EC.20}$$

Therefore, one has

$$\begin{aligned}
&\sum_{i=1} \sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL} \Big| \mathbb{E}[-cQ_n^j + h(S_{nm}^{ji} + O_t - D_t)^+ - (b-c)\min\{S_{nm}^{ji} + O_t, D_t\} | \hat{\mathcal{B}}_{\Gamma(nmj)}] \\
&\qquad\qquad - \mathbb{E}[-cQ_n^j + h(S_{nm}^{ji} + \hat{O}_t - D_t)^+ - (b-c)\min\{S_{nm}^{ji} + \hat{O}_t, D_t\} | \hat{\mathcal{B}}_{\Gamma(nmj)}] \Big| \\
&\le \sum_{i=1} \sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL} \mathbb{E}\big[\max\{h, b-c\} |O_t - \hat{O}_t| \big| \hat{\mathcal{B}}_{\Gamma(nmj)}\big] \\
&\le K_{19} \log T \log\log T L \\
&= K_{19} (\log T)^2 (\log\log T)^3,
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{i=2}^{I_m} \sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL} \Big| \mathbb{E}[-cQ_n^j + h(S_{nm}^{ji} + O_t - D_t)^+ - (b-c)\min\{S_{nm}^{ji} + O_t, D_t\} | \hat{\mathcal{B}}_{nmj}] \\
&\qquad\qquad - \mathbb{E}[-cQ_n^j + h(S_{nm}^{ji} + \hat{O}_t - D_t)^+ - (b-c)\min\{S_{nm}^{ji} + \hat{O}_t, D_t\} | \hat{\mathcal{B}}_{nmj}] \Big| \\
&\le \sum_{i=2}^{I_m} \sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL} \mathbb{E}\big[\max\{h, b-c\} |O_t - \hat{O}_t| \big| \hat{\mathcal{B}}_{nmj}\big] \\
&\le \sum_{i=2}^{I_m} \frac{K_{20}}{\sqrt{i}} L \\
&\le K_{21} \sqrt{T} \log T (\log\log T)^2.
\end{aligned}$$

Then one has

$$\begin{aligned}
&\sum_{i=1}^{I_m} \sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL} \Big| \mathbb{E}[-cQ_n^j + h(S_{nm}^{ji} + O_t - D_t)^+ - (b-c)\min\{S_{nm}^{ji} + O_t, D_t\}] \\
&\qquad\qquad - \mathbb{E}[-cQ_n^j + h(S_{nm}^{ji} + \hat{O}_t - D_t)^+ - (b-c)\min\{S_{nm}^{ji} + \hat{O}_t, D_t\}] \Big| \\
&\le \mathbb{P}(\hat{\mathcal{B}}_{\Gamma(nmj)}) K_{22} (\log T)^2 (\log\log T)^3 + (1 - \mathbb{P}(\hat{\mathcal{B}}_{\Gamma(nmj)})) \sum_{i=1} \sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL} K_{23} t \\
&\qquad\qquad + \mathbb{P}(\hat{\mathcal{B}}_{nmj}) K_{24} \sqrt{T} \log T (\log\log T)^2 + (1 - \mathbb{P}(\hat{\mathcal{B}}_{nmj})) \sum_{i=2}^{I_m} \sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL} K_{25} t \\
&\le K_{26} \sqrt{T} \log T (\log\log T)^2 + K_{27} T^2 \times \frac{1}{T^2}.
\end{aligned}$$

This leads to

$$\begin{aligned}&\Bigg|\mathbb{E}\Bigg[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\sum_{i=1}^{I_m}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}\big((-cQ^j_n+h(S^{ji}_{nm}+O_t-D_t)^+-(b-c)\min\{S^{ji}_{nm}+O_t,D_t\})\\&\qquad\qquad -(-cQ^j_n+h(S^{ji}_{nm}+\hat{O}_t-D_t)^+-(b-c)\min\{S^{ji}_{nm}+\hat{O}_t,D_t\})\big)\Bigg]\Bigg|\\&\le\mathbb{E}\Bigg[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\big(K_{28}\sqrt{T}\log T(\log\log T)^2+K_{29}\big)\Bigg].\end{aligned}\tag{EC.21}$$

The first term in (EC.18) is thus upper bounded.

To bound the second term in (EC.18), for any given $n,m,j\in\{l,c,r\}$ and any realizations of τ^j_{nm}, Q^j_n and S^{ji}_{nm}, we proceed as the following.

$$\begin{aligned}&\sum_{i=1}^{I_m}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}\Big|\mathbb{E}[-cQ^j_n+h(S^{ji}_{nm}+\hat{O}_t-D_t)^+-(b-c)\min\{S^{ji}_{nm}+\hat{O}_t,D_t\}]-\mathbb{E}[G(Q^j_n,S^{ji}_{nm})]\Big|\\&\le\sum_{i=1}^{I_m}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}h\Big|\mathbb{E}[(S^{ji}_{nm}+\hat{O}_t-D_t)^+]-\mathbb{E}[(S^{ji}_{nm}+O_\infty(Q^j_n)-D_t)^+]\Big|\\&\qquad+\sum_{i=1}^{I_m}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}(b-c)\Big|\mathbb{E}[(D_t-S^{ji}_{nm}-\hat{O}_t)^+]-\mathbb{E}[(D_t-S^{ji}_{nm}-O_\infty(Q^j_n))^+]\Big|.\end{aligned}\tag{EC.22}$$

To bound the first term in (EC.22), one has that for $i=1$,

$$\begin{aligned}&\sum_{i=1}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}h\Big|\mathbb{E}[(S^{ji}_{nm}+\hat{O}_t-D_t)^+]-\mathbb{E}[(S^{ji}_{nm}+O_\infty(Q^j_n)-D_t)^+]\Big|\\&\le\sum_{i=1}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}K_{30}(t-\tau^j_{nm})+\sum_{i=1}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}(S^h+\mathbb{E}[O_\infty(Q^j_n)])\\&\le\sum_{i=1}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}K_{31}(t-\tau^j_{nm})+\sum_{i=1}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL}(S^h+\frac{\sigma^2}{2(\mu-Q^j_n)})\\&\le K_{32}(\log T)^2(\log\log T)^4,\end{aligned}$$

where the second last inequality follows from the Kingman's inequality (Kingman 1970).

For $i=2,\dots,I_m$ and any realization $D_t=d$,

$$\begin{aligned}&\Big|\mathbb{E}[(S^{ji}_{nm}+\hat{O}_t-d)^+]-\mathbb{E}[(S^{ji}_{nm}+O_\infty(Q^j_n)-d)^+]\Big|\\&=\Big|\int_0^\infty\mathbb{P}(S^{ji}_{nm}+O_\infty(Q^j_n)-d>z)dz-\int_0^{S^{ji}_{nm}+Q^j_n(t-\tau^j_{nm})}\mathbb{P}(S^{ji}_{nm}+\hat{O}_t-d>z)dz\Big|\end{aligned}$$

$$\begin{aligned}
&\leq \int_{S^{ji}_{nm}+Q^j_n(t-\tau^j_{nm})}^{\infty} \mathbb{P}(S^{ji}_{nm}+O_\infty(Q^j_n)-d>z)dz \\
&\quad + \int_0^{S^{ji}_{nm}+Q^j_n(t-\tau^j_{nm})} |\mathbb{P}(S^{ji}_{nm}+O_\infty(Q^j_n)-d>z)-\mathbb{P}(S^{ji}_{nm}+\hat{O}_t-d>z)|dz \\
&\leq \int_{Q^j_n(t-\tau^j_{nm})}^{\infty} \mathbb{P}(O_\infty(Q^j_n)>z+d)dz+(S^{ji}_{nm}+Q^j_n(t-\tau^j_{nm}))\frac{2}{T^3} \\
&\leq \int_{Q^j_n(t-\tau^j_{nm})}^{\infty} e^{-\kappa z}dz+(S^{ji}_{nm}+Q^j_n(t-\tau^j_{nm}))\frac{2}{T^3} \\
&= \frac{1}{\kappa}e^{-\kappa Q^j_n(t-\tau^j_{nm})}+(S^{ji}_{nm}+Q^j_n(t-\tau^j_{nm}))\frac{2}{T^3} \\
&\leq \frac{1}{\kappa}e^{-\kappa Q^l(t-\tau^j_{nm})}+(S^h+Q^h(t-\tau^j_{nm}))\frac{2}{T^3},
\end{aligned}$$

where the second inequality is implied by Lemma 1, and the third inequality is based on Lemmas EC.2 and EC.3.

Therefore, the first term in (EC.22) can be bounded as

$$\begin{aligned}
&\sum_{i=1}^{I_m}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL} h\left|\mathbb{E}[(S^{ji}_{nm}+\hat{O}_t-D_t)^+]-\mathbb{E}[(S^{ji}_{nm}+O_\infty(Q^j_n)-D_t)^+]\right| \\
&\leq K_{33}(\log T)^2(\log\log T)^4+\sum_{i=2}^{I_m}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL} h\left(\frac{1}{\kappa}e^{-\kappa Q^l(t-\tau^j_{nm})}+(S^h+Q^h(t-\tau^j_{nm}))\frac{2}{T^3}\right) \\
&\leq K_{34}(\log T)^2(\log\log T)^4. \qquad \text{(EC.23)}
\end{aligned}$$

The second term in (EC.22) can be bounded as the following. For $i=1$,

$$\begin{aligned}
&\sum_{i=1}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL} (b-c)\left|\mathbb{E}[(D_t-S^{ji}_{nm}-\hat{O}_t)^+]-\mathbb{E}[(D_t-S^{ji}_{nm}-O_\infty(Q^j_n))^+]\right| \\
&\leq \sum_{i=1}\sum_{t=\tau^j_{nm}+(i-1)L+1}^{\tau^j_{nm}+iL} K_{35}(t-\tau^j_{nm}) \\
&\leq K_{36}(\log T)^2(\log\log T)^4.
\end{aligned}$$

For $i\geq 2$ and any realizations of d,

$$\begin{aligned}
&\left|\mathbb{E}[(d-S^{ji}_{nm}-\hat{O}_t)^+]-\mathbb{E}[(d-S^{ji}_{nm}-O_\infty(Q^j_n))^+]\right| \\
&\leq \int_0^d |\mathbb{P}(d-S^{ji}_{nm}-\hat{O}_t>z)-\mathbb{P}(d-O_\infty(Q^j_n)-\hat{O}_t>z)|dz \\
&\leq \frac{2d}{T^3},
\end{aligned}$$

where the second inequality is implied by Lemma 1, which yields

$$\left|\mathbb{E}[(D_t-S^{ji}_{nm}-\hat{O}_t)^+]-\mathbb{E}[(D_t-S^{ji}_{nm}-O_\infty(Q^j_n))^+]\right|\leq\frac{2\mu}{T^3}. \qquad \text{(EC.24)}$$

Therefore, the second term in (EC.22) is bounded as

$$\begin{aligned}
&\sum_{i=1}^{I_m}\sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL}(b-c)\left|\mathbb{E}[(D_t-S_{nm}^{ji}-\hat{O}_t)^+]-\mathbb{E}[(D_t-S_{nm}^{ji}-O_\infty(Q_n^j))^+]\right|\\
&\le K_{37}(\log T)^2(\log\log T)^4+\sum_{i=2}^{I_m}\sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL}\frac{2\mu(b-c)}{T^3}\\
&\le K_{38}(\log T)^2(\log\log T)^4.
\end{aligned}\tag{EC.25}$$

Combining (EC.23) and (EC.25), and plug them in (EC.22),

$$\begin{aligned}
&\sum_{i=1}^{I_m}\sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL}\left|\mathbb{E}[-cQ_n^j+h(S_{nm}^{ji}+\hat{O}_t-D_t)^+-(b-c)\min\{S_{nm}^{ji}+\hat{O}_t,D_t\}]-\mathbb{E}[G(Q_n^j,S_{nm}^{ji})]\right|\\
&\le K_{39}(\log T)^2(\log\log T)^4.
\end{aligned}\tag{EC.26}$$

Therefore, the second term in (EC.18) is upper bounded as

$$\begin{aligned}
&\left|\mathbb{E}\left[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\sum_{i=1}^{I_m}\sum_{t=\tau_{nm}^j+(i-1)L+1}^{\tau_{nm}^j+iL}\left((-cQ_n^j+h(S_{nm}^{ji}+\hat{O}_t-D_t)^+-(b-c)\min\{S_{nm}^{ji}+\hat{O}_t,D_t\})-G(Q_n^j,S_{nm}^{ji})\right)\right]\right|\\
&\le \mathbb{E}\left[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}K_{40}(\log T)^2(\log\log T)^4\right].
\end{aligned}\tag{EC.27}$$

To bound the third term in (EC.18), let $(\breve{n},\breve{m},\breve{j})$ be the last *complete* round that has L periods, then we proceed as

$$\begin{aligned}
\mathbb{E}[x_{T+1}]&=\mathbb{E}[x_{T+1}|\hat{\mathcal{B}}_{\breve{n}\breve{m}\breve{j}}]\mathbb{P}(\hat{\mathcal{B}}_{\breve{n}\breve{m}\breve{j}})+\mathbb{E}[x_{T+1}|(\hat{\mathcal{B}}_{\breve{n}\breve{m}\breve{j}})^C]\mathbb{P}((\hat{\mathcal{B}}_{\breve{n}\breve{m}\breve{j}})^C)\\
&\le K_{41}\log T(\log\log T)^2+K_{42}T\times\frac{1}{T^2},
\end{aligned}$$

where the inequality is valid because under $\hat{\mathcal{B}}_{\breve{n}\breve{m}\breve{j}}$, $O_{\tau_{\breve{n}\breve{m}\breve{j}}+N_{\breve{m}}L}\le\frac{1}{\kappa}\log T\log\log T$, and by the definition of $(\breve{n},\breve{m},\breve{j})$, $T-(\tau_{\breve{n}\breve{m}\breve{j}}+N_{\breve{m}}L)\le L$, which yields $x_{T+1}\le\frac{1}{\kappa}\log T\log\log T+K_{43}L=K_{44}\log T(\log\log T)^2$; on the other hand, under $(\hat{\mathcal{B}}_{\breve{n}\breve{m}\breve{j}})^C$, $x_{T+1}\le K_{45}T$.

Therefore, the third term in (EC.18) is upper bounded as

$$\mathbb{E}[x_{T+1}]\le K_{46}\log T(\log\log T)^2.\tag{EC.28}$$

Combining (EC.21), (EC.27) and (EC.28), and plug them into (EC.18), one has that

$$\left|\mathbb{E}\left[\sum_{t=1}^{T}C_t-C^{B1}\right]\right|\le\mathbb{E}\left[\sum_{n=1}^{N}\sum_{m=1}^{M_n}K_{47}\sqrt{T}\log T(\log\log T)^2\right].\tag{EC.29}$$

Proposition 3 is thus proved. **Q.E.D.**

Proof of Lemma 3. We proceed as the following, for any $n, m, j \in \{l, c, r\}$,

$$\begin{aligned}\mathbb{P}(\tilde{\mathcal{B}}_{nmj}) &= \mathbb{P}(\cap_{i\in\{1,\ldots,I_m\}}\tilde{\mathcal{A}}^i_{nmj}) \\ &\geq \mathbb{P}(\tilde{\mathcal{A}}^{I_m}_{nmj} | \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj})\mathbb{P}(\cap_{i\in\{1,\ldots,I_m-1\}}\tilde{\mathcal{A}}^i_{nmj}). \qquad \text{(EC.30)}\end{aligned}$$

Next we are going to develop a lower bound for $\mathbb{P}(\tilde{\mathcal{A}}^{I_m}_{nmj} | \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj})$.

One has

$$\begin{aligned}&\mathbb{P}(\tilde{\mathcal{A}}^{I_m}_{nmj} | \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj}) \qquad \text{(EC.31)} \\ &= \mathbb{P}\Big(\tilde{O}_{\tau^j_{nm}+I_mL} \leq \frac{1}{\kappa}\log T \log\log T,\ O_{\tau^j_{nm}+I_mL} = \tilde{O}_{\tau^j_{nm}+I_mL} | \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj}\Big).\end{aligned}$$

By Lemmas EC.2 and EC.3, one has

$$\mathbb{P}\Big(\tilde{O}_{\tau^j_{nm}+I_mL} > \frac{1}{\kappa}\log T \log\log T\Big) \leq \frac{1}{T^3},$$

and because $\tilde{O}_{\tau^j_{nm}+I_mL}$ is independent of $\cap_{i\in\{1,\ldots,I_m-1\}}\tilde{\mathcal{A}}^i_{nmj}$, this leads to

$$\mathbb{P}\Big(\tilde{O}_{\tau^j_{nm}+I_mL} \leq \frac{1}{\kappa}\log T \log\log T\Big| \cap_{i\in\{1,\ldots,I_m-1\}} \hat{\mathcal{A}}^i_{nmj}\Big) > 1 - \frac{1}{T^3}. \qquad \text{(EC.32)}$$

Now we are going to bound $\mathbb{P}\Big(O_{\tau^j_{nm}+I_mL} = \tilde{O}_{\tau^j_{nm}+I_mL} | \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj}\Big)$ as

$$\begin{aligned}&\mathbb{P}\Big(O_{\tau^j_{nm}+I_mL} = \tilde{O}_{\tau^j_{nm}+I_mL}\Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj}\Big) \\ &\geq \mathbb{P}\Big(O_{\tau^j_{nm}+I_mL} = \tilde{O}_{\tau^j_{nm}+I_mL}\Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj},\ \tilde{O}_{\tau^j_{nm}+(I_m-1)L+1} \leq \frac{1}{\kappa}\log T \log\log T\Big) \\ &\qquad\qquad \times \mathbb{P}\Big(\tilde{O}_{\tau^j_{nm}+(I_m-1)L+1} \leq \frac{1}{\kappa}\log T \log\log T\Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj}\Big) \\ &\geq \mathbb{P}\Big(O_{\tau^j_{nm}+I_mL} = \tilde{O}_{\tau^j_{nm}+I_mL}\Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}^i_{nmj},\ \tilde{O}_{\tau^j_{nm}+(I_m-1)L+1} \leq \frac{1}{\kappa}\log T \log\log T\Big) \qquad \text{(EC.33)} \\ &\qquad\qquad \times \big(1 - \frac{1}{T^3}\big),\end{aligned}$$

where the last inequality comes from Lemmas EC.2 and EC.3 and the fact that $\tilde{O}_{\tau^j_{nm}+(I_m-1)L+1}$ is independent of $\cap_{i\in\{1,\ldots,I_m-1\}}\tilde{\mathcal{A}}^i_{nmj}$.

Under $\cap_{i\in\{1,\ldots,I_m-1\}}\tilde{\mathcal{A}}^i_{nmj}$, one has that

$$O_{\tau^j_{nm}+(I_m-1)L+1} \leq \Big(\frac{1}{\kappa} + 1\Big)\log T \log\log T,$$

and on the other hand,

$$\tilde{O}_{\tau^j_{nm}+(I_m-1)L+1} \leq \frac{1}{\kappa}\log T \log\log T$$

by the conditional inequality definition in (EC.33).

Therefore, in order for $\hat{O}_t$ and O_t to couple during a period $t \in \{\tau_{nm}^j + (I_m - 1)L + 2, \ldots, \tau_{nm}^j + I_m L\}$, a sufficient condition is

$$\sum_{t=\tau_{nm}^j+(I_m-1)L+1}^{\tau_{nm}^j+I_mL-1} (D_t - Q_n^j) \geq \left(\frac{1}{\kappa} + 1\right) \log T \log\log T. \tag{EC.34}$$

Hence, one has

$$\begin{aligned}
&\mathbb{P}\Big(O_{\tau_{nm}^j+I_mL} = \tilde{O}_{\tau_{nm}^j+I_mL} \Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}_{nmj}^i, \tilde{O}_{\tau_{nm}^j+(I_m-1)L+1} \leq \frac{1}{\kappa}\log T \log\log T\Big) \\
&\geq \mathbb{P}\Big(\sum_{t=\tau_{nm}^j+(I_m-1)L+1}^{\tau_{nm}^j+I_mL-1} (D_t - Q_n^j) \geq \left(\frac{1}{\kappa}+1\right)\log T\log\log T \\
&\qquad\qquad\qquad \Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}_{nmj}^i, \tilde{O}_{\tau_{nm}^j+(I_m-1)L+1} \leq \frac{1}{\kappa}\log T\log\log T\Big) \\
&\geq \mathbb{P}\Big(\sum_{t=\tau_{nm}^j+(I_m-1)L+1}^{\tau_{nm}^j+I_mL-1} (D_t - Q_n^j) - (L-1)(\mu - Q_n^j) \geq \left(\frac{1}{\kappa}+1\right)\log T\log\log T - (L-1)(\mu - Q^h) \\
&\qquad\qquad\qquad \Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}_{nmj}^i, \tilde{O}_{\tau_{nm}^j+(I_m-1)L+1} \leq \frac{1}{\kappa}\log T\log\log T\Big) \\
&\geq \mathbb{P}\Big(\sum_{t=\tau_{nm}^j+(I_m-1)L+1}^{\tau_{nm}^j+I_mL-1} (D_t - Q_n^j) - (L-1)(\mu - Q_n^j) \geq -2\sigma\log T(\log\log T)^{\frac{3}{2}} \\
&\qquad\qquad\qquad \Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}_{nmj}^i, \tilde{O}_{\tau_{nm}^j+(I_m-1)L+1} \leq \frac{1}{\kappa}\log T\log\log T\Big) \\
&\geq 1 - \frac{1}{T^3},
\end{aligned} \tag{EC.35}$$

where the last inequality is implied by Lemma EC.4.

Plug (EC.35) back into (EC.33), one has that

$$\mathbb{P}\Big(O_{\tau_{nm}^j+I_mL} = \tilde{O}_{\tau_{nm}^j+I_mL} \Big| \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}_{nmj}^i\Big) \geq 1 - \frac{2}{T^3}. \tag{EC.36}$$

Combining (EC.32) and (EC.36), one obtains a lower bound for (EC.31) as

$$\mathbb{P}(\tilde{\mathcal{A}}_{nmj}^{I_m} | \cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}_{nmj}^i) \geq 1 - \frac{3}{T^3}. \tag{EC.37}$$

Plugging (EC.37) into (EC.30), one has that for any $n, m, j \in \{l, c, r\}$,

$$\mathbb{P}(\cap_{i\in\{1,\ldots,I_m\}} \tilde{\mathcal{A}}_{nmj}^i) \geq \left(1 - 3e^{-\log T\log\log T}\right)\mathbb{P}(\cap_{i\in\{1,\ldots,I_m-1\}} \tilde{\mathcal{A}}_{nmj}^i),$$

and continue the same procedure, one will obtain

$$\begin{aligned}
\mathbb{P}(\tilde{\mathcal{B}}_{nmj}) &= \mathbb{P}(\cap_{i\in\{1,\ldots,I_m\}} \tilde{\mathcal{A}}_{nmj}^i) \\
&\geq \left(1 - 3(I_m - 1)e^{-\log T\log\log T}\right)\mathbb{P}(\tilde{\mathcal{A}}_{nmj}^1) \\
&\geq \left(1 - 3(I_m - 1)e^{-\log T\log\log T}\right)\mathbb{P}(\tilde{\mathcal{A}}_{nmj}^1 | \tilde{\mathcal{B}}_{\Gamma(nmj)})\mathbb{P}(\tilde{\mathcal{B}}_{\Gamma(nmj)}).
\end{aligned} \tag{EC.38}$$

$\mathbb{P}(\tilde{\mathcal{A}}^1_{nmj}|\tilde{\mathcal{B}}_{\Gamma(nmj)})$ in (EC.12) can be bounded as

$$\mathbb{P}(\tilde{\mathcal{A}}^1_{nmj}\big|\tilde{\mathcal{B}}_{\Gamma(nmj)}) = \mathbb{P}\Big(\tilde{O}_{\tau^j_{nm}+L} \le \frac{1}{\kappa}\log T\log\log T, O_{\tau^j_{nm}+L} = \tilde{O}_{\tau^j_{nm}+L}\Big|\tilde{\mathcal{B}}_{\Gamma(nmj)}\Big). \tag{EC.39}$$

By Lemmas EC.2 and EC.3 and the fact that $\tilde{O}_{\tau^j_{nm}+L}$ is independent of $\tilde{\mathcal{B}}_{\Gamma(nmj)}$, one has

$$\mathbb{P}\Big(\tilde{O}_{\tau^j_{nm}+L} \le \frac{1}{\kappa}\log T\log\log T\Big|\tilde{\mathcal{B}}_{\Gamma(nmj)}\Big) > 1 - \frac{1}{T^3}. \tag{EC.40}$$

In order to bound $\mathbb{P}\Big(O_{\tau^j_{nm}+L} = \tilde{O}_{\tau^j_{nm}+L}|\tilde{\mathcal{B}}_{\Gamma(nmj)}\Big)$, we start by analyzing the behaviors of stochastic processes $\{O\}^{\tau^j_{nm}+L}_{\tau^j_{nm}+1}$ and $\{\tilde{O}\}^{\tau^j_{nm}+L}_{\tau^j_{nm}+1}$.

During periods $t = \tau^j_{nm}+1,\dots,\tau^j_{nm}+l-1$, O_t and $\tilde{O}_t$ follow different recursions as

$$O_{t+1} = ((O_t + S^{ji}_{nm} - d_t)^+ + q^R_{t+1-l} - S^{ji}_{nm})^+$$

as opposed to

$$\tilde{O}_{t+1} = (\tilde{O}_t + Q^j_n - D_t)^+ \text{ with } \tilde{O}_{\tau^j_{nm}+1} = O_\infty(Q^j_n).$$

Note that $q^R_{t+1-l} \ne Q^j_n$ due to the l periods delay from the regular source.

During periods $t = \tau^j_{nm}+l,\dots,\tau^j_{nm}+L$, they follow the same recursion, since the system of O_t starts to receive Q^j_n every period from the regular source, and because $Q^j_n \le S^{ji}_{nm}$, we have

$$O_{t+1} = (O_t + Q^j_n - D_t)^+,$$

and by definition, one has

$$\tilde{O}_{t+1} = (\tilde{O}_t + Q^j_n - D_t)^+.$$

Under $\tilde{\mathcal{B}}_{\Gamma(nmj)}$, it holds that $O_{\tau^j_{nm}+1} \le \frac{1}{\kappa}\log T\log\log T$, then $O_{\tau^j_{nm}+l} \le O_{\tau^j_{nm}+1} + K_{48} \le (\frac{1}{\kappa}+1)\log T\log\log T$. On the other hand, $\mathbb{P}(\tilde{O}_{\tau^j_{nm}+l} \le \frac{1}{\kappa}\log T\log\log T) \ge 1 - e^{-\log T\log\log T}$.

The stochastic processes O_t and $\tilde{O}_t$ will couple during a period $t \in \{\tau^j_{nm}+l+1,\dots,\tau^j_{nm}+L\}$, if there exists $t_0 \in \{\tau^j_{nm}+l,\dots,\tau^j_{nm}+L-1\}$ such that

$$\sum_{t=\tau^j_{nm}+l}^{t_0}(D_t - Q^j_n) \ge \max\{O_{\tau^j_{nm}+l}, \tilde{O}_{\tau^j_{nm}+l}\} = (\frac{1}{\kappa}+1)\log T\log\log T.$$

Therefore, one has

$$
\begin{aligned}
&\mathbb{P}\Big(O_{\tau^j_{nm}+L}=\tilde{O}_{\tau^j_{nm}+L}\Big|\tilde{\mathcal{B}}_{\Gamma(nmj)}\Big)\\
\geq\;&\mathbb{P}\Big(O_{\tau^j_{nm}+L}=\tilde{O}_{\tau^j_{nm}+L}\Big|\tilde{\mathcal{B}}_{\Gamma(nmj)},\ \tilde{O}_{\tau^j_{nm}+l}\leq\frac{1}{\kappa}\log T\log\log T\Big)\times\mathbb{P}\big(\tilde{O}_{\tau^j_{nm}+l}\leq\frac{1}{\kappa}\log T\log\log T\big|\tilde{\mathcal{B}}_{\Gamma(nmj)}\big)\\
\geq\;&\mathbb{P}\Big(O_{\tau^j_{nm}+L}=\tilde{O}_{\tau^j_{nm}+L}\Big|\tilde{\mathcal{B}}_{\Gamma(nmj)},\ \tilde{O}_{\tau^j_{nm}+l}\leq\frac{1}{\kappa}\log T\log\log T\Big)\times(1-\frac{1}{T^3})\\
\geq\;&\mathbb{P}\Big(\sum_{t=\tau^j_{nm}+l}^{\tau^j_{nm}+L-1}(D_t-Q^j_n)\geq(\frac{1}{\kappa}+1)\log T\log\log T\Big|\tilde{\mathcal{B}}_{\Gamma(nmj)},\ \tilde{O}_{\tau^j_{nm}+l}\leq\frac{1}{\kappa}\log T\log\log T\Big)(1-\frac{1}{T^3})\\
\geq\;&\mathbb{P}\Big(\sum_{t=\tau^j_{nm}+l}^{\tau^j_{nm}+L-1}(D_t-Q^j_n)-(L-l)(\mu-Q^j_n)\geq(\frac{1}{\kappa}+1)\log T\log\log T-(L-l)(\mu-Q^h)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\Big|\tilde{\mathcal{B}}_{\Gamma(nmj)},\ \tilde{O}_{\tau^j_{nm}+l}\leq\frac{1}{\kappa}\log T\log\log T\Big)(1-\frac{1}{T^3})\\
\geq\;&\mathbb{P}\Big(\sum_{t=\tau^j_{nm}+l}^{\tau^j_{nm}+L-1}(D_t-Q^j_n)-(L-l)(\mu-Q^j_n)\geq-2\sigma\log T(\log\log T)^{\frac{3}{2}}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\Big|\tilde{\mathcal{B}}_{\Gamma(nmj)},\ \tilde{O}_{\tau^j_{nm}+l}\leq\frac{1}{\kappa}\log T\log\log T\Big)(1-\frac{1}{T^3})\\
\geq\;&1-\frac{2}{T^3},
\end{aligned}
\tag{EC.41}
$$

where the last inequality is implied by Lemma EC.4.

Combining (EC.40) and (EC.41), one obtains a lower bound for (EC.39) as

$$
\mathbb{P}(\tilde{\mathcal{A}}^1_{nmj}|\tilde{\mathcal{B}}_{\Gamma(nmj)})\geq 1-\frac{3}{T^3}. \tag{EC.42}
$$

Plugging (EC.16) into (EC.12), one has that

$$
\mathbb{P}(\tilde{\mathcal{B}}_{nmj})\geq\big(1-\frac{3I_m}{T^3}\big)\mathbb{P}(\tilde{\mathcal{B}}_{\Gamma(nmj)}). \tag{EC.43}
$$

Set $O_1=0$ (or to achieve the same theoretical result, set $\mathbb{P}(\hat{\mathcal{B}}_{\Gamma(111)})=1$), and repeat the above procedure, one will obtain that

$$
\mathbb{P}(\tilde{\mathcal{B}}_{111})\geq 1-\frac{3I_1}{T^3},
$$

then by (EC.43), one has that for any $n,m,j\in\{l,c,r\}$,

$$
\mathbb{P}(\tilde{\mathcal{B}}_{nmj})\geq 1-\frac{3}{T^3}\sum_n\sum_m\sum_j I_m=1-\frac{3}{T^2}.
$$

The proof of Lemma 3 is thus complete. **Q.E.D.**

Proof of Proposition 4. We proceed as the following.

$$|\mathbb{E}[C^{B1}-C^{B2}]| \le \log T(\log\log T)^2\mathbb{E}\left[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\sum_{i=1}^{I_m}\left(G(Q_n^j,S_{nm}^{ji})-G(Q_n^j,S^*(Q_n^j))\right)\right], \tag{EC.44}$$

Based on Theorem 7 in Huh et al. (2009), for any n,m,j and Q_n^j, one has that

$$\begin{aligned}&\mathbb{E}[\sum_{i=1}^{I_m}\left(\mathbb{E}[G(Q_n^j,S_{nm}^{ji})]-G(Q_n^j,S^*(Q_n^j))\right)]\\ &\le (2\max\{h,b-c\}(S^h-S^l))\sqrt{I_m}+(S^h-S^l)\mathbb{E}[\sum_{i=1}^{I_m}|\delta(S_{nm}^{ji})|],\end{aligned} \tag{EC.45}$$

where $\delta(s)=\mathbb{E}[\bigtriangledown_{nm}^{ji}(s)|s]-\frac{\partial G(Q_n^j,S)}{\partial S}\Big|_{S=s}$. Therefore, for any $s\in[S^l,S^h]$, one has

$$\begin{aligned}&|\delta(s)|\\ &\le h|\mathbb{P}(O_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0)-\mathbb{P}(O_\infty(Q_n^j)+s-d_{\tau_{nm}^j+iL}\ge 0)|\\ &\qquad+(b-c)|\mathbb{P}(O_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}<0)-\mathbb{P}(O_\infty(Q_n^j)+s-d_{\tau_{nm}^j+iL}<0)|\\ &= h|\mathbb{P}(O_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0)-\mathbb{P}(\tilde{O}_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0)|\\ &\qquad+(b-c)|\mathbb{P}(O_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}<0)-\mathbb{P}(\tilde{O}_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}<0)|.\end{aligned} \tag{EC.46}$$

In (EC.46), the first term can be bounded as

$$\begin{aligned}&h\big|\mathbb{P}(O_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0)-\mathbb{P}(\tilde{O}_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0)\big|\\ &\le h\big|\mathbb{P}(O_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0\big|\tilde{\mathcal{B}}_{nmj})-\mathbb{P}(\tilde{O}_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0\big|\tilde{\mathcal{B}}_{nmj})\big|\times\mathbb{P}(\tilde{\mathcal{B}}_{nmj})\\ &\quad+h\big|\mathbb{P}(O_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0\big|(\tilde{\mathcal{B}}_{nmj})^C)-\mathbb{P}\big(\tilde{O}_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}\ge 0\big|(\tilde{\mathcal{B}}_{nmj})^C)\big|\times\mathbb{P}((\tilde{\mathcal{B}}_{nmj})^C)\\ &\le 0+\frac{6h}{T^2},\end{aligned} \tag{EC.47}$$

where the last inequality is true because on $\tilde{\mathcal{B}}_{nmj}$, it holds that $O_{\tau_{nm}^j+iL}=\tilde{O}_{\tau_{nm}^j+iL}$, and $\mathbb{P}((\tilde{\mathcal{B}}_{nmj})^C)$ is bounded by Lemma 3.

Similarly, the second term in (EC.46) can be bounded as

$$(b-c)|\mathbb{P}(O_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}<0)-\mathbb{P}(\tilde{O}_{\tau_{nm}^j+iL}+s-d_{\tau_{nm}^j+iL}<0)|\le\frac{6(b-c)}{T^2}. \tag{EC.48}$$

Combining (EC.47) and (EC.48), we have that, for any $s\in[S^l,S^h]$,

$$|\delta(s)|\le\frac{6(h+b-c)}{T^2}, \tag{EC.49}$$

which yields that in (EC.45),

$$\begin{aligned}&\sum_{i=1}^{I_m}\left(\mathbb{E}[G(Q_n^j,S_{nm}^{ji})]-\mathbb{E}[G(Q_n^j,S^*(Q_n^j))]\right)\\ &\le (2\max\{h,b-c\}(S^h-S^l))\sqrt{I_m}+(S^h-S^l)\sum_{i=1}^{I_m}\frac{6(h+b-c)}{T^2} \qquad \text{(EC.50)}\\ &\le (2\max\{h,b-c\}(S^h-S^l))\sqrt{T}+(S^h-S^l)\frac{6(h+b-c)}{T},\end{aligned}$$

and the last inequality is due to $I_m \le T$.

Therefore, in (EC.44) one has that

$$\begin{aligned}
&|\mathbb{E}[C^{B1} - C^{B2}]| \\
&\le \log T (\log\log T)^2 \mathbb{E}\Bigg[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\sum_{i=1}^{I_m}\left(G(Q_n^j, S_{nm}^{ji}) - G(Q_n^j, S^*(Q_n^j))\right)\Bigg] \\
&\le K_{49}\log T(\log\log T)^2\mathbb{E}\Bigg[\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}}\sqrt{T}\Bigg]. \qquad \text{(EC.51)}
\end{aligned}$$

The proof of Proposition 4 is thus complete. **Q.E.D.**

Proof of Lemma 4. We start by reviewing the two *equivalent characterizations* for sub-exponential variables with parameter v, b. For a zero mean random variable X, the following statements are equivalent:

(I) There are nonnegative numbers (v, b) such that

$$\mathbb{E}[e^{\lambda X}] \le e^{\frac{v^2\lambda^2}{2}}, \quad \text{for all } |\lambda| < \frac{1}{b}.$$

(II) There are constants c_1, c_2 such that

$$\mathbb{P}(|X| \ge t) \le c_1 e^{-c_2 t} \quad \text{for all } t > 0.$$

By Lemma 6 in Yuan et al. (2021), if condition (II) is satisfied, then X is sub-exponential with parameters $(\frac{2\sqrt{c_1}}{c_2}, \frac{2}{c_2})$. It can be seen that

$$\begin{aligned}
&|h(\tilde{O}_{\tau_{nm}^j+iL} + \tilde{S}_{nm}^{ji} - D_{\tau_{nm}^j+iL})^+ - (b-c)\min\{D_{\tau_{nm}^j+iL}, \tilde{O}_{\tau_{nm}^j+iL} + \tilde{S}_{nm}^{ji}\} \\
&\qquad - (h\mathbb{E}[(\tilde{O}_{\tau_{nm}^j+iL} + \tilde{S}_{nm}^{ji} - D_{\tau_{nm}^j+iL})^+] - (b-c)\mathbb{E}[\min\{D_{\tau_{nm}^j+iL}, \tilde{O}_{\tau_{nm}^j+iL} + \tilde{S}_{nm}^{ji}\}])| \\
&\le (h+b-c)(\tilde{O}_{\tau_{nm}^j+iL} + S^h) + (h+b-c)(\mathbb{E}[O_\infty(Q_n^j)] + S^h) \\
&\le (h+b-c)(\tilde{O}_{\tau_{nm}^j+iL} + S^h) + (h+b-c)(\frac{\sigma^2}{2(\mu - Q^h)} + S^h), \qquad \text{(EC.52)}
\end{aligned}$$

where the first inequality is because $\tilde{O}_{\tau_{nm}^j+iL}$ follows the same distribution as $O_\infty(Q_n^j)$, and the second inequality comes from the Kingman's inequality. Therefore, it is sufficient to prove that $(h+b-c)(\tilde{O}_{\tau_{nm}^j+iL} + S^h) + (h+b-c)(\frac{\sigma^2}{2(\mu-Q^h)} + S^h)$ is sub-exponential.

Because $\tilde{O}_{\tau_{nm}^j+iL}$ follows the same distribution as $O_\infty(Q_n^j)$, and then by Lemma EC.3, one has that for $x \ge 0$,

$$\mathbb{P}(\tilde{O}_{\tau_{nm}^j+iL} > x) \le e^{-\kappa x},$$

which leads to

$$\mathbb{P}\left((h+b-c)\left(\tilde{O}_{\tau_{nm}^j+iL}+2S^h+\frac{\sigma^2}{2(\mu-Q^h)}\right)>x\right)\le e^{\kappa\left(2S^h+\frac{\sigma^2}{2(\mu-Q^h)}\right)}e^{-\frac{\kappa}{h+b-c}x}.$$

Therefore, $(h+b-c)\left(\tilde{O}_{\tau_{nm}^j+iL}+2S^h+\frac{\sigma^2}{2(\mu-Q^h)}\right)$ is sub-exponential with parameter $\left(\frac{2(h+b-c)e^{\frac{1}{2}\kappa\left(2S^h+\frac{\sigma^2}{2(\mu-Q^h)}\right)}}{\kappa},\frac{2(h+b-c)}{\kappa}\right)$.

This implies that

$$\begin{aligned}&h(\tilde{O}_{\tau_{nm}^j+iL}+\tilde{S}_{nm}^{ji}-D_{\tau_{nm}^j+iL})^+-(b-c)\min\{D_{\tau_{nm}^j+iL},\tilde{O}_{\tau_{nm}^j+iL}+\tilde{S}_{nm}^{ji}\}\\&\qquad-(h\mathbb{E}[(\tilde{O}_{\tau_{nm}^j+iL}+\tilde{S}_{nm}^{ji}-d_{\tau_{nm}^j+iL})^+]-(b-c)\mathbb{E}[\min\{D_{\tau_{nm}^j+iL},\tilde{O}_{\tau_{nm}^j+iL}+\tilde{S}_{nm}^{ji}\}])\end{aligned}$$

is also sub-exponential with parameter $\left(\frac{2(h+b-c)e^{\frac{1}{2}\kappa\left(2S^h+\frac{\sigma^2}{2(\mu-Q^h)}\right)}}{\kappa},\frac{2(h+b-c)}{\kappa}\right)$.

The proof of Lemma 4 is thus complete. **Q.E.D.**

In order to prove Lemma 6, we first provide Lemma A1 and Lemma A2 in the following.

LEMMA A1. **(Lemma 1 in Agarwal et al. 2011)** *Suppose that the event $\mathcal{M}$ holds. If epoch n ends in round m, then the interval $[l_{n+1},r_{n+1}]$ contains every $Q\in[l_n,r_n]$ such that $\mathbb{E}[G(Q,S^*(Q))]\le\mathbb{E}[G(Q^*,S^*(Q^*))]+\gamma_m$. In particular, $Q^*\in[l_n,r_n]$ for all epochs n.*

Proof of Lemma A1. Under the learning algorithm, epoch n ends in round m iff one of the following two cases happen:

(1) $\max\{\text{LB}_{\gamma_m}(Q_n^l),\text{LB}_{\gamma_m}(Q_n^r)\}\ge\min\{\text{UB}_{\gamma_m}(Q_n^l),\text{UB}_{\gamma_m}(Q_n^r)\}+\gamma_m$

(2) $\max\{\text{LB}_{\gamma_m}(Q_n^l),\text{LB}_{\gamma_m}(Q_n^r)\}\ge\text{UB}_{\gamma_m}(Q_n^c)+\gamma_m$.

Suppose Case (1) happens. This means that either $\text{LB}_{\gamma_m}(Q_n^l)\ge\text{UB}_{\gamma_m}(Q_n^r)+\gamma_m$ or $\text{LB}_{\gamma_m}(Q_n^r)\ge\text{UB}_{\gamma_m}(Q_n^l)+\gamma_m$. Consider the former case, and the argument for the latter case is analogous. Since the event $\mathcal{M}$ holds, this implies that

$$\mathbb{E}[G(Q_n^l,S^*(Q_n^l))]\ge\mathbb{E}[G(Q_n^r,S^*(Q_n^r))]+\gamma_m.$$

Since $\mathbb{E}[G]$ is convex, we conclude that every $Q\in[l_n,l_{n+1}]=[l_n,Q_n^l]$ has $\mathbb{E}[G(Q,S^*(Q))]\ge\mathbb{E}[G(Q^*,S^*(Q^*))]+\gamma_m$.

Now suppose epoch n terminates in round m via Case (2). This means

$$\max\{\text{LB}_{\gamma_m}(Q_n^l),\text{LB}_{\gamma_m}(Q_n^r)\}\ge\text{UB}_{\gamma_m}(Q_n^c)+\gamma_m.\tag{EC.53}$$

Suppose $\text{LB}_{\gamma_m}(Q_n^l)\ge\text{LB}_{\gamma_m}(Q_n^r)$, and the argument for the case $\text{LB}_{\gamma_m}(Q_n^l)<\text{LB}_{\gamma_m}(Q_n^r)$ is analogous. The above inequality implies

$$\mathbb{E}[G(Q_n^l,S^*(Q_n^l))]\ge\mathbb{E}[G(Q_n^c,S^*(Q_n^c))]+\gamma_m.$$

By the same argument in Case (1), it can be shown that every $Q \in [l_n, l_{n+1}] = [l_n, Q_n^l]$ has $\mathbb{E}[G(Q, S^*(Q))] \geq \mathbb{E}[G(Q^*, S^*(Q^*))] + \gamma_m$.

The fact that every $Q \in [l_n, r_n]$ for all epochs n follows by induction. **Q.E.D.**

LEMMA A2. **(Lemma 2 in Agarwal et al. 2011)** *Suppose event $\mathcal{M}$ holds. For a given epoch n, if it continues from round m to round $m+1$, then one has that for $j \in \{l, c, r\}$,*

$$\mathbb{E}[G(Q_n^j, S^*(Q_n^j))] \leq \mathbb{E}[G(Q^*, S^*(Q^*))] + 12\gamma_m. \tag{EC.54}$$

Proof of Lemma A2. The algorithm continues from round m to $m+1$ if and only if

$$\max\{\mathrm{LB}_{\gamma_m}(Q_n^l), \mathrm{LB}_{\gamma_m}(Q_n^r)\} < \min\{\mathrm{UB}_{\gamma_m}(Q_n^l), \mathrm{UB}_{\gamma_m}(Q_n^r)\} + \gamma_m. \tag{EC.55}$$

and

$$\max\{\mathrm{LB}_{\gamma_m}(Q_n^l), \mathrm{LB}_{\gamma_m}(Q_n^r)\} < \mathrm{UB}_{\gamma_m}(Q_n^c) + \gamma_m. \tag{EC.56}$$

This implies that $G(Q_n^l, S^*(Q_n^l))$, $G(Q_n^c, S^*(Q_n^c))$ and $G(Q_n^r, S^*(Q_n^r))$ are contained in an interval of width at most $3\gamma_m$.

By Lemma A1, we have $Q^* \in [l_n, r_n]$. Assume $Q^* \leq Q_n^c$. (The case $Q^* > Q_n^c$ is analogous.) There exists $\nu \geq 0$ such that $Q^* = Q_n^l + \nu(Q_n^c - Q_n^r)$, so

$$Q_n^c = \frac{1}{1+\nu} Q^* + \frac{\nu}{1+\nu} Q_n^r.$$

Note that $\nu \leq 2$ because $|Q_n^c - l_n| = w_n/2$ and $Q_n^r - Q_n^c = w_n/4$, so

$$\nu = \frac{|Q^* - Q_n^c|}{|Q_n^r - Q_n^c|} \leq \frac{|l_n - Q_n^c|}{|Q_n^r - Q_n^c|} = \frac{w_n/2}{w_n/4} = 2. \tag{EC.57}$$

By convexity,

$$\begin{aligned}
\mathbb{E}[G(Q^*, S^*(Q^*))] &\geq (1+\nu)(\mathbb{E}[G(Q_n^c, S^*(Q_n^c))] - \frac{\nu}{1+\nu}\mathbb{E}[G(Q_n^r, S^*(Q_n^r))]) \\
&= \mathbb{E}[G(Q_n^r, S^*(Q_n^r))] + (1+\nu)(\mathbb{E}[G(Q_n^c, S^*(Q_n^c))] - \mathbb{E}[G(Q_n^r, S^*(Q_n^r))]) \\
&\geq \mathbb{E}[G(Q_n^r, S^*(Q_n^r))] - (1+\nu)|\mathbb{E}[G(Q_n^c, S^*(Q_n^c))] - \mathbb{E}[G(Q_n^r, S^*(Q_n^r))]| \\
&\geq \mathbb{E}[G(Q_n^r, S^*(Q_n^r))] - (1+\nu)3\gamma_m \\
&\geq \mathbb{E}[G(Q_n^r, S^*(Q_n^r))] - 9\gamma_m.
\end{aligned}$$

Then we conclude that for each $j \in \{l, c, r\}$,

$$\mathbb{E}[G(Q_n^j, S^*(Q_n^j))] \leq \mathbb{E}[G(Q_n^r, S^*(Q_n^r))] + 3\gamma_m \leq \mathbb{E}[G(Q^*, S^*(Q^*))] + 12\gamma_m.$$

The proof of Lemma A2 is thus complete. **Q.E.D.**

Proof of Lemma 6. If $M_n = 1$, then $\mathbb{E}[G(Q,S^*(Q))] - \mathbb{E}[G(Q^*,S^*(Q^*))] \le |Q-Q^*| \le \theta(Q^h - Q^l)$ for each $Q_n^j, j\in\{l,c,r\}$. Therefore, for the given n, one has

$$\sum_{m=1}\sum_{j\in\{l,c,r\}} \big(\mathbb{E}[G(Q_n^j,S^*(Q_n^j))] - \mathbb{E}[G(Q^*,S^*(Q^*))]\big)/(\gamma_m)^2 \le \frac{6\theta(Q^h-Q^l)}{\gamma_{M_n}}.$$

Now assume $M_n \ge 2$. By Lemma A2, one has for each $j\in\{l,c,r\}$,

$$\mathbb{E}[G(Q_n^j,S^*(Q_n^j))] \le \mathbb{E}[G(Q^*,S^*(Q^*))] + 12\gamma_{M_n-1}.$$

Therefore, one has

$$\begin{aligned}
&\sum_{m=1}^{M_n}\sum_{j\in\{l,c,r\}} \big(\mathbb{E}[G(Q_n^j,S^*(Q_n^j))] - \mathbb{E}[G(Q^*,S^*(Q^*))]\big)/(\gamma_m)^2\\
\le{}& \sum_{m=1}^{M_n} 36\gamma_{M_n-1}/(\gamma_m)^2\\
\le{}& \sum_{m=1}^{M_n-1} 36\gamma_m/(\gamma_m)^2 + 72\gamma_{M_n}/(\gamma_{M_n})^2\\
\le{}& 36/\gamma_{M_n} + 72/\gamma_{M_n}\\
={}& 108/\gamma_{M_n}.
\end{aligned}$$

Then note that $\gamma_m \ge (T/(\log T)^2(\log\log T)^3)^{-1/2}$ at all epochs and rounds. Indeed, if $\gamma_m \le (T/(\log T)^2(\log\log T)^3)^{-1/2}$, then the mth round will contain more than $\log T(\log\log T)^2 I_m = (\log T)^2(\log\log T)^3/(\gamma_m)^2 = T$ periods. Hence we set $\gamma_{\min} = (T/(\log T)^2(\log\log T)^3)^{-1/2}$ and one has

$$108/\gamma_{M_n} \le 108/\gamma_{\min}.$$

The proof of Lemma 6 is thus complete. **Q.E.D.**

Proof of Proposition 6. One has that

$$\begin{aligned}
&\left|\mathbb{E}[C^{B3} - \sum_{t=1}^{T} C_t^{(Q^*,S^*)}]\right|\\
={}& \sum_{t=1}^{T}\left|\mathbb{E}[-cQ^* + h(S^* + O_t^{(Q^*,S^*)} - D_t)^+ - (b-c)\min\{S^* + O_t^{(Q^*,S^*)}, D_t\}] - \mathbb{E}[G(Q^*,S^*)]\right| + c\mathbb{E}[x_{T+1}^{(Q^*,S^*)}]\\
\le{}& \sum_{t=1}^{T} h\left|\mathbb{E}[(S^* + O_t^{(Q^*,S^*)} - D_t)^+] - \mathbb{E}[(S^* + O_\infty(Q^*) - D_t)^+]\right|\\
&+\sum_{t=1}^{T}(b-c)\left|\mathbb{E}[(D_t - S^* - O_t^{(Q^*,S^*)})^+] - \mathbb{E}[(D_t - S^* - O_\infty(Q^*))^+]\right| + c\mathbb{E}[x_{T+1}^{(Q^*,S^*)}].
\end{aligned}\qquad\text{(EC.58)}$$

To bound the first term in (EC.58), similar to proving (EC.23), one has that

$$\sum_{t=1}^{T} h\left|\mathbb{E}[(S_{nm}^{ji} + O_t^{(Q^*,S^*)} - D_t)^+] - \mathbb{E}[(S_{nm}^{ji} + O_\infty(Q_n^j) - D_t)^+]\right| \le K_{50}(\log T)^2(\log\log T)^4. \qquad\text{(EC.59)}$$

The second term in (EC.58) can be bounded similar to deriving (EC.25) as

$$\sum_{t=1}^{T}(b-c)\left|\mathbb{E}[(D_t-S_{nm}^{ji}-O_t^{(Q^*,S^*)})^+]-\mathbb{E}[(D_t-S_{nm}^{ji}-O_\infty(Q_n^j))^+]\right|\le K_{51}(\log T)^2(\log\log T)^4. \tag{EC.60}$$

The third term in (EC.58) is bounded as the following,

$$\begin{aligned}\mathbb{E}[x_{T+1}]=\mathbb{E}[x_{T+1}|O_{T+1}^{(Q^*,S^*)}\le\frac{1}{\kappa}\log T\log\log T]\mathbb{P}(O_{T+1}^{(Q^*,S^*)}\le\frac{1}{\kappa}\log T\log\log T)\\+\mathbb{E}[x_{T+1}|O_{T+1}^{(Q^*,S^*)}>\frac{1}{\kappa}\log T\log\log T]\mathbb{P}(O_{T+1}^{(Q^*,S^*)}>\frac{1}{\kappa}\log T\log\log T)\\\le K_{52}\log T\log\log T+Q^hT\times\frac{3}{T^3},\end{aligned} \tag{EC.61}$$

where the last inequality is valid because $\mathbb{P}(O_{T+1}>\frac{1}{\kappa}\log T\log\log T)\le\frac{3}{T^3}$ by Lemmas EC.3 and 1.

Combining (EC.59), (EC.60) and (EC.61), and plug them in (EC.58), one has

$$|C^{B3}-\mathbb{E}[\sum_{t=1}^{T}C_t^{(Q^*,S^*)}]|\le K_{53}(\log T)^2(\log\log T)^4. \tag{EC.62}$$

The proof of Proposition 6 is thus complete. **Q.E.D.**

Proof of Lemma 7. Recall that $\gamma_{\min}=(T/(\log T)^2(\log\log T)^3)^{-1/2}\le\gamma_m$ for any m. By definition, $\gamma_m=2^{-m}$, therefore, one has

$$m\le\frac{1}{2}\log_2\frac{T}{(\log T)^2(\log\log T)^3}.$$

The bound of n is based on Lemma 4 in Agarwal et al. (2011). The rest of the analyses in this proof is conditional on the event $\mathcal{M}$. Define the interval $I_{\min}:=[Q^*-\frac{1}{\theta}\gamma_{\min},Q^*+\frac{1}{\theta}\gamma_{\min}]$ which has width $2\gamma_{\min}$. For any $Q\in I_{\min}$, $\mathbb{E}[G(Q,S^*(Q))]-\mathbb{E}[G(Q^*,S^*(Q^*))]\le\theta|Q-Q^*|\le\gamma_{\min}$. Moreover, for any epoch $\breve{n}$ which ends in round $\breve{m}$, $\gamma_{\min}\le\gamma_{\breve{m}}$, by definition and therefore by Lemma A1,

$$I_{\min}\subset\{Q\in[Q^l,Q^h]:\mathbb{E}[G(Q,S^*(Q))]\le\mathbb{E}[G(Q^*,S^*(Q^*))]+\gamma_{\breve{m}}\}\subset[l_{\breve{n}+1},r_{\breve{n}+1}].$$

This implies that $2\gamma_{\min}\le r_{\breve{n}+1}-l_{\breve{n}+1}=w_{\breve{n}+1}$. Furthermore, by the definitions of $r_{\breve{n}+1},l_{\breve{n}+1},w_{\breve{n}+1}$ in the algorithm, it follows that

$$w_{\breve{n}+1}\le\frac{3}{4}w_{\breve{n}}$$

for any $\breve{n}$. Therefore, we conclude that

$$2\gamma_{\min}\le w_{n+1}\le\left(\frac{3}{4}\right)^n w_1=(Q^h-Q^l)\left(\frac{3}{4}\right)^n,$$

which renders that

$$n\le\log_{4/3}\left(\frac{T}{4(Q^h-Q^l)(\log T)^2(\log\log T)^3}\right).$$

Because $\mathbb{P}(\mathcal{M}) \geq 1 - \frac{4}{T}$, then one has

$$\mathbb{P}\left(n \leq \log_{4/3}\left(\frac{T}{4(Q^h - Q^l)(\log T)^2(\log\log T)^3}\right)\right) \geq 1 - \frac{4}{T}.$$

The proof of Lemma 7 is thus complete. **Q.E.D.**

EC.3. Standard Technical Results

The section contains the previously known technical results. We omit the proofs for the most standard results and only present the proofs for the less standard ones, for the sake of completeness.

Definition EC.1 (**Adjustment Coefficient**). For a random variable X, let the moment generating function be $\lambda(z) = \mathbb{E}[e^{zX}]$. The adjustment coefficient is defined as the positive solution of $\lambda(z) = 1$.

Lemma EC.1 **(Existence and Uniqueness of Adjustment Coefficient)**. *For a random variable X, the adjustment coefficient of X exists and is unique, if the following conditions hold:*

1. *For some constant $z > 0$, the moment generating function $\lambda(z) = \mathbb{E}[e^{zX}] < \infty$;*
2. $\mathbb{E}[X] < 0$;
3. $\mathbb{P}(X > 0) > 0$;
4. $\lim_{z\to a^-} \lambda(z) \geq 1$ *for* $a := \sup\{z \geq 0, \lambda(z) < \infty\}$.

Lemma EC.2 **(Boundedness of Adjustment Coefficient)**. *For any $Q \in [Q^l, Q^h]$, the adjustment coefficient of $Q - D$ exists and is unique, and is bounded from below by a constant $\kappa > 0$.*

Proof of Lemma EC.2. In Lemma EC.1, let $X = Q - D$. Condition 1 in Lemma EC.1 is satisfied because $Q - D \leq Q^h$. Condition 2 is satisfied because $Q^h < \mu$. Conditions 3 and 4 are implied by Assumption 1 (3) and (4). Therefore, for any $Q \in [Q^l, Q^h]$, the adjustment coefficient of $Q - D$ exists and is the unique z as the positive solution of

$$\lambda(z, Q) = \mathbb{E}[e^{z(Q-D)}] = 1.$$

Moreover, for any $Q \in [Q^l, Q^h]$, it can be seen that $\lambda(z, Q)$ is convex in z, with $\lambda(0, Q) = 1$ and $\lambda'_z(0, Q) = Q - \mu \leq Q^h - \mu < 0$. $\lambda''_z(z, Q) = \mathbb{E}[e^{z(Q-D)}(Q - D)^2]$, and we will show that $\lambda''_z(z, Q)$ is continuous on $[0, 2] \times [Q^l, Q^h]$. This can be shown by the Dominated Convergence Theorem, because

$$|e^{z(Q-D)}(Q - D)^2| \leq 2e^{2Q^h}((Q^h)^2 + D^2),$$

and $\mathbb{E}[2e^{2Q^h}((Q^h)^2 + D^2)]$ is finite if $\mathbb{E}[D^2] < \infty$, which is implied by Assumption 1 (2).

So there exists $0 < \beta < \infty$ such that $\lambda''_z(z, Q) \leq \beta$ for any $z \in [0, 2]$ and $Q \in [Q^l, Q^h]$, which yields that for any $Q \in [Q^l, Q^h]$, the adjustment coefficient of $Q - D$ is lower bounded by $\kappa \geq \min\{\frac{\mu - Q^h}{\beta}, 2\} > 0$. **Q.E.D.**

Lemma EC.3 **(Lundberg's Inequality)**. *Recall that $O_\infty(Q)$ is the stationary distribution for the stochastic process $O_{t+1} = (O_t + Q - D_t)^+$, then it holds that*

$$\mathbb{P}(O_\infty(Q) > a) \le e^{-\rho a}, \tag{EC.63}$$

for $a \ge 0$ and ρ is the adjustment coefficient of the random variable $Q - D$.

Lemma EC.4 **(Concentration Inequality for Independent Samples)**. *Let ξ_t be i.i.d. random variables with mean 0 and standard deviation σ. If the moment generating function of ξ_1 around 0 is finite, i.e., there exists a constant $\rho > 0$, such that for any $s \in (-\rho, \rho)$ it holds that*

$$\mathbb{E}[e^{s\xi_1}] < +\infty, \tag{EC.64}$$

then one has

$$\mathbb{P}\left(\frac{1}{L}\sum_{t=1}^{L} \xi_t \ge -2\sigma(\log\log T)^{-1/2}\right) > 1 - e^{-\log T \log\log T}. \tag{EC.65}$$

Proof of Lemma EC.4. Under the condition of (EC.64), and following similar lines of proving Lemma A2 in Besbes and Zeevi (2015), (EC.65) can be proved as the following. For $s \in (-\rho, \rho)$, define

$$\Psi(s) = \log \mathbb{E}[e^{s\xi_1}].$$

For $x > 0$ and $s \in [0, \rho)$, by Markov's inequality one has

$$\mathbb{P}\left(\frac{1}{L_i}\sum_{t=1}^{L_i} \xi_t > x\right) \le e^{L_i(\Psi(s) - sx)}.$$

Let $x = 2\sigma(\log\log T)^{-1/2}$ and $s^* = x/\sigma^2$. Apply Taylor's expansion to the third order around 0, one has

$$\Psi(s^*) = \frac{1}{2}\sigma^2(s^*)^2 + \frac{1}{6}\Psi'''(s)(s^*)^3,$$

where $s \in [0, s^*]$. Therefore one has

$$\Psi(s^*) - s^* x \le -\frac{x^2}{2\sigma^2} + K_{60}\frac{x^3}{\sigma^6},$$

where $K_{60} = \max_{s\in[0,s^*]} \Psi'''(s)$. This yields

$$\mathbb{P}\left(\frac{1}{L}\sum_{t=1}^{L} \xi_t > 2\sigma(\log\log T)^{-1/2}\right) \le e^{L(-\frac{x^2}{2\sigma^2} + K_{60}\frac{x^3}{\sigma^6})} \le e^{L(-\frac{x^2}{4\sigma^2})} = e^{-\log T \log\log T},$$

and similarly one has

$$\mathbb{P}\left(\frac{1}{L}\sum_{t=1}^{L} \xi_t < -2\sigma(\log\log T)^{-1/2}\right) \le e^{L(-\frac{x^2}{2\sigma^2} + K_{60}\frac{x^3}{\sigma^6})} \le e^{L(-\frac{x^2}{4\sigma^2})} = e^{-\log T \log\log T},$$

The proof is thus complete. **Q.E.D.**

Lemma EC.5 **(Azuma's Inequality)**. *Let* $\{(D_k, \mathcal{F}_k)\}_{k=1}^{\infty}$ *be a martingale difference sequence, and suppose that for any* $|\lambda| < \frac{1}{b_k}$*, we have* $\mathbb{E}[e^{\lambda D_k} | \mathcal{F}_{k-1}] \le e^{\lambda^2 v_k^2 / 2}$ *almost surely. Then the sum* $\sum_{k=1}^n D_k$ *is sub-exponential with parameters* $(\sqrt{\sum_{k=1}^n v_k^2}, b_*)$*, where* $b_* = \max_{k\in\{1,\dots,n\}} b_k$*. Consequently, for all* $\xi \ge 1$*,*

$$\mathbb{P}\left(\left|\sum_{k=1}^n D_k\right| \ge \xi\right) \le \begin{cases} 2e^{-\frac{\xi^2}{2\sum_{k=1}^n v_k^2}} & \text{if } 0 \le \xi \le \frac{\sum_{k=1}^n v_k^2}{b_*}, \text{ and} \\ 2e^{-\frac{\xi}{2b_*}} & \text{if } \xi > \frac{\sum_{k=1}^n v_k^2}{b_*}. \end{cases}$$

Lemma EC.6 **(High Probability Regret Bound for SGD)**. *Let* $f(\cdot): \mathcal{K} \to \mathbb{R}$ *be a convex function and* $\{z_i\}_{i=1}^n$ *be a sequence generated by the projected stochastic gradient descent with respect to* $f(\cdot)$*, i.e.,*

$$z_1 \in \mathcal{K} \quad \text{and} \quad z_{i+1} = \boldsymbol{Proj}_{\mathcal{K}}[z_i - \eta_i \tilde{\bigtriangledown}_i] \text{ with } \eta_i = \frac{\beta}{c}\frac{1}{\sqrt{i}}, \quad \text{for } i = 1, \dots, n-1,$$

where $\tilde{\bigtriangledown}_i$ *is a stochastic gradient of* f *at* z_i *and* η_i *is the step size in the* i^{th} *iteration. We make the following assumptions:*

(a) The diameter of function domain $\mathcal{K}$ *is bounded by* β*, i.e.,* $\sum_{z_1, z_2 \in \mathcal{K}} \|z_1 - z_2\|_2 \le \beta$*.*

(b) For $i = 1, \dots, n-1$*, the stochastic gradient* $\|\tilde{\bigtriangledown}_i\|_2 \le c$ *almost surely for some constant* $c > 0$*.*

Then we have that for any $\xi > 0$*,*

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n (f(z_i) - f(z^*)) \le \frac{3\beta c}{2\sqrt{n}} + \xi\right) \ge 1 - 2e^{-\frac{n\xi^2}{2c^2\beta^2}}.$$

EC.4. Figures for Numerical Studies in Section 6

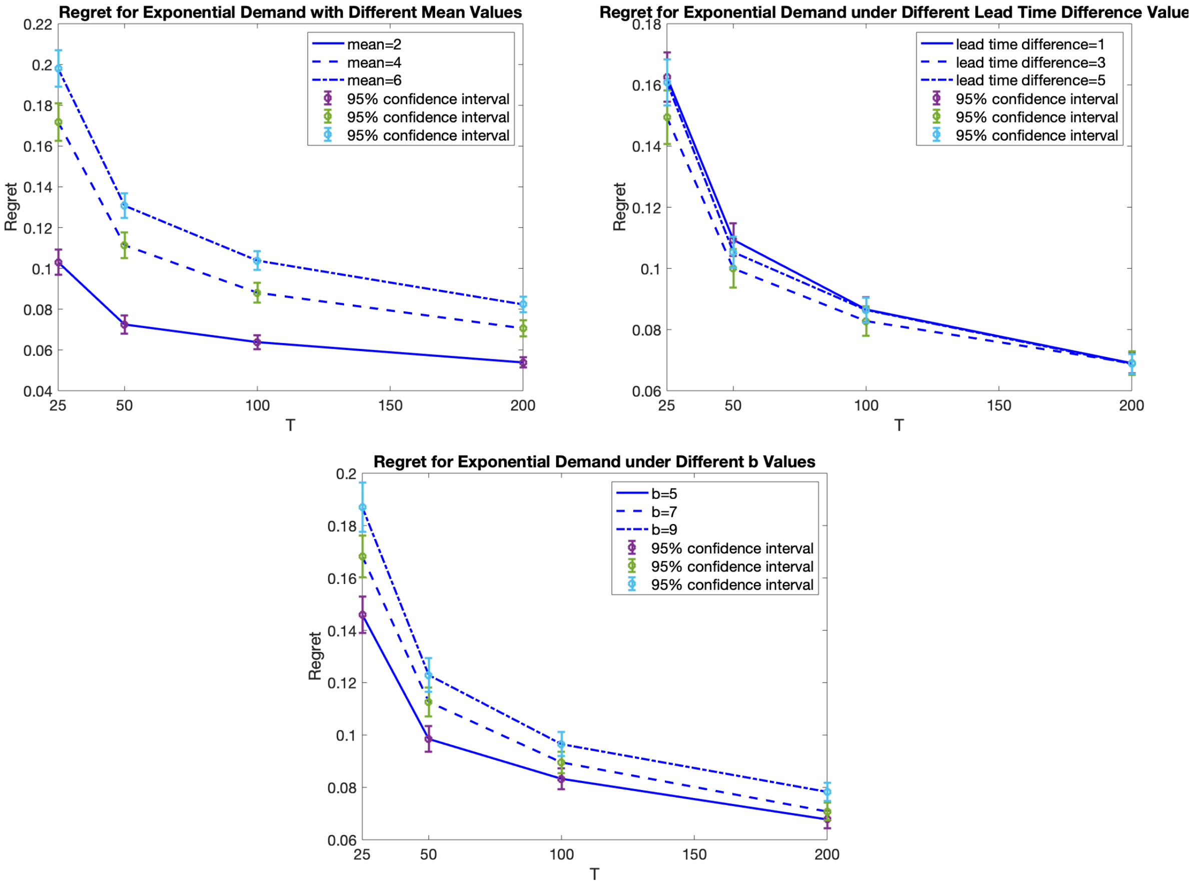


Figure EC.1 Regret for Exponential Demand

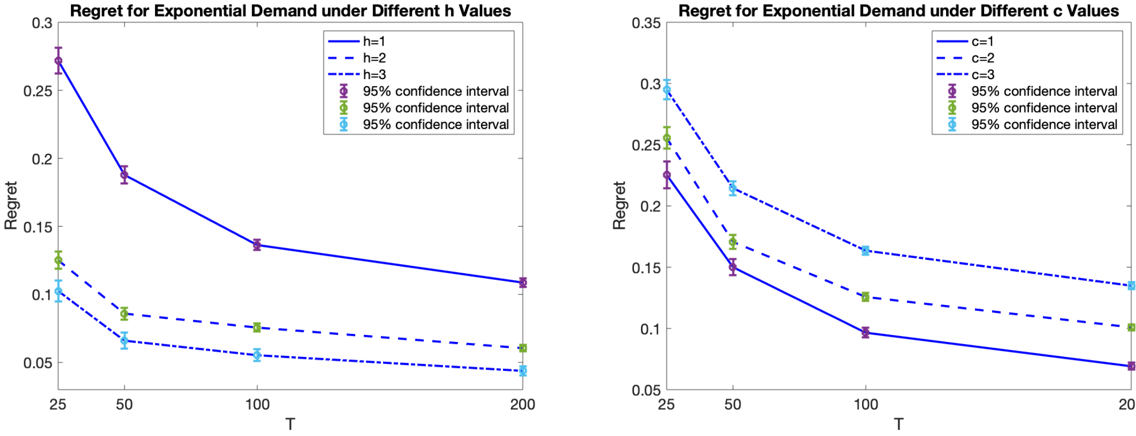
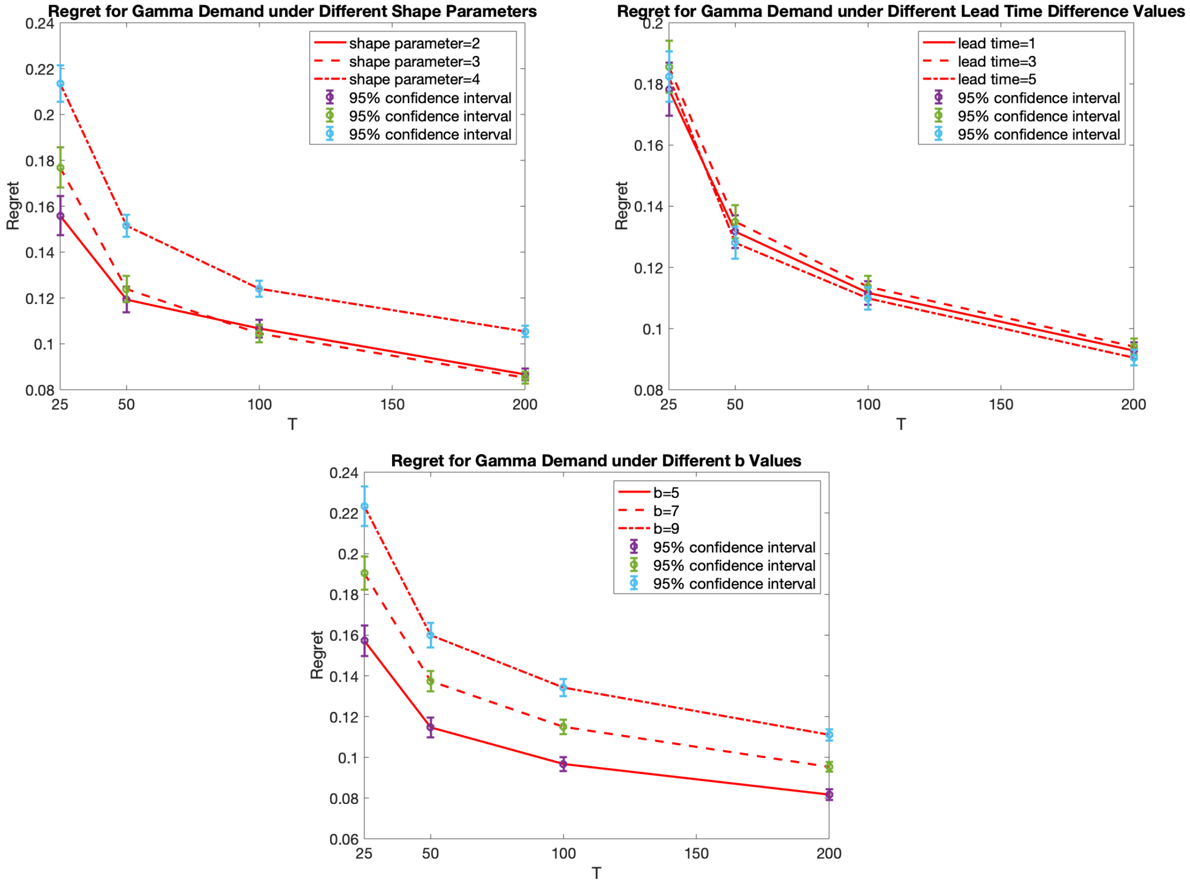


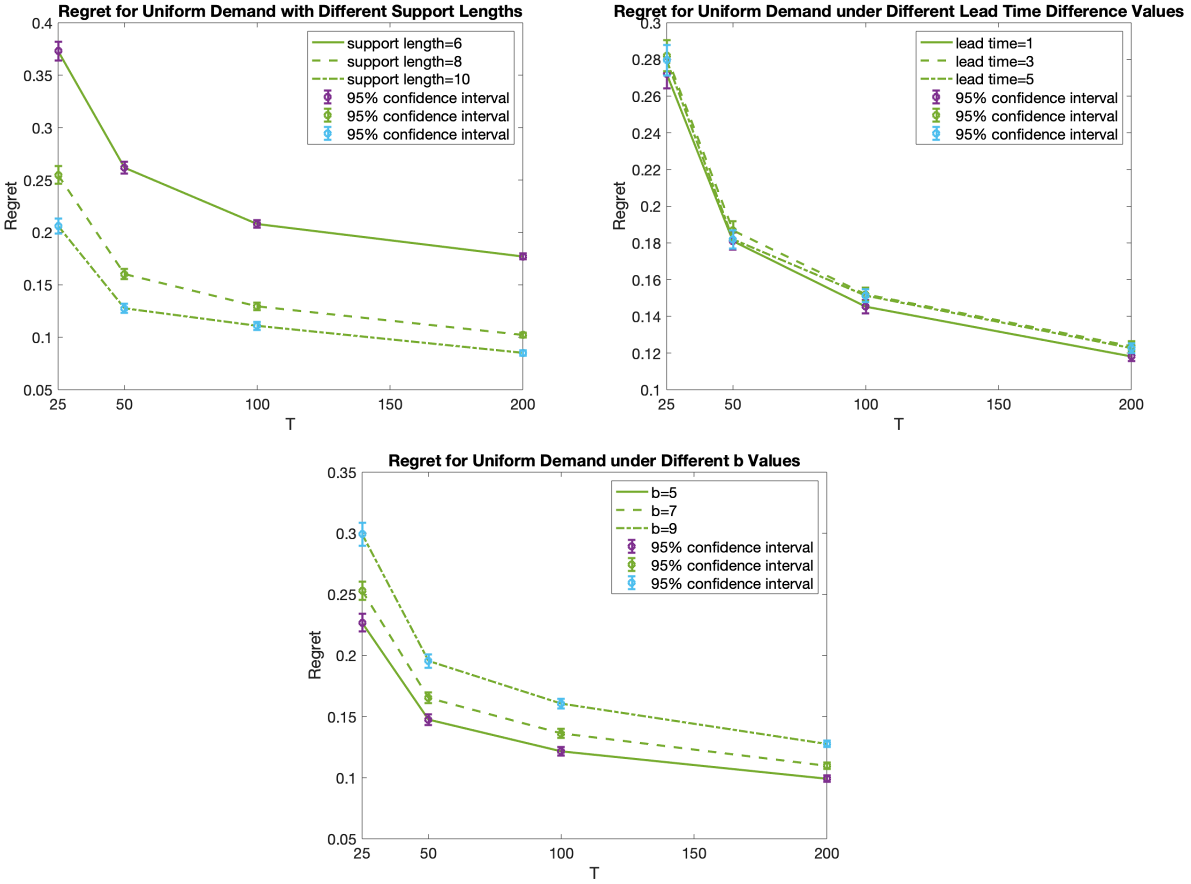
Figure EC.2 Regret for Exponential Demand under Different h and c Values


Regret for Gamma Demand under Different Shape Parameters

Regret for Gamma Demand under Different Lead Time Difference Values

Regret for Gamma Demand under Different b Values

Figure EC.3 **Regret for Gamma Demand**


Regret for Uniform Demand with Different Support Lengths

Regret for Uniform Demand under Different Lead Time Difference Values

Regret for Uniform Demand under Different b Values

Figure EC.4 **Regret for Uniform Demand**

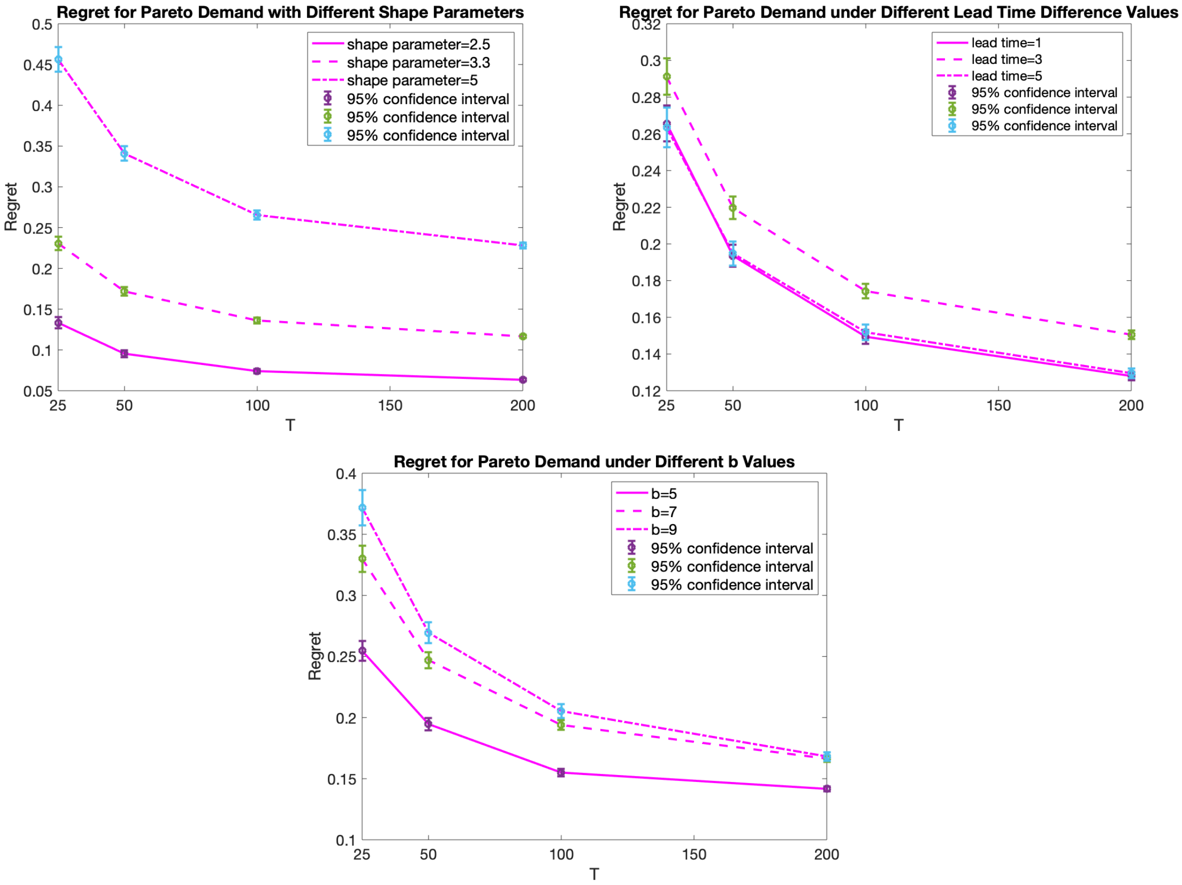


Figure EC.5 Regret for Pareto Demand

Figure EC.6 Average Regret over All Problem Instances