Appendix

Appendix A: Concentration Inequality In this paper, we mainly use Hoeffding's inequality in Hoeffding (1963) [29]. To be self-contained, we present the inequality in the following lemma.

LEMMA 5 (**Hoeffding's Inequality**). *Let $V_1, V_2, \ldots, V_k$ be i.i.d. Bernoulli random variables with mean $\bar{v}$. Then, we have*

$$\mathbb{P}\left(\sum_{\ell=1}^{k} V_\ell - k\bar{v} \geq \sqrt{k\log\xi}\right) \leq \frac{1}{\xi^2}$$
$$\mathbb{P}\left(\sum_{\ell=1}^{k} V_\ell - k\bar{v} \leq -\sqrt{k\log\xi}\right) \leq \frac{1}{\xi^2}$$
$$\mathbb{P}\left(\left|\sum_{\ell=1}^{k} V_\ell - k\bar{v}\right| \geq \sqrt{k\log\xi}\right) \leq \frac{2}{\xi^2}.$$

Appendix B: Discussion of ϵ Consider a simple problem with two types of customers and one type of resource. The arrival probability of either customer type is 0.5. The rewards of the two types of customers are $r_1 = 2$ and $r_2 = 1$, and each customer consumes one unit of resource. Given the time horizon T, the initial inventory is ρT with $\rho \geq 0.5$. Given the resolving schedule $\left\{1, T - T^{1/2+\epsilon}\right\}$, the time horizon is divided into two parts, $[1, T - T^{1/2+\epsilon})$ and $[T - T^{1/2+\epsilon}, T]$. Consider the event when all the following four conditions hold:

$$\Lambda_1\left(1, T - T^{1/2+\epsilon}\right) \leq \frac{1}{2}\left(T - T^{1/2+\epsilon}\right) - 2\sqrt{T} \qquad \Lambda_1\left(T - T^{1/2+\epsilon}, T\right) \leq \frac{1}{2}T^{1/2+\epsilon}$$
$$\Lambda_2\left(1, T - T^{1/2+\epsilon}\right) \geq \frac{1}{2}\left(T - T^{1/2+\epsilon}\right) + 2\sqrt{T} \qquad \Lambda_2\left(T - T^{1/2+\epsilon}, T\right) \leq \frac{1}{2}T^{1/2+\epsilon},$$

where $\Lambda_i(t_1, t_2)$ is the random number of type-i arrivals during the time interval $[t_1, t_2)$. Following the proof of Proposition 5, we can similarly prove that the probability of the above event is greater than a constant.

If we set $\epsilon = 0$, under the above event, the proposed policy will accept all type-1 customers and at most $\left(\rho - \frac{1}{2}\right)T + \frac{1}{2}\sqrt{T}$ type-2 customers. In contrast, the hindsight benchmark will accept all type-1 customers and at least $\left(\rho - \frac{1}{2}\right)T + \frac{3}{2}\sqrt{T}$ type-2 customers, implying an $\Omega(\sqrt{T})$ regret. This is because the error in the first part cannot be fully compensated for by the second part, i.e., $\frac{1}{2}T^{1/2+\epsilon} \leq 2\sqrt{T}$ when $\epsilon = 0$. Therefore, a positive ϵ is necessary for our work even in the non-degenerate case.

Then, a natural question is why the ϵ term does not appear in the finite-resolving bound (under the non-degeneracy assumption) in Jasin and Kumar (2012) [31]. We highlight that the finite-resolving schedule in Jasin and Kumar (2012) [31] is adaptively determined based on the realization of the arrival process. In contrast, our schedule is independent of the arrival process and is determined at the beginning of the time horizon.

Under an adaptive schedule, we resolve the LP when the number of arrivals significantly deviates from the expectation, and hence control the wrong actions more carefully. Thus, an adaptive schedule has the potential to improve the performance. However, the analysis of adaptive schedules is complicated, so the proof in Jasin and Kumar (2012) [31] requires both the non-degeneracy assumption and a finite number of resolvings M independent of T. Moreover, there are no existing results showing that in the degenerate case, solving LPs twice can lead to an $O(T^{\frac{1}{4}})$ regret bound. Therefore, we leave the analysis of adaptive schedules without the degeneracy assumption to future work.

Appendix C: Omitted Proofs In this section, we provide omitted proofs in the main text.

C.1. Proof of Lemma 2 Let μ^* denote the optimal policy of $V^*(T)$. First, for any sample path ω with the demand $\boldsymbol{Z}^1(\omega)$, let $\bar{y}_j(\omega) = \sum_{t=1}^{T} x_{t,j}^{\mu^*}(\omega)$ for each j. Then, it can be verified that $\bar{\boldsymbol{y}}(\omega)$ is always a feasible solution to the problem $\phi(\boldsymbol{b}^1, \boldsymbol{Z}^1(\omega))$ because of the feasibility constraints in (2). Then, we have $\mathbb{E}\left[\mathbb{E}[\phi(\boldsymbol{b}^1, \boldsymbol{Z}^1) : \omega]\right] \geq \mathbb{E}[\boldsymbol{r}^\top \bar{\boldsymbol{y}}(\omega)] = V^*(T)$. □

C.2. Proof of Proposition 1 In this proof, we fix $(\boldsymbol{b}, \boldsymbol{Z}, j)$ with $\boldsymbol{b} \geq \boldsymbol{A}_j x$ and $\boldsymbol{Z} \geq \boldsymbol{e}_j$, and define $\bar{a} := \max_{i,j} a_{ij}$ and $\bar{r} = \max_j r_j$. First, according to Theorem 2.4 in Mangasarian and Shiau (1987) [40], for any optimal solution $\boldsymbol{y}_1^*$ of $\phi(\boldsymbol{b}, \boldsymbol{Z})$, there exists an optimal solution $\boldsymbol{y}_2^*$ of $\phi(\boldsymbol{b} - \boldsymbol{A}_j, \boldsymbol{Z} - \boldsymbol{e}_j)$ such that

$$\|\boldsymbol{y}_1^* - \boldsymbol{y}_2^*\|_\infty \leq \kappa_1 \cdot \max\{\max_i a_{ij}, 1\} \leq \kappa_1 \cdot \max\{\bar{a}, 1\},$$

where κ_1 only depends on the matrix $\boldsymbol{A}$. Then, we have

$$\Delta(\boldsymbol{b}, \boldsymbol{Z}, j, x) = \phi(\boldsymbol{b}, \boldsymbol{Z}) - \phi(\boldsymbol{b} - x\boldsymbol{A}_j, \boldsymbol{Z} - \boldsymbol{e}_j) - x r_j \leq \sum_{j=1}^{n} \kappa_1 r_j \cdot \max\{\bar{a}, 1\} \leq n\kappa_1 \bar{r} \cdot \max\{\bar{a}, 1\}.$$

Therefore, we have $\Delta(\boldsymbol{b}, \boldsymbol{Z}, j, x) \leq r_\phi := n\kappa_1 \bar{r} \cdot \max\{\bar{a}, 1\}$, which is independent of T.

Let $\boldsymbol{y}^*$ be an optimal solution of $\phi(\boldsymbol{b}, \boldsymbol{Z})$ with $y_j^* \geq 1$. Then, we prove that $\boldsymbol{y}^* - \boldsymbol{e}_j$ is an optimal solution to $\phi(\boldsymbol{b} - \boldsymbol{A}_j, \boldsymbol{Z} - \boldsymbol{e}_j)$. First, it is obvious that $\boldsymbol{y}^* - \boldsymbol{e}_j$ is a feasible solution. Second, suppose there exists a feasible solution $\tilde{\boldsymbol{y}}$ of $\phi(\boldsymbol{b} - \boldsymbol{A}_j, \boldsymbol{Z} - \boldsymbol{e}_j)$ such that $\boldsymbol{r}^\top \tilde{\boldsymbol{y}} > \boldsymbol{r}^\top(\boldsymbol{y}^* - \boldsymbol{e}_j)$. Since $\tilde{\boldsymbol{y}} + \boldsymbol{e}_j$ is a feasible solution to $\phi(\boldsymbol{b}, \boldsymbol{Z})$, we have $\boldsymbol{r}^\top(\tilde{\boldsymbol{y}} + \boldsymbol{e}_j) > \boldsymbol{r}^\top \boldsymbol{y}^*$, which contradicts the optimality of $\boldsymbol{y}^*$. Therefore, we can deduce that $\boldsymbol{y}^* - \boldsymbol{e}_j$ is an optimal solution to $\phi(\boldsymbol{b} - \boldsymbol{A}_j, \boldsymbol{Z} - \boldsymbol{e}_j)$, implying that $\Delta(\boldsymbol{b}, \boldsymbol{Z}, j, 1) = 0$. Similarly, we can prove that $\Delta(\boldsymbol{b}, \boldsymbol{Z}, j, 0) = 0$ if $Z_j - y_j^* \geq 1$. □

C.3. Proof of Proposition 2 For $t \in [T_{k-1}, T_k)$, we have

$$
\begin{aligned}
|d_j^t - Z_j^t| &= \left| \frac{(\sum_{\ell=1}^{T_{k-1}-1} Y_j^\ell)\cdot(T - T_{k-1}+1)}{T_{k-1}-1} - \sum_{\ell=T_{k-1}}^{t-1} Y_j^\ell - \sum_{\ell=t}^{T} Y_j^\ell \right| \\
&= \left| \frac{(\sum_{\ell=1}^{T_{k-1}-1} Y_j^\ell)\cdot(T - T_{k-1}+1)}{T_{k-1}-1} - \sum_{\ell=T_{k-1}}^{T} Y_j^\ell \right| \\
&\le (T - T_{k-1}+1)\left(\left| \frac{(\sum_{\ell=1}^{T_{k-1}-1} Y_j^\ell)}{T_{k-1}-1} - p_j \right| + \left| p_j - \frac{\sum_{\ell=T_{k-1}}^{T} Y_j^\ell}{T - T_{k-1}+1} \right| \right),
\end{aligned}
$$

where the first equality follows from the definition of d_j^t. Note that the arrival process is i.i.d. across time. By Hoeffding's inequality, it holds that

$$
\begin{aligned}
&\mathbb{P}\left(|d_j^t - Z_j^t| \ge (T - T_{k-1}+1)\sqrt{\frac{\log(t-1)}{T_{k-1}-1}} + \sqrt{(T-T_{k-1}+1)\log(T-t+1)} \right) \\
\le&\mathbb{P}\left(\left| \frac{(\sum_{\ell=1}^{T_{k-1}-1} Y_j^\ell)}{T_{k-1}-1} - p_j \right| \ge \sqrt{\frac{\log(t-1)}{T_{k-1}-1}} \right) + \mathbb{P}\left(\left| p_j - \frac{\sum_{\ell=T_{k-1}}^{T} Y_j^\ell}{T - T_{k-1}+1} \right| \ge \sqrt{\frac{\log(T-t+1)}{T-T_{k-1}+1}} \right) \\
\le&\frac{2}{(t-1)^2} + \frac{2}{(T-t+1)^2}.
\end{aligned}
$$

Similarly, we have $|d_j^t - Z_j^t| \le (T - T_{k-1}+1)\sqrt{\frac{\log(T-t+1)}{T_{k-1}-1}} + \sqrt{(T-T_{k-1}+1)\log(T-t+1)}$ with probability at least $1 - \frac{4}{(T-t+1)^2}$. □

C.4. Proof of Lemma 1 Given $t \in [T_{k-1}, T_k)$, there exist n_1 and n_2 such that $T_{k-1} \ge \max\{\lceil T^{\alpha^{n_1+1}} \rceil, \lceil T - T^{\beta^{n_2}} \rceil\}$ and $T_k \le \min\{\lceil T^{\alpha^{n_1}} \rceil, \lceil T - T^{\beta^{n_2+1}} \rceil\}$. Then, we have

$$
\begin{aligned}
&(t-1)^\alpha \le (T_k - 1)^\alpha \le (\lceil T^{\alpha^{n_1}} \rceil - 1)^\alpha \le (T^{\alpha^{n_1}})^\alpha = T^{\alpha^{n_1+1}} \le T_{k-1} \\
&(T-t+1)^{1/\beta} \ge (T - T_k + 1)^{1/\beta} \ge (T - \lceil T - T^{\beta^{n_2+1}} \rceil + 1)^{1/\beta} \ge (T^{\beta^{n_2+1}})^{1/\beta} = T^{\beta^{n_2}} \ge T - T_{k-1}.
\end{aligned}
$$
□

C.5. Proof of Proposition 3 We prove the two statements in Proposition 3 one by one. Before proceeding, we first simplify $d_j^t - (T-t+1)p_j$ as follows:

$$
\begin{aligned}
d_j^t - (T-t+1)p_j &= \frac{(\sum_{\ell=1}^{T_{k-1}-1} Y_j^\ell)\cdot(T - T_{k-1}+1)}{T_{k-1}-1} - \sum_{\ell=T_{k-1}}^{t-1} Y_j^\ell - (T-t+1)p_j \\
&= (T - T_{k-1}+1)\left(\frac{\sum_{\ell=1}^{T_{k-1}-1} Y_j^\ell}{T_{k-1}-1} - p_j \right) + (t - T_{k-1})\left(p_j - \frac{\sum_{\ell=T_{k-1}}^{t-1} Y_j^\ell}{t - T_{k-1}} \right).
\end{aligned}
$$

(i) Consider the time period $t \in [3, T-3]$ such that there exist T_{k-1} and T_k satisfying $t \in [T_{k-1}, T_k)$. There are two possible cases:

(a) When $T_{k-1} < \lceil \frac{T}{2} \rceil$, similar to the proof of Proposition 2, by Hoeffding's inequality, we have

$$\begin{aligned}
&\mathbb{P}\left(d_j^t - (T-t+1)p_j \leq -(T-T_{k-1}+1)\sqrt{\frac{\log(t-1)}{T_{k-1}-1}} - \sqrt{(t-T_{k-1})\log(T-t+1)}\right) \\
\leq &\mathbb{P}\left(\frac{(\sum_{\ell=1}^{T_{k-1}-1} Y_j^\ell)}{T_{k-1}-1} - p_j \leq -\sqrt{\frac{\log(t-1)}{T_{k-1}-1}}\right) + \mathbb{P}\left(p_j - \frac{\sum_{\ell=T_{k-1}}^{t-1} Y_j^\ell}{t-T_{k-1}} \leq -\sqrt{\frac{\log(T-t+1)}{t-T_{k-1}}}\right) \\
\leq &\frac{1}{(t-1)^2} + \frac{1}{(T-t+1)^2}.
\end{aligned}$$

Since $\lceil \frac{T}{2} \rceil \in \mathcal{T}$ and $T_{k-1} < \lceil \frac{T}{2} \rceil$, we have $t - T_{k-1} \leq T - T_{k-1} + 1 \leq T \leq 2(T - T_k + 1) \leq 2(T - t + 1)$. Then, according to Lemma 1, with probability at least $1 - \frac{1}{(t-1)^2} - \frac{1}{(T-t+1)^2}$, it holds

$$d_j^t \geq (T-t+1)p_j - 2(T-t+1)\sqrt{\frac{\log(t-1)}{(t-1)^\alpha - 1}} - \sqrt{2(T-t+1)\log(T-t+1)}.$$

Since $\alpha > 0$, there exists a constant η_j such that when $t \geq \eta_j$, we have $2\sqrt{\frac{\log(t-1)}{(t-1)^\alpha-1}} \leq \frac{p_j}{4}$. Moreover, there exists a constant η'_j such that when $t \leq T - \eta'_j$, we have

$$\frac{p_j}{4}(T-t+1) - \sqrt{2(T-t+1)\log(T-t+1)} \geq 0 \text{ and } \frac{p_j(T-t+1)}{2} \geq 2.$$

Therefore, when $t \in [\eta_j, \min\{T - \eta'_j, \lceil \frac{T}{2} \rceil\}]$, we have

$$d_j^t \geq \frac{3p_j(T-t+1)}{4} - \sqrt{2(T-t+1)\log(T-t+1)} \geq \frac{p_j(T-t+1)}{2} \geq 2$$

with probability at least $1 - \frac{1}{(t-1)^2} - \frac{1}{(T-t+1)^2}$.

(b) When $T_{k-1} \geq \lceil \frac{T}{2} \rceil$, according to Hoeffding's inequality, we have

$$d_j^t \geq (T-t+1)p_j - (T-T_{k-1}+1)\sqrt{\frac{\log(T-t+1)}{T_{k-1}-1}} - \sqrt{(t-T_{k-1})\log(T-t+1)},$$

with probability at least $1 - \frac{2}{(T-t+1)^2}$. Since $T_{k-1} \geq \lceil \frac{T}{2} \rceil$, we have $T_{k-1} - 1 \geq \frac{T-T_{k-1}+1}{2}$. According to Lemma 1, we have

$$\begin{aligned}
d_j^t &\geq (T-t+1)p_j - (T-T_{k-1}+1)\sqrt{\frac{2\log(T-t+1)}{T-T_{k-1}+1}} - \sqrt{(T-T_{k-1}+1)\log(T-t+1)} \\
&= (T-t+1)p_j - (\sqrt{2}+1)\sqrt{(T-T_{k-1}+1)\log(T-t+1)} \\
&\geq (T-t+1)p_j - (\sqrt{2}+2)(T-t+1)^{\frac{1}{2\beta}}\sqrt{\log(T-t+1)}.
\end{aligned}$$

Since $\frac{1}{2\beta} < 1$, there exists a constant η''_j such that when $t \leq T - \eta''_j$, we have $d_j^t \geq \frac{p_j(T-t+1)}{2} \geq 2$ with probability $1 - \frac{2}{(T-t+1)^2}$.

Let $c_1 = \max\{\max_j \eta_j,\, 3\}$ and $c_2 = \max\{\max_j \eta'_j, \max_j \eta''_j, 3\}$, which are independent of T. The above proof implies that, when $t \in [c_1, T-c_2]$, we have $\boldsymbol{d}^t \geq \frac{p_j(T-t+1)}{2} \geq \mathbf{2}$ with probability $1 - \frac{n}{(T-t+1)^2} - \frac{n}{(\min\{t-1, T-t+1\})^2}$.

(ii) According to the proof of (i), when $t \in [c_1, T-c_2] \cap [T_{k-1}, T_k)$, we have $\boldsymbol{d}^t \geq \mathbf{2}$ with high probability. Then, it suffices to prove that if $\boldsymbol{d}^t \geq \mathbf{2}$, then $\boldsymbol{u}^t$ is an optimal solution to the surrogate LP $\phi(\boldsymbol{b}^t, \boldsymbol{d}^t)$. Given that $\boldsymbol{d}^t \geq \mathbf{2}$, we have $\boldsymbol{d}^\ell \geq \mathbf{2}$ when $\ell \in [T_{k-1}, t-1]$ due to the monotonicity of $\boldsymbol{d}^t$. When $t = T_{k-1}$, the solution $\boldsymbol{u}^t$ is an optimal solution obtained from $\phi(\boldsymbol{b}^t, \boldsymbol{d}^t)$. Then, we prove the statement by induction.

Suppose $\boldsymbol{u}^\tau$ is an optimal solution to $\phi(\boldsymbol{b}^\tau, \boldsymbol{d}^\tau)$ with $\tau \in [T_{k-1}, t)$. Since $\boldsymbol{d}^\tau \geq \mathbf{2}$, if the AIR algorithm accepts the arriving customer of type j^t, then we have $u^\tau_{j^t} \geq \frac{1}{2} d^\tau_{j^t} \geq 1$. Thus, we have $\boldsymbol{u}^{\tau+1} = \boldsymbol{u}^\tau - \boldsymbol{e}_{j^t} \geq \mathbf{0}$ is a feasible solution of $\phi(\boldsymbol{b}^{\tau+1}, \boldsymbol{d}^{\tau+1})$. If there exists a feasible solution $\tilde{\boldsymbol{u}}$ of $\phi(\boldsymbol{b}^{\tau+1}, \boldsymbol{d}^{\tau+1})$ such that $\boldsymbol{r}^\top \tilde{\boldsymbol{u}} > \boldsymbol{r}^\top \boldsymbol{u}^{\tau+1}$, then we have $\tilde{\boldsymbol{u}} + \boldsymbol{e}_{j^t}$ feasible to $\phi(\boldsymbol{b}^\tau, \boldsymbol{d}^\tau)$ and $\boldsymbol{r}^\top(\tilde{\boldsymbol{u}} + \boldsymbol{e}_{j^t}) > \boldsymbol{r}^\top \boldsymbol{u}^\tau$, which contradicts the optimality of $\boldsymbol{u}^\tau$. Then, we can deduce that $\boldsymbol{u}^{\tau+1}$ is an optimal solution to $\phi(\boldsymbol{b}^{\tau+1}, \boldsymbol{d}^{\tau+1})$. If the AIR algorithm rejects the arriving customer, we can similarly prove the optimality of $\boldsymbol{u}^{\tau+1}$. Therefore, we can prove that $\boldsymbol{u}^t$ is optimal to $\phi(\boldsymbol{b}^t, \boldsymbol{d}^t)$. □

C.6. Proof of Proposition 4 In the following, we always consider the good event. First, we bound the difference $|\phi(\boldsymbol{b}^t, \boldsymbol{Z}^t) - \phi(\boldsymbol{b}^t, \boldsymbol{d}^t)|$. According to Theorem 2.4 in Mangasarian and Shiau (1987) [40], for any optimal solution $\boldsymbol{y}_1$ to $\phi(\boldsymbol{b}^t, \boldsymbol{Z}^t)$, there exists an optimal solution $\boldsymbol{y}_2$ to $\phi(\boldsymbol{b}^t, \boldsymbol{d}^t)$ such that $\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_\infty \leq \kappa_2 \|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty$, where κ_2 is a constant independent of T. Therefore, we have

$$|\phi(\boldsymbol{b}^t, \boldsymbol{Z}^t) - \phi(\boldsymbol{b}^t, \boldsymbol{d}^t)| = |\boldsymbol{r}^\top(\boldsymbol{y}_1 - \boldsymbol{y}_2)| \leq \left(\sum_{j=1}^n r_j\right) \|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_\infty \leq \kappa_2 \left(\sum_j r_j\right) \|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty.$$

Note that $\bar{\mathfrak{S}}(\boldsymbol{b}, \boldsymbol{d}, j)$ can be formulated as an LP:

$$\begin{aligned} \max_{\boldsymbol{y} \geq \mathbf{0}} \quad & y_j \\ \text{s.t.} \quad & \boldsymbol{r}^\top \boldsymbol{y} \geq \phi(\boldsymbol{b}, \boldsymbol{d}), \\ & \boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{b}, \\ & \boldsymbol{y} \leq \boldsymbol{d}. \end{aligned}$$

According to Theorem 2.4 in Mangasarian and Shiau (1987) [40], we have

$$\|\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) - \bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{d}^t, j)\|_\infty \leq \kappa_3(\|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty + |\phi(\boldsymbol{b}^t, \boldsymbol{Z}^t) - \phi(\boldsymbol{b}^t, \boldsymbol{d}^t)|) \leq \kappa_4 \|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty,$$

where $\kappa_4 = \kappa_3 + \kappa_2 \sum_j r_j$.

Consider the case when a type-j customer arrives at time t and the AIR policy accepts this request. We then have

$$\begin{aligned}\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) &\geq \bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{d}^t, j) - \kappa_4 \|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty \geq u_j^t - \kappa_4 \|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty \\ &\geq \frac{d_j^t}{2} - \kappa_4 \|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty \geq \frac{p_j(T-t+1)}{4} - \kappa_4 \|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty,\end{aligned}$$

where the second inequality holds because $\boldsymbol{u}^t$ is an optimal solution to $\phi(\boldsymbol{b}^t, \boldsymbol{d}^t)$, the third inequality is due to the argmax operation in Algorithm 1, and the last inequality follows from Proposition 3.

Given $t \in [T_{k-1}, T_k)$, we discuss two possible cases:

(a) When $T_{k-1} < \lceil \frac{T}{2} \rceil$, since $\lceil \frac{T}{2} \rceil \in \mathcal{T}$ and $T_{k-1} < \lceil \frac{T}{2} \rceil$, we have $t \leq \lceil \frac{T}{2} \rceil$ and $T - T_{k-1} + 1 \leq T \leq 2(T - t + 1)$. According to Proposition 2, we have

$$\begin{aligned}&\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \\ \geq &\frac{p_j(T-t+1)}{4} - \kappa_4 (T - T_{k-1} + 1)\sqrt{\frac{\log(t-1)}{T_{k-1} - 1}} - \kappa_4 \sqrt{(T - T_{k-1} + 1)\log(T - t + 1)} \\ \geq &\frac{p_j(T-t+1)}{4} - 2\kappa_4 (T - t + 1)\sqrt{\frac{\log(t-1)}{(t-1)^\alpha - 1}} - \kappa_4 \sqrt{2(T - t + 1)\log(T - t + 1)}.\end{aligned}$$

where the last inequality follows from Lemma 1. Since $\alpha > 0$, there exists a constant θ_j such that when $t \geq \theta_j$, we have $2\kappa_4 \sqrt{\frac{\log(t-1)}{(t-1)^\alpha - 1}} \leq \frac{p_j}{8}$. Moreover, there exists a constant θ_j' such that when $t \leq T - \theta_j'$, we have $\frac{(T-t+1)p_j}{8} - \kappa\sqrt{2(T-t+1)\log(T-t+1)} \geq 1$. Therefore, when $t \in [\theta_j, \min\{\lceil \frac{T}{2} \rceil, T - \theta_j'\}]$, we have $\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq \frac{(T-t+1)p_j}{8} - \kappa_4 \sqrt{2(T-t+1)\log(T-t+1)} \geq 1$.

(b) When $T_{k-1} \geq \lceil \frac{T}{2} \rceil$, we have $T_{k-1} - 1 \geq \frac{T - T_{k-1} + 1}{2}$. According to Proposition 2, we have

$$\begin{aligned}&\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \\ \geq &\frac{p_j(T-t+1)}{4} - \kappa_4 (T - T_{k-1} + 1)\sqrt{\frac{\log(T-t+1)}{T_{k-1} - 1}} - \kappa_4 \sqrt{(T - T_{k-1} + 1)\log(T - t + 1)} \\ \geq &\frac{p_j(T-t+1)}{4} - \kappa_4 (T - T_{k-1} + 1)\sqrt{\frac{2\log(T-t+1)}{T - T_{k-1} + 1}} - \kappa_4 \sqrt{(T - T_{k-1} + 1)\log(T - t + 1)} \\ = &\frac{p_j(T-t+1)}{4} - \kappa_4 (\sqrt{2} + 1)\sqrt{(T - T_{k-1} + 1)\log(T - t + 1)} \\ \geq &\frac{p_j(T-t+1)}{4} - \kappa_4 (2 + \sqrt{2})(T - t + 1)^{\frac{1}{2\beta}} \sqrt{\log(T - t + 1)}.\end{aligned}$$

where the last inequality follows from Lemma 1. Similar to the above discussion, since $\beta > \frac{1}{2}$, there exists a constant θ_j'' such that $\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq 1$ when $\lceil \frac{T}{2} \rceil \leq t \leq T - \theta_j''$.

Let $\kappa_5^A = \max_j \theta_j$ and $\kappa_6^A = \max\{\max_j \theta'_j, \max_j \theta''_j\}$, which are independent of T. When $t \in [\kappa_5^A, T - \kappa_6^A] \cap [c_1, T - c_2]$, if the algorithm accepts a type-j customer at time t, then we have $\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq 1$. Similarly, for the rejection action, we can derive κ_5^R and κ_6^R. By setting $c_5 = \max\{\kappa_5^A, \kappa_5^R, c_1\}$ and $c_6 = \max\{\kappa_6^A, \kappa_6^R, c_2\}$, we have $\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j^t) \geq 1$ if $x_{j^t}^t = 1$ and $Z_j^t - \bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j^t) \geq 1$ if $x_{j^t}^t = 0$. Finally, according to Proposition 1, we have $\Delta(\boldsymbol{b}^t, \boldsymbol{Z}^t, j^t, x_{j^t}^t) = 0$. $\square$

C.7. Proof of Theorem 2 We discuss different time intervals.

1. For the time interval $\left[1, \left\lceil T^{\beta^{M-1}} \right\rceil\right)$, Algorithm 1 rejects all customers and incurs at most $T^{\beta^{M-1}} \cdot r_\phi$ regret.
2. For the time interval $\left[\left\lceil T^{\beta^{M-1}} \right\rceil, \left\lceil T - T^{\beta^{M-1}} \right\rceil\right]$, the proofs of Propositions 2, 3 and 4 still hold, and hence the regret at time t can be bounded by $\left(\frac{c_3}{(T-t+1)^2} + \frac{c_4}{(t-1)^2}\right) r_\phi$.
3. For the time interval $\left[\left\lceil T - T^{\beta^{M-1}} \right\rceil + 1, T\right]$, the regret is no greater than $T^{\beta^{M-1}} r_\phi$.

To summarize, given the resolving time set $\mathcal{T}^F(M)$, the regret can be bounded as

$$\begin{aligned}
\mathrm{Reg}^{\mathcal{A}}(T) &\leq T^{\beta^{M-1}} \cdot r_\phi + \sum_{t=\left\lceil T^{\beta^{M-1}} \right\rceil}^{\left\lceil T - T^{\beta^{M-1}} \right\rceil} \left(\frac{c_3}{(T-t+1)^2} + \frac{c_4}{(t-1)^2}\right) r_\phi + T^{\beta^{M-1}} r_\phi \\
&\leq 2T^{\beta^{M-1}} r_\phi + \frac{\pi^2}{6}(c_3 + c_4) r_\phi = O(T^{\beta^{M-1}}).
\end{aligned}$$ $\square$

C.8. Proof of Theorem 3 Similar to Lemma 1, we have $T - T_{k-1} \leq (T - t + 1)^{1/\beta}$ for any $t \in [T_{k-1}, T_k)$. In this case, the regret decomposition is the same as (6) but the proof is simpler than the unknown-probability case. In the following, we present the main steps of the proof. First, we bound the approximation error, i.e., the difference between $\boldsymbol{d}^t$ and $\boldsymbol{Z}^t$.

LEMMA 6. *Consider the resolving schedule $\mathcal{T}^{\mathcal{K}}$. Given a time $t \in [T_{k-1}, T_k)$, we have $|d_j^t - Z_j^t| \leq \sqrt{(T - T_{k-1} + 1)\log(T - t + 1)}$ with probability larger than $1 - \frac{2}{(T-t+1)^2}$.*

Second, we prove the relationship between $\boldsymbol{u}^t$ and a surrogate LP.

LEMMA 7. *Consider the known-probability case. For the AIR-KP policy with the resolving schedule $\mathcal{T}^{\mathcal{K}}$ with $\beta \in (1/2, 1)$, there exists a constant c_7 independent of T such that when $t \leq T - c_7$, with probability larger than $1 - \frac{n}{(T-t+1)^2}$, we have*

1. *$d_j^t \geq \frac{p_j(T-t+1)}{2} \geq 2$ for any j.*
2. *$\boldsymbol{u}^t$ is an optimal solution to the surrogate LP $\phi(\boldsymbol{b}^t, \boldsymbol{d}^t)$.*

Third, we bound the difference between $\boldsymbol{u}^t$ and the hindsight solution, and then prove the regret bound. The *good event* is the event on which the conditions in both Lemmas 6 and 7 are satisfied, and the probability is no less than $1-\frac{c_8}{(T-t+1)^2}$.

LEMMA 8. *Consider the known-probability case. For the AIR-KP policy with the resolving schedule $\mathcal{T}^{\mathcal{K}}$ with $\beta \in (\frac{1}{2},1)$, there exists a constant c_{10} independent of T such that when $t \le T-c_{10}$, under the good event, we have*

1. $\bar{\mathfrak{S}}(\boldsymbol{b}^t,\boldsymbol{Z}^t,j)\ge 1$ *if* $x^{\mathcal{A}}_{t,j^t}=1$ *and* $Z^t_j-\underline{\mathfrak{S}}(\boldsymbol{b}^t,\boldsymbol{Z}^t,j)\ge 1$ *if* $x^{\mathcal{A}}_{t,j^t}=0$.
2. $\Delta(\boldsymbol{b}^t,\boldsymbol{Z}^t,j^t,x^{\mathcal{A}}_{t,j^t})=0$.

Lastly, we can prove the constant regret bound.

$$\begin{aligned}\mathrm{Reg}^{\mathcal{A}}(T) &\le \sum_{t=1}^{T-c_{10}} r_\phi \mathbb{P}\left(\Delta(\boldsymbol{b}^t,\boldsymbol{Z}^t,j^t,x^{\mathcal{A}}_{t,j^t})>0\right)+c_{10}r_\phi \\ &\le \sum_{t=1}^{T-c_{10}} r_\phi \frac{c_8}{(T-t+1)^2}+c_{10}r_\phi \le \left(\frac{\pi^2}{6}c_8+c_{10}\right) r_\phi. \qquad \square\end{aligned}$$

C.9. Proof of Lemma 6 In this case, for $t\in[T_{k-1},T_k)$, by Hoeffding's inequality, it holds that

$$\begin{aligned}&\mathbb{P}\left(|d^t_j-Z^t_j|\ge\sqrt{(T-T_{k-1}+1)\log(T-t+1)}\right)\\ =&\mathbb{P}\left(\left|p_j(T-T_{k-1}+1)-\sum_{\ell=T_{k-1}}^{T} Y^\ell_j\right|\ge\sqrt{(T-T_{k-1}+1)\log(T-t+1)}\right)\\ \le&\frac{2}{(T-t+1)^2}. \qquad \square\end{aligned}$$

C.10. Proof of Lemma 7 First, for $t\in[T_{k-1},T_k)$, by Hoeffding's inequality, we have

$$\begin{aligned}&\mathbb{P}\left(d^t_j\le(T-t+1)p_j-\sqrt{(t-T_{k-1})\log(T-t+1)}\right)\\ =&\mathbb{P}\left((t-T_{k-1})p_j-\sum_{\ell=T_{k-1}}^{t-1} Y^\ell_j\le-\sqrt{(t-T_{k-1})\log(T-t+1)}\right)\\ \le&\frac{1}{(T-t+1)^2}.\end{aligned}$$

Therefore, with probability greater than $1-\frac{1}{(T-t+1)^2}$, we have $d^t_j\ge(T-t+1)p_j-\sqrt{(t-T_{k-1})\log(T-t+1)}\ge(T-t+1)p_j-\sqrt{(T-t+1)^{1/\beta}\log(T-t+1)}$. Since $\frac{1}{2\beta}<1$, there exists a constant c_7 such that when $t\le T-c_7$, $\boldsymbol{d}^t\ge\frac{T-t+1}{2}\boldsymbol{p}\ge\boldsymbol{2}$. Similar to Proposition 3, we can prove that $\boldsymbol{u}^t$ is an optimal solution to $\phi(\boldsymbol{b}^t,\boldsymbol{d}^t)$ when $\boldsymbol{d}^t\ge\boldsymbol{2}$. $\square$

C.11. Proof of Lemma 8 In the following, we always consider the good event. Consider the case when a type-j customer arrives at time t and the AIR-KP policy accepts the request. Similar to Proposition 4, given $t \in [T_{k-1}, T_k)$ and $t \le T - c_7$, we have

$$\begin{aligned}\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) &\ge \frac{p_j(T-t+1)}{4} - \kappa_7 \|\boldsymbol{d}^t - \boldsymbol{Z}^t\|_\infty \ge \frac{p_j(T-t+1)}{4} - \kappa_7 \sqrt{(T-T_{k-1}+1)\log(T-t+1)} \\ &\ge \frac{p_j(T-t+1)}{4} - \kappa_7 \sqrt{2(T-t+1)^{1/\beta}\log(T-t+1)}.\end{aligned}$$

Since $\frac{1}{2\beta} < 1$, there exists a constant c_9 such that when $t \le T - c_9$, we have $\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j^t) \ge 1$ if $x^{\mathcal{A}}_{t,j^t} = 1$ and $Z^t_{j^t} - \underline{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j^t) \ge 1$ otherwise. Then, we set $c_{10} = \max\{c_7, c_9, 3\}$, and finish the proof. □

C.12. Proof of Theorem 4 We discuss different time intervals.

1. For the time interval $\left[1, \left\lceil T - T^{\beta^M} \right\rceil\right)$, the proofs of Lemmas 6, 7 and 8 still hold, and hence the regret at time t can be bounded by $\frac{c_8}{(T-t+1)^2} r_\phi$.
2. For the time interval $\left[\left\lceil T - T^{\beta^M} \right\rceil, T\right]$, the regret is no greater than $T^{\beta^M} r_\phi$.

To summarize, given the resolving time set $\mathcal{T}^{\mathcal{K}}$, the regret can be bounded as

$$\mathrm{Reg}^{\mathcal{A}}(T) \le \sum_{t=1}^{\left\lceil T - T^{\beta^M} \right\rceil} \frac{c_8}{(T-t+1)^2} r_\phi + T^{\beta^M} \cdot r_\phi \le \frac{c_8 \pi^2}{6} r_\phi + T^{\beta^M} r_\phi = O(T^{\beta^M}). \qquad \square$$

C.13. Proof of Proposition 5 According to Theorem 3, we have $\mathbb{E}\left[\phi(T\boldsymbol{\rho}, \boldsymbol{Z}^1)\right] - V^*(T) = O(1)$, i.e., the difference between the hindsight problem and the optimal problem is upper bounded by a constant. Thus, in the following, we compare the policy with the hindsight problem.

Let T_1 and T_2 (with $T_1 < T_2$) denote the two resolving time points. Similar to Bumpensanti and Wang (2020) [15], in this proof, we consider two types of customers and one type of resource. The arrival probabilities of either customer type is 0.5. The rewards of the two types of customers are $r_1 = 2$ and $r_2 = 1$, and each customer consumes one unit of resource. The budget per period is 0.5, resulting in a degenerate case.

We start with the case where $T_1 = \omega(T^{1/4})$, which means that $\lim_{T\to\infty} \frac{T_1}{T^{1/4}} = \infty$. In this case, due to the initialization in Algorithm 2, the algorithm rejects all customers during the time interval $[1, T_1)$. Let $\Lambda_j(t_1, t_2)$ denote the random number of type-j arrivals during the time interval $[t_1, t_2]$. We consider the event where $\frac{1}{2}(T_1 - 1) - \sqrt{T_1 - 1} \le \Lambda_1(1, T_1 - 1) \le \frac{1}{2}(T_1 - 1)$ and $\Lambda_1(T_1, T) \le \frac{1}{2}(T - T_1 + 1)$. In this case, we have $\Lambda_1(1, T) = \Lambda_1(1, T_1 - 1) + \Lambda_1(T_1, T) \le \frac{1}{2}T$ and hence the hindsight problem will accept all type-1 customers. However, the algorithm rejects $\Lambda_1(1, T_1 - 1)$ type-1 customers, and hence the corresponding revenue loss is at least $\Lambda_1(1,T_1-1) \geq \frac{1}{2}(T_1-1) - \sqrt{T_1-1} = \Omega(T^{1/4})$. Then, similar to Bumpensanti and Wang (2020) [15], according to the Berry-Esseen theorem (see Shevtsova 2011 [49]), we bound the probability of the event as follows:

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{2}(T_1-1) - \sqrt{T_1-1} \leq \Lambda_1(1,T_1-1) \leq \frac{1}{2}(T_1-1) \,\&\, \Lambda_1(T_1,T) \leq \frac{1}{2}(T-T_1+1)\right) \\
=&\mathbb{P}\left(-\frac{\sqrt{T_1-1}}{\frac{1}{2}\sqrt{T_1-1}} \leq \frac{\Lambda_1(1,T_1-1)-\frac{1}{2}(T_1-1)}{\frac{1}{2}\sqrt{T_1-1}} \leq 0\right) \cdot \mathbb{P}\left(\frac{\Lambda_1(T_1,T)-\frac{1}{2}(T-T_1+1)}{\frac{1}{2}\sqrt{T-T_1+1}} \leq 0\right) \\
\geq &\left(0.477 - 2 \times \frac{0.4748}{\frac{1}{8}\sqrt{T_1-1}}\right) \cdot \left(0.5 - \frac{0.4748}{\frac{1}{8}\sqrt{T-T_1+1}}\right),
\end{aligned}$$

which is greater than a constant when T is no less than a threshold. Therefore, the regret will of order $\Omega(T^{1/4})$. Then, we consider the case where $T_1 = O(T^{1/4})$, and study different choices of T_2.

Case I: There exists a constant $c_{12} < 1$ **such that** $T_1 \leq T_2 \leq c_{12} \cdot T$**.**

In this case, we consider the event when $\frac{1}{2}(T_2-1) \leq \Lambda_1(1,T_2-1) \leq \frac{1}{2}(T_2-1) + \sqrt{T_2-1}$ and $\Lambda_1(T_2,T) \leq \frac{1}{2}(T-T_2+1) - \sqrt{T}$. Under this event, the total number of type-1 arrivals is $\Lambda_1(1,T_2-1) + \Lambda_1(T_2,T) \leq \frac{1}{2}T + \sqrt{T_2-1} - \sqrt{T} \leq \frac{1}{2}T - (1-\sqrt{c_{12}})\sqrt{T}$, and hence the hindsight optimal policy accepts at least $(1-\sqrt{c_{12}})\sqrt{T}$ type-2 customers. Then, we analyze the performance of the AIR-KP policy with the resolving schedule $\{T_1, T_2\}$. During the time interval $[1, T_1)$, the policy rejects all customers. Then, the algorithm resolves the fluid model and derives the solution $\boldsymbol{u}^{T_1} = \left(\frac{1}{2}(T-T_1), \frac{1}{2}T_1\right)$ and the demand estimation $\boldsymbol{d}^{T_1} = \left(\frac{1}{2}(T-T_1), \frac{1}{2}(T-T_1)\right)$. Thus, during the interval $[T_1, T_2)$, the algorithm accepts all type-1 customers and no greater than $\frac{1}{2}T_1$ type-2 customers, resulting in $v \geq \Lambda_1(1,T_2-1) - T_1$ resource consumption. When the algorithm resolves the fluid model at time T_2, we can derive the solution $\boldsymbol{u}^{T_2}$ with $u_2^{T_2} = \max\left\{\frac{1}{2}T - v - \frac{1}{2}(T-T_2+1), 0\right\} \leq \max\left\{\frac{1}{2}(T_2-1) - \frac{1}{2}(T_2-1) - T_1, 0\right\} \leq T_1$. Therefore, the algorithm accept at most $\frac{3}{2}T_1$ type-2 customers and induce at least $(1-\sqrt{c_{12}})\sqrt{T} - \frac{3}{2}T_1 = \Omega(\sqrt{T})$ revenue loss under this event. Then, we provide a lower bound for the event probability as follows (Let $\Phi(\cdot)$ denote the c.d.f of a standard normal distribution):

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{2}(T_2-1) \leq \Lambda_1(1,T_2-1) \leq \frac{1}{2}(T_2-1) + \sqrt{T_2-1} \,\&\, \Lambda_1(T_2,T) \leq \frac{1}{2}(T-T_2+1) - \sqrt{T}\right) \\
=&\mathbb{P}\left(0 \leq \frac{\Lambda_1(1,T_2-1)-\frac{1}{2}(T_2-1)}{\frac{1}{2}\sqrt{T_2-1}} \leq 2\right) \cdot \mathbb{P}\left(\frac{\Lambda_1(T_2,T)-\frac{1}{2}(T-T_2+1)}{\frac{1}{2}\sqrt{T-T_2+1}} \leq -\frac{\sqrt{T}}{\frac{1}{2}\sqrt{T-T_2+1}}\right) \\
\geq&\mathbb{P}\left(0 \leq \frac{\Lambda_1(1,T_2-1)-\frac{1}{2}(T_2-1)}{\frac{1}{2}\sqrt{T_2-1}} \leq 2\right) \cdot \mathbb{P}\left(\frac{\Lambda_1(T_2,T)-\frac{1}{2}(T-T_2+1)}{\frac{1}{2}\sqrt{T-T_2+1}} \leq -\frac{2}{\sqrt{1-c_{12}}}\right)
\end{aligned}$$

$$\geq \left(0.477 - 2\times\frac{0.4748}{\frac{1}{8}\sqrt{T_2-1}}\right)\cdot\left(\Phi\left(-\frac{2}{\sqrt{1-c_{12}}}\right)-\frac{0.4748}{\frac{1}{8}\sqrt{T-T_2+1}}\right),$$

which is greater than a constant when T exceeds a threshold. Therefore, the regret of the algorithm is at least $\Omega(\sqrt{T})$.

Case II: $T-T_2+1=\Omega(\sqrt{T})$ **and** $T-T_2+1=o(T)$**.**

In this case, we consider the event when $\Lambda_1(1,T_2-1)\leq\frac{1}{2}(T_2-1)-\sqrt{T-T_2+1}$ and $\frac{1}{2}(T-T_2+1)+\frac{1}{2}\sqrt{T-T_2+1}\leq\Lambda_1(T_2,T)\leq\frac{1}{2}(T-T_2+1)+\sqrt{T-T_2+1}$. Under this event, we have $\Lambda_1(1,T)=\Lambda_1(1,T_2-1)+\Lambda_1(T_2,T)\leq\frac{1}{2}T$, and hence the hindsight optimal policy accepts all type-1 customers. Then, we analyze the performance of the AIR-KP policy with the resolving schedule $\{T_1,T_2\}$. Similar to Case I, the algorithm rejects all customers during the interval $[1,T_1)$. During the interval $[T_1,T_2)$, the algorithm accepts all type-1 customers and at most $\frac{1}{2}T_1$ type-2 customers. Then, the algorithm resolves the fluid model and obtains the solution $\boldsymbol{u}^{T_2}$ with $u_1^{T_2}=\min\left\{\frac{1}{2}T-\nu,\frac{1}{2}(T-T_2+1)\right\}\leq\frac{1}{2}(T-T_2+1)$. However, since $\Lambda_1(T_2,T)\geq\frac{1}{2}(T-T_2+1)+\frac{1}{2}\sqrt{T-T_2+1}$, the algorithm rejects at least $\frac{1}{2}\sqrt{T-T_2+1}$ type-1 customers, resulting in $\Omega(T^{1/4})$ revenue loss. Then, given a sufficiently large T, we provide a lower bound for the event probability as follows:

$$\begin{aligned}
&\mathbb{P}\Bigg(\Lambda_1(1,T_2-1)\leq\frac{1}{2}(T_2-1)-\sqrt{T-T_2+1}\\
&\qquad\&\ \frac{1}{2}(T-T_2+1)+\frac{1}{2}\sqrt{T-T_2+1}\leq\Lambda_1(T_2,T)\leq\frac{1}{2}(T-T_2+1)+\sqrt{T-T_2+1}\Bigg)\\
=&\mathbb{P}\left(\frac{\Lambda_1(1,T_2-1)-\frac{1}{2}(T_2-1)}{\frac{1}{2}\sqrt{T_2-1}}\leq-\frac{\sqrt{T-T_2+1}}{\frac{1}{2}\sqrt{T_2-1}}\right)\cdot\mathbb{P}\left(1\leq\frac{\Lambda_1(T_2,T)-\frac{1}{2}(T-T_2+1)}{\frac{1}{2}\sqrt{T-T_2+1}}\leq 2\right)\\
\geq&\mathbb{P}\left(\frac{\Lambda_1(1,T_2-1)-\frac{1}{2}(T_2-1)}{\frac{1}{2}\sqrt{T_2-1}}\leq-2\right)\cdot\mathbb{P}\left(1\leq\frac{\Lambda_1(T_2,T)-\frac{1}{2}(T-T_2+1)}{\frac{1}{2}\sqrt{T-T_2+1}}\leq 2\right)\\
\geq&\left(0.022-\frac{0.4748}{\frac{1}{8}\sqrt{T_2-1}}\right)\cdot\left(0.135-2\times\frac{0.4748}{\frac{1}{8}\sqrt{T-T_2+1}}\right),
\end{aligned}$$

which is greater than a constant when T exceeds a threshold. Therefore, the regret of the algorithm is at least $\Omega(T^{1/4})$.

Case III: $T-T_2+1=o(\sqrt{T})$**.**

In this case, we consider the event when $\Lambda_1(1,T_2-1)\leq\frac{1}{2}(T_2-1)-\sqrt{T_2-1}$. Under this event, given a sufficiently large T, we have $\Lambda_1(1,T)=\Lambda_1(1,T_2-1)+\Lambda_1(T_2,T)\leq\frac{1}{2}(T_2-1)-\sqrt{T_2-1}+(T-T_2+1)\leq\frac{1}{2}(T_2-1)-\frac{1}{2}\sqrt{T_2-1}$, and hence the hindsight optimal policy accepts at least $\frac{1}{2}\sqrt{T_2-1}$

type-2 customers. Then, similar to the analysis in Case I, the algorithm rejects all customers during the interval $[1, T_1)$. During the interval $[T_1, T_2)$, the algorithm accepts all type-1 customers and at most $\frac{1}{2}T_1$ type-2 customers. During the interval $[T_2, T]$, the algorithm accepts at most $T - T_2 + 1$ type-2 customers. Thus, the algorithm accepts at most $\frac{1}{2}T_1 + T - T_2 + 1$ type-2 customers, resulting in $\frac{1}{2}\sqrt{T_2 - 1} - \frac{1}{2}T_1 - (T - T_2 + 1) = \Omega(\sqrt{T}) - O(T^{1/4}) - o(T^{1/4}) = \Omega(\sqrt{T})$. Then, we provide a lower bound for the event probability as follows:

$$\mathbb{P}\left(\Lambda_1(1, T_2 - 1) \le \frac{1}{2}(T_2 - 1) - \sqrt{T_2 - 1}\right) = \mathbb{P}\left(\frac{\Lambda_1(1, T_2 - 1) - \frac{1}{2}(T_2 - 1)}{\frac{1}{2}\sqrt{T_2 - 1}} \le -2\right) \ge \left(0.022 - \frac{0.4748}{\frac{1}{8}\sqrt{T_2 - 1}}\right),$$

which is greater than a constant when T exceeds a threshold. Therefore, the regret is $\Omega(\sqrt{T})$.

To summarize, given two resolving time points, the regret of the AIR-KP policy is $\Omega(T^{1/4})$. □

C.14. Proof of Lemma 3 For the known-probability problem, following the proof of Theorem 3, we need to prove results analogous to Lemmas 7 and 8. Similar to Lemma 7, we have

$$d_j^t \ge (T - t + 1)p_j - \sqrt{(t - T_{k-1})\log(T - t + 1)} \ge (T - t + 1)p_j - \sqrt{\omega \log(T - t + 1)},$$

with probability $1 - \frac{1}{(T-t+1)^2}$. There exists a constant $\tilde{\zeta}^P$ such that when $t \le T - \tilde{\zeta}^P$, $\boldsymbol{d}^t \ge \frac{(T-t+1)}{2}\boldsymbol{p} \ge \boldsymbol{2}$. Moreover, $\tilde{\zeta}^P$ is the minimal positive integer x satisfying $\frac{(x+1)\min_j p_j}{2} - \sqrt{\omega \log(x+1)} \ge 0$. Therefore, we can deduce that $\tilde{\zeta}^P = O(\sqrt{\omega}\log\omega) = \tilde{O}(\sqrt{\omega})$.

Similar to Lemma 8, under the good event, we have

$$\begin{aligned}\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) &\ge \frac{p_j(T - t + 1)}{4} - \kappa_7\sqrt{(T - T_{k-1} + 1)\log(T - t + 1)} \\ &\ge \frac{p_j(T - t + 1)}{4} - \kappa_7\sqrt{(T - t + 1 + \omega)\log(T - t + 1)}.\end{aligned}$$

Then, there exists a constant $\tilde{\theta}^P$ such that when $t \le T - \tilde{\theta}^P$, we have $\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j^t) \ge 1$ if $x^{\mathcal{A}}_{t,j^t} = 1$ and $Z_j^t - \underline{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j^t) \ge 1$ if $x^{\mathcal{A}}_{t,j^t} = 0$. Moreover, $\tilde{\theta}^P$ is the minimal integer x satisfying $\frac{(x+1)\min_j p_j}{4} - \sqrt{(x + 1 + \omega)\log(x+1)} \ge 1$, implying that $\tilde{\theta}^P = O(\sqrt{\omega}\log\omega) = \tilde{O}(\sqrt{\omega})$. Then, the regret can be bounded as $\mathrm{Reg}^{\mathcal{A}}(T) \le \left(\frac{\pi^2}{6}c_8 + \tilde{\theta}^P\right) r_\phi = \tilde{O}(\sqrt{\omega})$. □

C.15. Proof of Proposition 6 For the unknown-probability case, following the proof of Theorem 1, we only need to prove results analogous to Propositions 3 and 4.

Lemma 9. *For the resolving schedule $\mathcal{T}^{\mathcal{K},P}(\omega)$, there exist two constants $\eta^P = O(\omega)$ and $\zeta^P = \mathcal{P}(\omega\log\omega)$ independent of T. When $t \in [\eta^P, T - \zeta^P]$, with probability $1 - \frac{c_{10}}{(T-t+1)^2} - \frac{c_{11}}{(\min\{T-t+1, t-1\})^2}$, we have*

1. *$d_j^t \geq \frac{p_j(T-t+1)}{2} \geq 2$ for any j.*
2. *$\boldsymbol{u}^t$ is an optimal solution to the surrogate LP $\phi(\boldsymbol{b}^t, \boldsymbol{d}^t)$.*

LEMMA 10. *For the resolving schedule $\mathcal{T}^{\mathcal{K},P}(\omega)$, there exist two constants $\iota^P = O(\omega)$ and $\theta^P = O(\omega\log\omega)$ independent of T such that when $t \in [\iota^P, T-\theta^P]$, under the good event, we have*

1. *$\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq 1$ if $x_{t,j^t}^{\mathcal{A}} = 1$ and $Z_j^t - \underline{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq 1$ if $x_{t,j^t}^{\mathcal{A}} = 0$.*
2. *$\Delta(\boldsymbol{b}^t, \boldsymbol{Z}^t, j^t, x_{t,j^t}^{\mathcal{A}}) = 0$.*

Then, the regret bound can be bounded as $\mathrm{Reg}^{\mathcal{A}}(T) \leq \left(\frac{\pi^2}{6}(c_{10}+c_{11}) + \iota^P + \theta^P\right) r_\phi = O(\omega)$.

Proof of Lemma 9. To derive a result analogous to Proposition 3, we only need to show that $\boldsymbol{d}^t \geq \frac{T-t+1}{2}\boldsymbol{p} \geq \boldsymbol{2}$ with high probability. Consider the time period $t \in [T_{k-1}, T_k)$.

1. When $T_{k-1} < \lceil \frac{T}{2} \rceil$, we have

$$\begin{aligned} d_j^t &\geq (T-t+1)p_j - (T-T_{k-1}+1)\sqrt{\frac{\log(t-1)}{T_{k-1}-1}} - \sqrt{(t-T_{k-1})\log(T-t+1)} \\ &\geq (T-t+1)p_j - (T-t+\omega+1)\sqrt{\frac{\log(t-1)}{t-\omega-1}} - \sqrt{\omega\log(T-t+1)}, \end{aligned}$$

with probability $1 - \frac{1}{(t-1)^2} - \frac{1}{(T-t+1)^2}$. There exist constants η_j^P and ζ_j^P such that when $t \in [\eta_j^P, \min\{T-\zeta_j^P, \lceil \frac{T}{2} \rceil + \omega\}]$, we have $\sqrt{\frac{\log(t-1)}{t-\omega-1}} \leq \frac{p_j}{4}$, $\frac{p_j}{4}(T-t+1) - \frac{p_j\omega}{4} - \sqrt{\omega\log(T-t+1)} \geq 0$, and hence $d_j^t \geq \frac{p_j(T-t+1)}{2} \geq 2$. Moreover, we have $\eta_j^P = O(\omega)$ and $\zeta_j^P = O(\omega\log\omega)$.

2. When $T_{k-1} \geq \lceil \frac{T}{2} \rceil$, we have

$$\begin{aligned} d_j^t &\geq (T-t+1)p_j - \sqrt{2(T-T_{k-1}+1)\log(T-t+1)} - \sqrt{(t-T_{k-1})\log(T-t+1)} \\ &\geq (T-t+1)p_j - \sqrt{2(T-t+\omega+1)\log(T-t+1)} - \sqrt{\omega\log(T-t+1)} \\ &\geq (T-t+1)p_j - \sqrt{2(T-t+1)\log(T-t+1)} - \sqrt{2\omega\log(T-t+1)} - \sqrt{\omega\log(T-t+1)}, \end{aligned}$$

with probability $1 - \frac{2}{(T-t+1)^2}$. Then, there exists a constant $\bar{\zeta}_j^P$ such that when $t \leq T - \bar{\zeta}_j^P$, we have $d_j^t \geq \frac{p_j(T-t+1)}{2} \geq 2$. Moreover, we have $\bar{\zeta}_j^P = O(\omega\log\omega)$.

To summarize, by defining $\eta^P = \max_j \eta_j^P = O(\omega)$ and $\zeta^P = \max\{\max_j \zeta_j^P, \max_j \bar{\zeta}_j^P\} = O(\omega\log\omega)$, we finish the proof.

Proof of Lemma 10. Consider the case when a type-j customer arrives at time $t \in [T_{k-1}, T_k)$ and the AIR policy accepts the request. Similar to Proposition 4, we discuss two cases.

1. When $T_{k-1} < \lceil \frac{T}{2} \rceil$, we have

$$\begin{aligned}
&\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \\
\geq &\frac{p_j(T-t+1)}{4} - \kappa_4(T-T_{k-1}+1)\sqrt{\frac{\log(t-1)}{T_{k-1}-1}} - \kappa_4\sqrt{(T-T_{k-1}+1)\log(T-t+1)} \\
\geq &\frac{p_j(T-t+1)}{4} - \kappa_4(T-t+1+\omega)\sqrt{\frac{\log(t-1)}{t-\omega-1}} - \kappa_4\sqrt{(T-t+\omega+1)\log(T-t+1)}.
\end{aligned}$$

There exists a constant ι_j^P such that when $t \geq \iota_j^P$, we have $2\kappa_4\sqrt{\frac{\log(t-1)}{t-\omega-1}} \leq \frac{p_j}{8}$. Moreover, there exists a constant θ_j^P such that when $t \leq T - \theta_j^P$, we have $\frac{(T-t+1)p_j}{8} - \frac{\kappa_4 \omega p_j}{8} - \kappa_4\sqrt{(T-t+\omega+1)\log(T-t+1)} \geq 1$. Therefore, when $t \in [\iota_j^P, \min\{\lceil \frac{T}{2} \rceil + \omega, T - \theta_j^P\}]$, we have $\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq 1$. Moreover, we have $\iota_j^P = O(\omega)$ and $\theta_j^P = O(\omega \log \omega)$.

2. When $T_{k-1} \geq \lceil \frac{T}{2} \rceil$, we have

$$\begin{aligned}
\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq &\frac{p_j(T-t+1)}{4} - \kappa_4(\sqrt{2}+1)\sqrt{(T-T_{k-1}+1)\log(T-t+1)} \\
\geq &\frac{p_j(T-t+1)}{4} - \kappa_4(\sqrt{2}+1)\sqrt{(T-t+\omega+1)\log(T-t+1)}.
\end{aligned}$$

There exists a constant $\bar{\theta}_j^P$ such that when $\lceil \frac{T}{2} \rceil \leq t \leq T - \bar{\theta}_j^P$, we have $\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq 1$. Moreover, we have $\bar{\theta}_j^P = O(\omega \log \omega)$. Then, similar to Proposition 4, we can finish the proof.

□

C.16. Proof of Lemma 4 Similar to Lemma 3, we only need the following inequalities:

$$d_j^t \geq (T-t+1)p_j - \sqrt{(t-T_{k-1})\log(T-t+1)} \geq (T-t+1)p_j - \sqrt{2(T-t+1)\log(T-t+1)}$$

$$\begin{aligned}
\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j) \geq &\frac{p_j(T-t+1)}{4} - \kappa_7\sqrt{(T-T_{k-1}+1)\log(T-t+1)} \\
\geq &\frac{p_j(T-t+1)}{4} - \kappa_7\sqrt{(2(T-t+1)+1)\log(T-t+1)}.
\end{aligned}$$

Then, following a similar proof, we can derive the constant bound $O(1)$. □

C.17. Proof of Proposition 7 Similar to Proposition 6, we only need the following inequalities:

1. When $T_{k-1} < \lceil \frac{T}{2} \rceil$, we have

$$d_j^t \geq (T-t+1)p_j - (T-T_{k-1}+1)\sqrt{\frac{\log(t-1)}{T_{k-1}-1}} - \sqrt{(t-T_{k-1})\log(T-t+1)}$$

$$\geq (T-t+1)p_j - (2(T-t+1)+1)\sqrt{\frac{2\log(t-1)}{t+1}} - \sqrt{(2(T-t+1)+1)\log(T-t+1)},$$

and

$$\begin{aligned}&\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j)\\ \geq &\frac{p_j(T-t+1)}{4} - \kappa_4(T-T_{k-1}+1)\sqrt{\frac{\log(t-1)}{T_{k-1}-1}} - \kappa_4\sqrt{(T-T_{k-1}+1)\log(T-t+1)}\\ \geq &\frac{p_j(T-t+1)}{4} - \kappa_4\,(2(T-t+1)+1)\sqrt{\frac{2\log(t-1)}{t+1}} - \kappa_4\sqrt{(2(T-t+1)+1)\log(T-t+1)}.\end{aligned}$$

2. When $T_{k-1} \geq \lceil \frac{T}{2} \rceil$, we have

$$\begin{aligned}d_j^t &\geq (T-t+1)p_j - (T-T_{k-1}+1)\sqrt{\frac{\log(T-t+1)}{T_{k-1}-1}} - \sqrt{(t-T_{k-1})\log(T-t+1)}\\ &\geq (T-t+1)p_j - (2(T-t+1)+1)\sqrt{\frac{\log(T-t+1)}{(2(T-t+1)+1)}} - \sqrt{(2(T-t+1)+1)\log(T-t+1)},\end{aligned}$$

and

$$\begin{aligned}&\bar{\mathfrak{S}}(\boldsymbol{b}^t, \boldsymbol{Z}^t, j)\\ \geq &\frac{p_j(T-t+1)}{4} - \kappa_4(T-T_{k-1}+1)\sqrt{\frac{\log(t-1)}{T_{k-1}-1}} - \kappa_4\sqrt{(T-T_{k-1}+1)\log(T-t+1)}\\ \geq &\frac{p_j(T-t+1)}{4} - \kappa_4\,(2(T-t+1)+1)\sqrt{\frac{\log(t-1)}{(2(T-t+1)+1)}} - \kappa_4\sqrt{(2(T-t+1)+1)\log(T-t+1)}.\end{aligned}$$

Then, following a similar proof, we can derive the constant bound $O(1)$. □

C.18. Proof of Proposition 8 In order to derive the same result as Theorem 3, we need to prove results analogous to Lemmas 6 and 7. Let $\boldsymbol{\pi}^\ell$ denote the state distribution at period ℓ. First, for $t \in [T_{k-1}, T_k)$, we have

$$\begin{aligned}|d_j^t - Z_j^t| = \left|d_j^{T_{k-1}} - Z_j^{T_{k-1}}\right| &= \left|\sum_{i=1}^{L} \pi_i^* \cdot \sigma_j^i \cdot (T-T_{k-1}+1) - \sum_{\ell=T_{k-1}}^{T} Y_j^\ell\right|\\ &\leq \sum_{\ell=T_{k-1}}^{T}\sum_{i=1}^{L}\left|\pi_i^* - \pi_i^\ell\right| \cdot \sigma_j^i + \left|\sum_{\ell=T_{k-1}}^{T} Y_j^\ell - \sum_{\ell=T_{k-1}}^{T}\sum_{i=1}^{L}\pi_i^\ell \cdot \sigma_j^i\right|.\end{aligned}$$

According to Theorem 15.0.1 in Meyn and Tweedie (2012) [42], there exist constants $c_{15} > 0$ and $\psi \in (0,1)$ such that $\|\boldsymbol{\pi}^\ell - \boldsymbol{\pi}^*\|_\infty \leq c_{15} \cdot \psi^\ell$. Thus, we have

$$\sum_{\ell=T_{k-1}}^{T}\sum_{i=1}^{L}\left|\pi_i^* - \pi_i^\ell\right| \cdot \sigma_j^i \leq \sum_{\ell=T_{k-1}}^{T} L \cdot \|\boldsymbol{\pi}^\ell - \boldsymbol{\pi}^*\|_\infty \leq \sum_{\ell=T_{k-1}}^{T} L \cdot c_{15} \cdot \psi^\ell \leq \frac{L \cdot c_{15}}{1-\psi}.$$

Since $\sum_{\ell=T_{k-1}}^{T} Y_j^\ell - \sum_{\ell=T_{k-1}}^{T} \sum_{i=1}^{L} \pi_i^\ell \cdot \sigma_j^i$ is a martingale, Azuma's inequality (e.g., Theorem 7.2.1 in Alon and Spencer 2016 [2]) implies that

$$\mathbb{P}\left(\left| \sum_{\ell=T_{k-1}}^{T} Y_j^\ell - \sum_{\ell=T_{k-1}}^{T} \sum_{i=1}^{L} \pi_i^\ell \cdot \sigma_j^i \right| \geq \sqrt{(T - T_{k-1} + 1)\log(T - t + 1)} \right) \leq \frac{2}{(T - t + 1)^2}.$$

Therefore, for $t \in [T_{k-1}, T_k)$, with probability at least $1 - \frac{2}{(T-t+1)^2}$, we have

$$|d_j^t - Z_j^t| \leq \frac{L \cdot c_{15}}{1 - \psi} + \sqrt{(T - T_{k-1} + 1)\log(T - t + 1)},$$

which is similar to Lemma 6 except for an additional constant $\frac{L \cdot c_{15}}{1-\psi}$.

Similarly, for $t \in [T_{k-1}, T_k)$, by Azuma's inequality, we have

$$\mathbb{P}\left(d_j^t \leq \sum_{i=1}^{L} \pi_i^* \cdot \sigma_j^i \cdot (T - t + 1) - \sqrt{(t - T_{k-1})\log(T - t + 1)} - \frac{L \cdot c_{15}}{1 - \psi} \right) \leq \frac{1}{(T - t + 1)^2},$$

which is similar to Lemma 7 except for an additional constant $\frac{L \cdot c_{15}}{1-\psi}$.

Since the constant term $\frac{L \cdot c_{15}}{1-\psi}$ is dominated by polynomial terms, we can still derive a constant regret bound following the proof idea of Theorem 3. □

Appendix D: Numerical Details In this section, we provide the omitted details in Section 4.

D.1. OLP Algorithms

Algorithm 3 Argmax with frequent resolving (AFR) policy

Initialize $\boldsymbol{b}^1 \leftarrow T\boldsymbol{\rho}$ and $\boldsymbol{N}^1 \leftarrow \boldsymbol{0}$.
for $t = 1, 2, 3, \ldots, T$ **do**
  if $t = 1$ **then**
    $\hat{p}_j^t \leftarrow 1/n$ for each j
  else
    Compute the empirical estimations $\hat{p}_j^t \leftarrow N_j^t/(t-1)$ for each j.
  end if
  Solve the fluid problem $\phi(\boldsymbol{b}^t, (T-t+1)\hat{\boldsymbol{p}}^t)$ and obtain its optimal solution $\boldsymbol{y}^*$.
  Observe arrival type j and set $\boldsymbol{N}^{t+1} \leftarrow \boldsymbol{N}^t + \boldsymbol{e}_j$.
  if $\boldsymbol{A}_j \leq \boldsymbol{b}^t$ and $y_j^* \geq (T-t+1)\hat{p}_j^t - y_j^*$ **then** ▷ Argmax between y_j^* and $(T-t+1)\hat{p}_j^t - y_j^*$
    Accept the request.
    Set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t - \boldsymbol{A}_j$.

else

Reject the request.

Set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t$.

end if

end for

Algorithm 4 Adaptive allocation (ADA) policy

Initialize $\boldsymbol{b}^1 \leftarrow T\boldsymbol{\rho}$ and $\boldsymbol{N}^1 \leftarrow \boldsymbol{0}$.

for $t = 1, 2, 3, \ldots, T$ **do**

if $t = 1$ **then**

$\hat{p}^t_j \leftarrow 1/n$ for each j

else

Compute the empirical estimations $\hat{p}^t_j \leftarrow N^t_j/(t-1)$ for each j.

end if

Solve the fluid problem $\phi(\boldsymbol{b}^t, (T-t+1)\hat{\boldsymbol{p}}^t)$ and obtain its optimal solution $\boldsymbol{y}^*$.

Observe arrival type j and set $\boldsymbol{N}^{t+1} \leftarrow \boldsymbol{N}^t + \boldsymbol{e}_j$.

if $\boldsymbol{A}_j \le \boldsymbol{b}^t$ **then**

Accept the request with probability $y^*_j / \Big((T-t+1)\hat{p}^t_j\Big)$. ▷ Probabilistic Allocation.

If accepted, set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t - \boldsymbol{A}_j$; otherwise, set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t$.

else

Reject the request.

Set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t$.

end if

end for

Algorithm 5 Simple and fast (SFA) policy

Initialize $\boldsymbol{b}^1 \leftarrow T\boldsymbol{\rho}$ and $\boldsymbol{q}^1 = \boldsymbol{0}$.

for $t = 1, 2, 3, \ldots, T$ **do**

Observe arrival type j.

Set $x^t \leftarrow 1$ if $r_j > \boldsymbol{A}_j \cdot \boldsymbol{q}^t$ and $x^t \leftarrow 0$ otherwise.

Compute $\boldsymbol{q}^{t+1} \leftarrow \boldsymbol{q}^t + \frac{1}{\sqrt{t}}(\boldsymbol{A}_j x^t - \boldsymbol{\rho})$.

Compute $\boldsymbol{q}^{t+1} \leftarrow \max\{\boldsymbol{q}^{t+1}, \boldsymbol{0}\}$.

if $\boldsymbol{A}_j \le \boldsymbol{b}^t$ and $x^t = 1$ **then**

Accept the request and set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t - \boldsymbol{A}_j$.

else

Reject the request and set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t$.

end if

end for

Algorithm 6 Decoupling learning and decision (DLD) policy

Input: $T_e = \lfloor T^{2/3} \rfloor$, $\alpha_e = T^{-1/3}$ and $\alpha_p = T^{-2/3}$.

Initialize $\boldsymbol{b}^1 \leftarrow T\boldsymbol{\rho}$, $\boldsymbol{q}_D^1 \leftarrow \boldsymbol{0}$ and $\boldsymbol{q}_L^1 \leftarrow \boldsymbol{0}$.

for $t = 1, 2, 3, \ldots, T_e$ **do**

Observe arrival type j.

Set $\tilde{x}_D^t \leftarrow 1$ if $r_j > \boldsymbol{A}_j \cdot \boldsymbol{q}_D^t$ and $\tilde{x}_D^t \leftarrow 0$ otherwise.

Compute $\boldsymbol{q}_D^{t+1} \leftarrow \boldsymbol{q}_D^t + \alpha_e(\boldsymbol{A}_j \tilde{x}_D^t - \boldsymbol{\rho})$.

Compute $\boldsymbol{q}_D^{t+1} \leftarrow \max\{\boldsymbol{q}_D^{t+1}, \boldsymbol{0}\}$.

if $\boldsymbol{A}_j \le \boldsymbol{b}^t$ and $\tilde{x}_D^t = 1$ **then**

Accept the request and set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t - \boldsymbol{A}_j$.

else

Reject the request and set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t$.

end if

Set $\tilde{x}_L^t \leftarrow 1$ if $r_j > \boldsymbol{A}_j \cdot \boldsymbol{q}_L^t$ and $\tilde{x}_L^t \leftarrow 0$ otherwise.

Compute $\boldsymbol{q}_L^{t+1} \leftarrow \boldsymbol{q}_L^t + \frac{1}{t}(\boldsymbol{A}_j \tilde{x}_L^t - \boldsymbol{\rho})$.

Compute $\boldsymbol{q}_L^{t+1} \leftarrow \max\{\boldsymbol{q}_L^{t+1}, \boldsymbol{0}\}$.

end for

Set $\boldsymbol{q}_D^{T_e+1} \leftarrow \boldsymbol{q}_L^{T_e+1}$.

for $t = T_e + 1, T_e + 2, \ldots, T$ **do**

Observe arrival type j.

Set $\tilde{x}_D^t \leftarrow 1$ if $r_j > \boldsymbol{A}_j \cdot \boldsymbol{q}_D^t$ and $\tilde{x}_D^t \leftarrow 0$ otherwise.

Compute $\boldsymbol{q}_D^{t+1} \leftarrow \boldsymbol{q}_D^t + \alpha_p(\boldsymbol{A}_j \tilde{x}_D^t - \boldsymbol{\rho})$.

Compute $\boldsymbol{q}_D^{t+1} \leftarrow \max\{\boldsymbol{q}_D^{t+1}, \boldsymbol{0}\}$.

if $\boldsymbol{A}_j \le \boldsymbol{b}^t$ and $\tilde{x}_D^t = 1$ **then**

Accept the request and set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t - \boldsymbol{A}_j$.

else

Reject the request and set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t$.

end if

end for

Algorithm 7 Budget-updating fast (BUF) policy

Input: Time set $\mathcal{T} = \left\{T - \left\lceil \frac{T}{2^k} \right\rceil : k = 1, 2, \ldots, \lceil \log_2 T \rceil \right\}$.

Initialize $\boldsymbol{b}^1 \leftarrow T\boldsymbol{\rho}$, $\boldsymbol{d}^1 \leftarrow \boldsymbol{\rho}$, $\boldsymbol{q}^1 \leftarrow \boldsymbol{0}$, and $l \leftarrow 1$.

for $t = 1, 2, 3, \ldots, T$ **do**

  Observe arrival type j.

  Set $\tilde{x}^t \leftarrow 1$ if $r_j > \boldsymbol{A}_j \cdot \boldsymbol{q}^t$ and $\tilde{x}^t \leftarrow 0$ otherwise.

  if $\boldsymbol{b}^t - \boldsymbol{A}_j \geq \boldsymbol{0}$ and $\tilde{x}^t = 1$ **then**

    Accept the request and set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t - \boldsymbol{A}_j$.

  else

    Reject the request and set $\boldsymbol{b}^{t+1} \leftarrow \boldsymbol{b}^t$.

  end if

  if $t + 1 \in \mathcal{T}$ **then**

    Set $l \leftarrow t + 1$ and $\boldsymbol{d}^{t+1} \leftarrow \frac{\boldsymbol{b}^{t+1}}{T-t}$.

  else

    Set $\boldsymbol{d}^{t+1} \leftarrow \boldsymbol{d}^t$

  end if

  Set $\boldsymbol{q}^{t+1} \leftarrow \boldsymbol{q}^t + \frac{1}{t-l+2}(\boldsymbol{A}_j\tilde{x}^t - \boldsymbol{d}^{t+1})$

end for

D.2. Numerical Setup of the Multi-resource Case In this section, we present the parameters randomly generated for the multi-resource case with $m = 10$ and $n = 2$.

$$\boldsymbol{A} = \begin{bmatrix} 0.226 & 0.146 \\ 0.957 & 0.916 \\ 0.005 & 0.876 \\ 0.457 & 0.790 \\ 0.285 & 0.960 \\ 0.572 & 0.736 \\ 0.701 & 0.206 \\ 0.093 & 0.642 \\ 0.903 & 0.923 \\ 0.743 & 0.789 \end{bmatrix} \quad \boldsymbol{\rho} = \begin{bmatrix} 0.128 \\ 0.805 \\ 0.770 \\ 0.695 \\ 0.844 \\ 0.647 \\ 0.181 \\ 0.564 \\ 0.812 \\ 0.694 \end{bmatrix} \quad \boldsymbol{p} = \begin{bmatrix} 0.121 \\ 0.879 \end{bmatrix} \quad \boldsymbol{r} = \begin{bmatrix} 0.689 \\ 0.710 \end{bmatrix}$$