This page is intentionally blank. Proper e-companion title page, with INFORMS branding and exact metadata of the main paper, will be produced by the INFORMS office when the issue is being assembled.

Proofs and Auxiliary Results

In this e-companion, we provide the proofs of results in the main text, and discuss few extensions of our baseline model.

EC.1. Proofs from Section 3

In this section, we present the missing proofs from Section 3.

Proof of Lemma 2. The proof is immediate for $\lambda < 1$, since the sets Π_{AP} and Π_{SM} are compact (under the topology of weak convergence) and $W(\pi,\theta)$ is continuous in π for each $\theta \in [0,1]$. To see this, note that for any $\pi \in \Pi_{\mathsf{AP}}$, by Lemma EC.1 (stated at the end of this section), we have $\pi_k \le \lambda^k \pi_0$ for all k. Hence, for $\lambda < 1$, Prohorov's theorem (Aliprantis and Border 2006) directly implies compactness of Π_{AP} and Π_{SM}.

Thus, for the rest of the proof, suppose $\lambda = 1$. Let $\Pi^{\mathsf{fi}}_{\mathsf{AP}}$ and $\Pi^{\mathsf{fi}}_{\mathsf{SM}}$ denote the set of admission policies and signaling mechanisms that are not Pareto-dominated by the full-information mechanism. We again use Prohorov's theorem to show that these sets are relatively compact, implying the existence of a maximizer of $W(\cdot,\theta)$ over the closures of these sets. The result then follows since any maximizer of $W(\pi,\theta)$ over the closure of $\Pi^{\mathsf{fi}}_{\mathsf{AP}}$ ($\Pi^{\mathsf{fi}}_{\mathsf{SM}}$) is also a maximizer over Π_{AP} (resp., Π_{SM}).

Thus, we first show that the set of distributions $\Pi^{\mathsf{fi}}_{\mathsf{AP}}$ is tight. Fix an $\epsilon > 0$. Let $W_{\mathsf{L}}(\mathsf{fi})$ and $W_{\mathsf{H}}(\mathsf{fi})$ denote the welfare of each type under the full information mechanism. Next, fix some large enough N to be chosen later. Consider a steady-state distribution $\pi \in \Pi_{\mathsf{AP}}$. We have the following expressions:

$$
\begin{aligned}
W_{\mathsf{H}}(\pi) &= \lambda_{\mathsf{H}} \sum_{n=0}^{\infty} \pi_n u_{\mathsf{H}}(n) \\
&\le \lambda_{\mathsf{H}} u_{\mathsf{H}}(0) \left(\sum_{n<N} \pi_n \right) + \lambda_{\mathsf{H}} u_{\mathsf{H}}(N) \left(\sum_{n \ge N} \pi_n \right) \\
&\le \lambda_{\mathsf{H}} u_{\mathsf{H}}(0) + \lambda_{\mathsf{H}} u_{\mathsf{H}}(N) \left(\sum_{n \ge N} \pi_n \right).
\end{aligned}
$$

where the first inequality follows from Assumption 1 and the second follows because $u_{\mathsf{H}}(0) > 0$. Similarly, we have for large enough $N > m_{\mathsf{fi}}$,

$$
\begin{aligned}
W_{\mathsf{L}}(\pi) &= \sum_{n=0}^{\infty} (\pi_{n+1} - \lambda_{\mathsf{H}} \pi_n) u_{\mathsf{L}}(n) \\
&\le u_{\mathsf{L}}(0) \left(\sum_{n<N} (\pi_{n+1} - \lambda_{\mathsf{H}} \pi_n) \right) + u_{\mathsf{L}}(N) \left(\sum_{n \ge N} (\pi_{n+1} - \lambda_{\mathsf{H}} \pi_n) \right) \\
&= u_{\mathsf{L}}(0) \left(\pi_N - \pi_0 + (1-\lambda_{\mathsf{H}}) \sum_{n=0}^{N-1} \pi_n \right) + u_{\mathsf{L}}(N) \left(-\lambda_{\mathsf{H}} \pi_N + (1-\lambda_{\mathsf{H}}) \sum_{n>N} \pi_n \right)
\end{aligned}
$$

$$
\begin{aligned}
&\leq (1-\lambda_{\mathsf{H}})u_{\mathsf{L}}(0) + u_{\mathsf{L}}(N)\left(-\pi_N + (1-\lambda_{\mathsf{H}})\sum_{n\geq N}\pi_n\right)\\
&\leq \lambda_{\mathsf{L}} u_{\mathsf{L}}(0) + u_{\mathsf{L}}(N)\left(-\frac{1}{N+1} + \lambda_{\mathsf{L}}\sum_{n\geq N}\pi_n\right),
\end{aligned}
$$

where the final inequality follows from (1) $\lambda_{\mathsf{L}} = 1-\lambda_{\mathsf{H}}$, (2) $u_{\mathsf{L}}(N) < 0$, and (3) $(N+1)\pi_N \leq \sum_{n=0}^{N}\pi_n \leq 1$ because of the detailed-balance conditions $\pi_n \leq \pi_{n-1}$. Thus, for $N \geq N_0^{\epsilon} = \frac{2}{\epsilon\lambda_{\mathsf{L}}}$, we obtain

$$
W_{\mathsf{L}}(\pi) \leq \lambda_{\mathsf{L}} u_{\mathsf{L}}(0) + \lambda_{\mathsf{L}} u_{\mathsf{L}}(N)\left(-\frac{\epsilon}{2} + \sum_{n\geq N}\pi_n\right).
$$

Since $\lim_{n\to\infty} u_i(n) = -\infty$, let N_1^{ϵ} be large enough so that $\max\{u_{\mathsf{L}}(k), u_{\mathsf{H}}(k)\} < -\frac{2}{\epsilon^2}$ for $k \geq N_1^{\epsilon}$. Then, for all $N \geq N^{\epsilon} = \max\{N_0^{\epsilon}, N_1^{\epsilon}\}$, we have

$$
\begin{aligned}
W_{\mathsf{H}}(\pi) &\leq \lambda_{\mathsf{H}}\left(u_{\mathsf{H}}(0) - \frac{2}{\epsilon^2}\sum_{n\geq N}\pi_n\right)\\
W_{\mathsf{L}}(\pi) &\leq \lambda_{\mathsf{L}}\left(u_{\mathsf{L}}(0) - \frac{2}{\epsilon^2}\left(-\frac{\epsilon}{2} + \sum_{n\geq N}\pi_n\right)\right).
\end{aligned}
$$

Now, for any $\pi \in \Pi_{\mathsf{AP}}$, if $\sum_{n\geq N}\pi_n \geq \epsilon$, then we have $W_i(\pi) \leq \lambda_i u_i(0) - \frac{\lambda_i}{\epsilon}$ for $i \in \{\mathsf{L},\mathsf{H}\}$. For small enough $\epsilon > 0$, we obtain that $W_i(\pi) < W_i(\mathsf{fi})$ and hence π is Pareto dominated by the full-information mechanism, implying $\pi \notin \Pi_{\mathsf{AP}}^{\mathsf{fi}}$. Thus, we conclude that for all small enough $\epsilon > 0$, there exists an N^{ϵ} such that for all $\pi \in \Pi_{\mathsf{AP}}^{\mathsf{fi}}$, we have $\sum_{n\geq N^{\epsilon}}\pi_n < \epsilon$. Thus, the set of distributions $\Pi_{\mathsf{AP}}^{\mathsf{fi}}$ is tight.

Using Prohorov's theorem, we then conclude that $\Pi_{\mathsf{AP}}^{\mathsf{fi}}$ is relatively compact (under weak topology). The set $\Pi_{\mathsf{SM}}^{\mathsf{fi}}$, being a subset of $\Pi_{\mathsf{AP}}^{\mathsf{fi}}$, is also relatively compact. Since $W(\pi,\theta)$ is continuous in $\pi \in \Pi_{\mathsf{AP}}^{\mathsf{fi}}$ for any $\theta \in [0,1]$, we obtain that the maximizer of $W(\pi,\theta)$ over the closure of $\Pi_{\mathsf{AP}}^{\mathsf{fi}}$ (and, separately, $\Pi_{\mathsf{SM}}^{\mathsf{fi}}$) exists and is Pareto-efficient within Π_{AP} (resp., Π_{SM}). $\square$

Proof of Theorem 1. Let $\pi \in \Pi_{\mathsf{AP}}$ be such that there exists an $m \geq 0$ with $\pi_{m+1} < \lambda\pi_m$ and $\lambda_{\mathsf{H}}\pi_{m+1} < \pi_{m+2}$. In words, this implies that under π, an arriving L-type user is asked to leave with positive probability if the queue length is m, and asked to join with positive probability if the queue length is $m+1$. We now show that such a π cannot be Pareto-efficient within Π_{SM}. We do this by constructing an $\hat{\pi} \in \Pi_{\mathsf{SM}}$ that Pareto-dominates π.

Towards that end, consider the following perturbation of π for small enough $\delta > 0$:

$$
\hat{\pi}_k = \begin{cases} \pi_k & \text{if } k < m+1;\\ \pi_{m+1} + \delta\sum_{n>m+1}\pi_n & \text{if } k = m+1;\\ \pi_k(1-\delta) & \text{if } k > m+1. \end{cases}
$$

First, it is straightforward to verify that $\hat{\pi}$ satisfies the detailed balance constraints in Lemma 1 for all small $\delta > 0$. In addition, we have

$$\begin{aligned} J(\hat{\pi}) &= \sum_{k=0}^{\infty}(\pi_{k+1} - \lambda_{\mathsf{H}}\pi_k)u_{\mathsf{L}}(k) + \delta\left(\sum_{k>m+1}\pi_k\right)(u_{\mathsf{L}}(m) - \lambda_{\mathsf{H}}u_{\mathsf{L}}(m+1)) \\ &\quad - \delta\pi_{m+2}u_{\mathsf{L}}(m+1) - \delta\sum_{k>m+1}(\pi_{k+1} - \lambda_{\mathsf{H}}\pi_k)u_{\mathsf{L}}(k) \\ &= J(\pi) + \delta\cdot\sum_{k>m+1}\pi_k\cdot\left(\ u_{\mathsf{L}}(m) - u_{\mathsf{L}}(k-1) - \lambda_{\mathsf{H}}(u_{\mathsf{L}}(m+1) - u_{\mathsf{L}}(k))\ \right). \end{aligned}$$

Now, as $\lambda_{\mathsf{H}} < 1$, for any $k > m+1$, we have

$$\begin{aligned} u_{\mathsf{L}}(m) - u_{\mathsf{L}}(k-1) - \lambda_{\mathsf{H}}(u_{\mathsf{L}}(m+1) - u_{\mathsf{L}}(k)) &> u_{\mathsf{L}}(m) - u_{\mathsf{L}}(k-1) - u_{\mathsf{L}}(m+1) - u_{\mathsf{L}}(k) \\ &= (u_{\mathsf{L}}(m) - u_{\mathsf{L}}(m+1)) - (u_{\mathsf{L}}(k-1) - u_{\mathsf{L}}(k)) \\ &\geq 0, \end{aligned}$$

where we have used Assumption 1 in both inequalities. Specifically, the first inequality follows from the fact that $u_{\mathsf{L}}(k)$ is strictly decreasing in k and hence $u_{\mathsf{L}}(m+1) - u_{\mathsf{L}}(k) > 0$, and the second inequality follows from the fact that $u_{\mathsf{L}}(n) - u_{\mathsf{L}}(n+1)$ is non-increasing in n. Using this and the fact that $\pi_{m+2} > \lambda_{\mathsf{H}}\pi_{m+1} \geq 0$, we obtain that $J(\hat{\pi}) > J(\pi) \geq 0$. Hence the obedience constraint (JOIN) holds for $\hat{\pi}$.

By similar algebraic steps, we have

$$L(\hat{\pi}) = L(\pi) - \delta\cdot\sum_{k>m+1}\pi_k\cdot\left(\ u_{\mathsf{L}}(m) - u_{\mathsf{L}}(k-1) - \lambda(u_{\mathsf{L}}(m+1) - u_{\mathsf{L}}(k))\ \right).$$

Using the fact that $\lambda \leq 1$, by a similar argument as before, we obtain that the parenthetical term is non-negative, and hence $L(\hat{\pi}) \leq L(\pi) \leq 0$. Thus, the obedience constraint (LEAVE) also holds for $\hat{\pi}$. Taken together, this implies we have $\hat{\pi} \in \Pi_{\mathsf{SM}}$.

Next, note that

$$\begin{aligned} W_{\mathsf{H}}(\hat{\pi}) &= \lambda_{\mathsf{H}}\sum_{n=0}^{\infty}\hat{\pi}_n u_{\mathsf{H}}(n) \\ &= W_{\mathsf{H}}(\pi) + \lambda_{\mathsf{H}}\delta\cdot\left(\sum_{k>m+1}\pi_k\cdot(u_{\mathsf{H}}(m+1) - u_{\mathsf{H}}(k))\right). \end{aligned}$$

Since $u_{\mathsf{H}}(k)$ is strictly decreasing in k, we obtain $W_{\mathsf{H}}(\hat{\pi}) \geq W_{\mathsf{H}}(\pi)$. Finally, we have $W_{\mathsf{L}}(\hat{\pi}) = J(\hat{\pi}) > J(\pi) = W_{\mathsf{L}}(\pi)$. Thus, we obtain that $\hat{\pi}$ Pareto-dominates π.

From the above, we conclude that for any Pareto-efficient signaling mechanism $\pi \in \Pi_{\mathsf{SM}}$, it must be the case that whenever there exists an $m \geq 0$ with $\pi_{m+1} < \lambda\pi_m$, we have $\pi_{m+2} = \lambda_{\mathsf{H}}\pi_{m+1}$. This implies that π must have one of the following two structures:

1. for all $m \geq 0$, we have $\pi_{m+1} = \lambda \pi_m$; OR
2. there exists an $m \geq 0$ such that $\pi_{k+1} = \lambda \pi_k$ for $k < m$, $\pi_{m+1} < \lambda \pi_m$ and $\pi_{k+1} = \lambda_{\mathsf{H}} \pi_k$ for $k > m$.

In the first case, we have L-type users being asked to join the queue for all queue length, implying that π trivially has a threshold structure (with threshold equal to ∞). In the second case, the L-type users are asked to join with probability 1 for queue-lengths strictly less than m and asked to leave with probability 1 for queue-lengths strictly greater than m. Again, this implies a threshold structure for π, with threshold in the interval $[m, m+1]$.

Having shown the threshold structure of Pareto efficient signaling mechanisms, next we show that the corresponding threshold is less than or equal to the full-information threshold m_{fi}. Let $\pi \in \Pi_{\mathsf{SM}}$ have a threshold structure, with a threshold $x > m_{\mathsf{fi}}$, where $x = m + a$ with $m \in \mathbb{N}_0$ and $a \in [0,1]$. Thus, we have $\pi_{k+1} = \lambda \pi_k$ for all $k < m$, and $\pi_{k+1} = \lambda_{\mathsf{H}} \pi_k$ for all $k > m$. Note, we allow $m = \infty$, which captures the case where $\pi_{k+1} = \lambda \pi_k$ for all $k \in \mathbb{N}_0$. Observe that $x > m_{\mathsf{fi}}$ implies that $m \geq m_{\mathsf{fi}}$, and hence the threshold structure of π implies $\pi_{m_{\mathsf{fi}}} > 0$.

We prove that such a distribution π cannot be Pareto efficient by constructing a $\hat{\pi} \in \Pi_{\mathsf{SM}}$ which Pareto dominates π. Consider $\hat{\pi}$ defined as follows:

$$\hat{\pi}_k = \begin{cases} \frac{1}{Z} \pi_k & \text{if } k \leq m_{\mathsf{fi}}; \\ \frac{1}{Z} \lambda_{\mathsf{H}}^{k - m_{\mathsf{fi}}} \pi_{m_{\mathsf{fi}}} & \text{if } k > m_{\mathsf{fi}}, \end{cases}$$

where $Z = \sum_{k \leq m_{\mathsf{fi}}} \pi_k + \pi_{m_{\mathsf{fi}}} \sum_{k > m_{\mathsf{fi}}} \lambda_{\mathsf{H}}^{k - m_{\mathsf{fi}}}$. Using the detailed balance constraints in Lemma 1, it follows that $\pi_k \geq \lambda_{\mathsf{H}}^{k - m_{\mathsf{fi}}} \pi_{m_{\mathsf{fi}}}$ for all $k > m_{\mathsf{fi}}$. Thus, as $\sum_k \pi_k = 1$, we have $Z \leq 1$.

Next, consider

$$\begin{aligned} J(\hat{\pi}) = \sum_{k=1}^{\infty} (\hat{\pi}_{k+1} - \lambda_{\mathsf{H}} \hat{\pi}_k)\, u_{\mathsf{L}}(k) &= \frac{1}{Z} \sum_{k < m_{\mathsf{fi}}} (\pi_{k+1} - \lambda_{\mathsf{H}} \pi_k)\, u_{\mathsf{L}}(k) \\ &> \frac{1}{Z} \sum_{k=1}^{\infty} (\pi_{k+1} - \lambda_{\mathsf{H}} \pi_k)\, u_{\mathsf{L}}(k) = \frac{1}{Z} \cdot J(\pi). \end{aligned}$$

Here, the inequality follows from the fact that $u_{\mathsf{L}}(k) < 0$ for $k \geq m_{\mathsf{fi}}$ and that $\pi_{m_{\mathsf{fi}}+1} - \lambda_{\mathsf{H}} \pi_{m_{\mathsf{fi}}} = \lambda_{\mathsf{L}} \min\{x - m_{\mathsf{fi}}, 1\} \pi_{m_{\mathsf{fi}}} > 0$. Since $J(\pi) \geq 0$ and $Z \leq 1$, we obtain that $J(\hat{\pi}) > J(\pi) \geq 0$. Hence, the obedience constraint (JOIN) holds for $\hat{\pi}$. Moreover, the threshold structure of π implies

$$L(\hat{\pi}) = \sum_{k=1}^{\infty} (\lambda \hat{\pi}_k - \hat{\pi}_{k+1})\, u_{\mathsf{L}}(k) = \sum_{k=m_{\mathsf{fi}}}^{\infty} (\lambda \hat{\pi}_k - \hat{\pi}_{k+1})\, u_{\mathsf{L}}(k) \leq 0.$$

Thus, the obedience constraint (LEAVE) also holds for $\hat{\pi}$. Hence, we obtain that $\hat{\pi} \in \Pi_{\mathsf{SM}}$.

Furthermore, for $\ell \leq m_{\mathsf{fi}}$, we have $\sum_{k \leq \ell} \hat{\pi}_k = \frac{1}{Z} \cdot \sum_{k \leq \ell} \pi_k$. Since, $Z \leq 1$, this implies $\sum_{k \leq \ell} \hat{\pi}_k \geq \sum_{k \leq \ell} \pi_k$ for all $\ell \leq m_{\mathsf{fi}}$. For $\ell > m_{\mathsf{fi}}$, after some algebra, we obtain

$$\sum_{k \leq \ell} \hat{\pi}_k - \sum_{k \leq \ell} \pi_k = \frac{1}{Z} \left(\sum_{q \leq m_{\mathsf{fi}}} \sum_{k > \ell} \pi_q \left(\pi_k - \pi_{m_{\mathsf{fi}}} \lambda_{\mathsf{H}}^{k - m_{\mathsf{fi}}} \right) + \sum_{q = m_{\mathsf{fi}}+1}^{\ell} \sum_{k > \ell} \pi_{m_{\mathsf{fi}}} \lambda_{\mathsf{H}}^{q - m_{\mathsf{fi}}} \left(\pi_k - \lambda_{\mathsf{H}}^{k-q} \pi_q \right) \right).$$

In Lemma EC.1 (stated at the end of this section), we show that $\pi_k \geq \pi_q \lambda_{\mathsf{H}}^{k-q}$ for all $k > q$. Thus, the right-hand side is non-negative, and hence, $\sum_{k\leq \ell} \hat{\pi}_k \geq \sum_{k\leq \ell} \pi_k$ for $\ell > m_{\mathsf{fi}}$ as well. Together, this implies that $\hat{\pi}$ is stochastically dominated by π. Since $u_{\mathsf{H}}(k)$ is strictly decreasing in k, we have

$$W_{\mathsf{H}}(\hat{\pi}) = \lambda_{\mathsf{H}} \sum_{k=0}^{\infty} \hat{\pi}_k u_{\mathsf{H}}(k) \geq \lambda_{\mathsf{H}} \sum_{k=0}^{\infty} \pi_k u_{\mathsf{H}}(k) = W_{\mathsf{H}}(\pi).$$

Finally, since $W_{\mathsf{L}}(\hat{\pi}) = J(\hat{\pi}) > J(\pi) = W_{\mathsf{L}}(\pi)$, we conclude that $\hat{\pi} \in \Pi_{\mathsf{SM}}$ Pareto-dominates π, and hence π cannot be Pareto-efficient within Π_{SM}. $\square$

LEMMA EC.1. *For any $\pi \in \Pi_{\mathsf{AP}}$, and for any $k > q \in \mathbb{N}_0$, we have $\pi_k \geq \lambda_{\mathsf{H}}^{k-q} \pi_q$ and $\pi_k \leq \lambda^{k-q} \pi_q$. In particular, when $\lambda_{\mathsf{H}} > 0$, for any $\pi \in \Pi_{\mathsf{AP}}$, we have $\pi_k \in (0,1)$ for all $k \in \mathbb{N}_0$.*

Proof. The proof follows immediately from the detailed balance constraints in Lemma 1. $\square$

Proof of Theorem 3. Since $\Pi_{\mathsf{SM}} \subset \Pi_{\mathsf{AP}}$, any signaling mechanism $\pi \in \Pi_{\mathsf{SM}}$ that is Pareto efficient within the class of admission policies must be so within the class of signaling mechanisms. Thus, it remains to show that for any signaling mechanism $\pi \in \Pi_{\mathsf{SM}}$ with $L(\pi) < 0$, if π is Pareto dominated by an admission policy, then it is Pareto dominated by a signaling mechanism.

By Theorem 1, we obtain that if π does not have a threshold structure, or if it has a threshold structure with threshold greater than the full-information threshold m_{fi}, then π is Pareto dominated within the class Π_{SM} of signaling mechanisms, and there is nothing to prove. Hence, suppose that π has a threshold structure with threshold smaller or equal to m_{fi}. This in turn implies that $J(\pi) > 0$, as a L-type user always receives non-negative utility upon joining the queue, and receives a positive utility if the queue is empty (which occurs with positive probability).

Since π is not Pareto-efficient within the class Π_{AP}, there exists an admission policy $\hat{\pi} \in \Pi_{\mathsf{AP}}$ that Pareto-dominates π. In particular, we have $W_i(\hat{\pi}) \geq W_i(\pi)$ for $i \in \{\mathsf{H}, \mathsf{L}\}$, with at least one inequality strict.

Next, let $\tilde{\pi} = (1-\epsilon)\pi + \epsilon\hat{\pi}$ for some $\epsilon \in (0,1]$ to be chosen later. By convexity of Π_{AP}, we have $\tilde{\pi} \in \Pi_{\mathsf{AP}}$. Furthermore, by linearity, we have $J(\tilde{\pi}) = (1-\epsilon)J(\pi) + \epsilon J(\hat{\pi})$ and $L(\tilde{\pi}) = (1-\epsilon)L(\pi) + \epsilon L(\hat{\pi})$. Since $J(\pi) > 0$ and $L(\pi) < 0$, for all small enough $\epsilon > 0$ we have $J(\tilde{\pi}) \geq 0$ and $L(\tilde{\pi}) \leq 0$. Thus, the obedience constraints (JOIN) and (LEAVE) hold for $\tilde{\pi}$, and hence $\tilde{\pi} \in \Pi_{\mathsf{SM}}$. Finally, again by linearity, we have

$$W_{\mathsf{L}}(\tilde{\pi}) = (1-\epsilon)W_{\mathsf{L}}(\pi) + \epsilon W_{\mathsf{L}}(\hat{\pi}) \geq W_{\mathsf{L}}(\pi)$$
$$W_{\mathsf{H}}(\tilde{\pi}) = (1-\epsilon)W_{\mathsf{H}}(\pi) + \epsilon W_{\mathsf{H}}(\hat{\pi}) \geq W_{\mathsf{H}}(\pi),$$

with at least one inequality strict. Thus, we obtain that the signaling mechanism $\tilde{\pi}$ Pareto-dominates π and hence π is not Pareto-efficient within the class Π_{SM}. $\square$

EC.2. Structural Results

Before proceeding to present the missing proofs of Sections 4 and 5, we present three structural results. First, in Lemma EC.2, we characterize the equilibrium structure under the no-information mechanism. Then, in Lemma EC.3, we study the shape of welfare functions $W_{\mathsf{L}}(\cdot)$ and $W_{\mathsf{H}}(\cdot)$ for threshold mechanisms. Finally, in Lemma EC.4, we study the function $L(\cdot)$ defined in (LEAVE). We remark that the last two lemmas are used in the proofs of results in Sections 4 and 5.

Lemma EC.2 **(Equilibrium structure under no-information mechanism)**. *For $p \in [0,1]$, let $\pi(p) \in \Pi_{\mathsf{AP}}$ be given by $\pi_n(p) = (1-\lambda_{\mathsf{L}} p - \lambda_{\mathsf{H}})(\lambda_{\mathsf{L}} p + \lambda_{\mathsf{H}})^n$. Then, the steady state distribution under the no-information mechanism* ni *is given by $\pi(p^{\mathsf{ni}}) \in \Pi_{\mathsf{SM}}$, for $p^{\mathsf{ni}} \in [0,1]$ that satisfies the following conditions:*

1. *if $\sum_{k=0}^{\infty} \lambda^k u_{\mathsf{L}}(k) \geq 0$ then $p^{\mathsf{ni}} = 1$;*
2. *if $\sum_{k=0}^{\infty} \lambda_{\mathsf{H}}^k u_{\mathsf{L}}(k) \leq 0$ then $p^{\mathsf{ni}} = 0$;*
3. *otherwise, $p^{\mathsf{ni}} \in (0,1)$ satisfies $\sum_{k=0}^{\infty} (\lambda_{\mathsf{L}} p^{\mathsf{ni}} + \lambda_{\mathsf{H}})^k u_{\mathsf{L}}(k) = 0$.*

Here, $p^{\mathsf{ni}} \in [0,1]$ denotes the probability under the no-information mechanism that a L*-type user joins the queue upon arrival.*

Proof of Lemma EC.2. First note that an arriving L-type user has no information about the queue length. Therefore, a symmetric equilibrium strategy consists of a probability p with which she joins the queue. Let $\pi_n(p)$ be the steady-state distribution corresponding to such a strategy. By detailed balance constraint, we have:

$$\pi_{n+1}(p) = (\lambda_{\mathsf{H}} + p\lambda_{\mathsf{L}})\pi_n(p), \; n \in \mathbb{N}_0$$

This implies $\pi_n(p) = (1-\lambda_{\mathsf{L}} p - \lambda_{\mathsf{H}})(\lambda_{\mathsf{L}} p + \lambda_{\mathsf{H}})^n$, $n \in \mathbb{N}_0$. A L-type users chooses p that maximizes her utility. This gives rise to the three cases listed in the statement of the lemma. $\square$

Lemma EC.3 **(Properties of welfare functions)**. 1. *The welfare function $W_{\mathsf{H}}(x)$ is strictly decreasing in $x \in \mathbb{R}_+$.*

2. *The welfare function $W_{\mathsf{L}}(x)$ is unimodal over $x \in \mathbb{R}_+$. Furthermore, $W_{\mathsf{L}}(x)$ is monotone between consecutive integers, initially increasing up to a maximum, and then decreasing.*
3. *The function $W(x,\theta) = \theta W_{\mathsf{L}}(x) + (1-\theta) W_{\mathsf{H}}(x)$ attains its maximum at an integer $m \leq m_{\mathsf{fi}}$.*

Proof of Lemma EC.3. The proof of the first statement follows from the fact the steady-state distribution under the threshold policy x is stochastically dominated by that under the threshold policy $\hat{x} > x$. Since $u_{\mathsf{H}}(k)$ is strictly decreasing in k, we thus obtain that $W_{\mathsf{H}}(x) > W_{\mathsf{H}}(\hat{x})$.

For the second statement, we show that (i) $W_{\mathsf{L}}(x)$ is monotone between consecutive integers, and (ii) $W_{\mathsf{L}}(x)$ is unimodal over integers, initially increasing up to a maximum, and then decreasing.

Together, theses two properties imply the unimodality of $W_{\mathsf{L}}(x)$ for $x \in \mathbb{R}_+$. Note that for $x = m + a$, where $m \in \mathbb{N}_0$ and $a \in [0,1)$, we have

$$W_{\mathsf{L}}(x) = \lambda_{\mathsf{L}} \cdot \frac{1}{\sum_{k=0}^{m} \lambda^k + \frac{\lambda^m(\lambda_{\mathsf{H}} + a\lambda_{\mathsf{L}})}{1-\lambda_{\mathsf{H}}}} \cdot \left(\sum_{k=0}^{m-1} \lambda^k u_{\mathsf{L}}(k) + a\lambda^m u_{\mathsf{L}}(m) \right).$$

Since this is of the form $\frac{\alpha+\beta a}{\gamma+\delta a}$, where $\alpha, \beta, \gamma, \delta$ are independent of a, we obtain that $W_{\mathsf{L}}(m+a)$ is monotone in $a \in [0,1)$. Thus, we conclude that $W_{\mathsf{L}}(x)$ is monotone between consecutive integers. It is straightforward to verify that $W_{\mathsf{L}}(x)$ is continuous, and hence the maximum of $W_{\mathsf{L}}(x)$ is attained at an integer.

For $m \in \mathbb{N}_0$, we have

$$W_{\mathsf{L}}(m) = \lambda_{\mathsf{L}} \cdot \frac{1}{\sum_{k=0}^{m} \lambda^k + \frac{\lambda^m \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}}} \cdot \sum_{k=0}^{m-1} \lambda^k u_{\mathsf{L}}(k) = \lambda_{\mathsf{L}} \cdot \Gamma(m) \cdot \Lambda(m) = \lambda_{\mathsf{L}} \Phi(m),$$

where $\Gamma(m) = \frac{1}{\sum_{k=0}^{m} \lambda^k + \frac{\lambda^m \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}}}$, $\Lambda(m) = \sum_{k=0}^{m-1} \lambda^k u_{\mathsf{L}}(k)$, and $\Phi(m) = \Gamma(m)\Lambda(m)$.

In the following, we show Φ is unimodal by establishing that if Φ decreases at some integer $m \in \mathbb{N}_0$, then it decreases at all integers $k \geq m$. Towards that goal, for any function $f : \mathbb{N}_0 \to \mathbb{R}$, let $\Delta f(m) \triangleq f(m) - f(m-1)$ denote the finite difference at m. Then, we have

$$\begin{aligned}
\Delta\Gamma(m) &= \frac{1}{\sum_{k=0}^{m} \lambda^k + \frac{\lambda^m \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}}} - \frac{1}{\sum_{k=0}^{m-1} \lambda^k + \frac{\lambda^{m-1} \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}}} = -\lambda^{m-1} \left(\frac{\lambda - \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}} \right) \Gamma(m)\Gamma(m-1), \\
\Delta\Lambda(m) &= \lambda^{m-1} u_{\mathsf{L}}(m-1),
\end{aligned}$$

and hence,

$$\begin{aligned}
\Delta\Phi(m) &= \Lambda(m)\Delta\Gamma(m) + \Gamma(m-1)\Delta\Lambda(m) \\
&= -\lambda^{m-1} \left(\frac{\lambda - \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}} \right) \Gamma(m)\Gamma(m-1)\Lambda(m) + \lambda^{m-1} u_{\mathsf{L}}(m-1)\Gamma(m-1) \\
&= \lambda^{m-1}\Gamma(m-1) \left(u_{\mathsf{L}}(m-1) - \left(\frac{\lambda - \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}} \right) \Phi(m) \right).
\end{aligned} \tag{EC.1}$$

Substituting $\Phi(m) = \Phi(m-1) + \Delta\Phi(m)$ into (EC.1) and after some algebra, we obtain for all $k \in \mathbb{N}_0$,

$$\left(1 + \left(\frac{\lambda - \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}} \right) \lambda^{k-1}\Gamma(k-1) \right) \Delta\Phi(k) = \lambda^{k-1}\Gamma(k-1) \left(u_{\mathsf{L}}(k-1) - \left(\frac{\lambda - \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}} \right) \Phi(k-1) \right). \tag{EC.2}$$

Now, suppose $\Delta\Phi(m) = \Phi(m) - \Phi(m-1) \leq 0$ for some $m \geq 1$. Since $\Gamma(m-1)$ is positive, from (EC.1) we obtain $u_{\mathsf{L}}(m-1) \leq \left(\frac{\lambda - \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}} \right) \Phi(m)$. Using the expression (EC.2) with $k = m+1$, we get

$$\begin{aligned}
\left(1 + \left(\frac{\lambda - \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}} \right) \lambda^m \Gamma(m) \right) \Delta\Phi(m+1) &= \lambda^m \Gamma(m) \left(u_{\mathsf{L}}(m) - \left(\frac{\lambda - \lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}} \right) \Phi(m) \right) \\
&\leq \lambda^m \Gamma(m) \left(u_{\mathsf{L}}(m) - u_{\mathsf{L}}(m-1) \right) \\
&\leq 0,
\end{aligned}$$

where we have used $u_{\mathsf{L}}(m-1) \leq \left(\frac{\lambda-\lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}}\right)\Phi(m)$ in the first inequality, and the fact that $u_{\mathsf{L}}(\cdot)$ is decreasing in the second inequality. This in turn implies that $\Delta\Phi(m+1) \leq 0$.

Thus, by induction, we obtain that if $\Delta\Phi(m) \leq 0$ then $\Delta\Phi(m+k) \leq 0$ for all $k \geq 0$. This proves the unimodality of Φ and hence that of $W_{\mathsf{L}} = \lambda_{\mathsf{L}}\Phi$. Finally, note that $W_{\mathsf{L}}(1) - W_{\mathsf{L}}(0) = \lambda_{\mathsf{L}}(1-\lambda_{\mathsf{H}})\frac{u_{\mathsf{L}}(0)}{1-\lambda_{\mathsf{H}}+\lambda} > 0$. Thus, $W_{\mathsf{L}}(m)$ initially increases up to a maximum, and then decreases subsequently.

For the third statement, note that for $x = m + a$, where $m \in \mathbb{N}_0$ and $a \in [0,1)$, we have

$$\begin{aligned}
W(x,\theta) &= \theta W_{\mathsf{L}}(x) + (1-\theta)W_{\mathsf{H}}(x) \\
&= \frac{1}{\sum_{k=0}^{m}\lambda^k + \frac{\lambda^m(\lambda_{\mathsf{H}}+a\lambda_{\mathsf{L}})}{1-\lambda_{\mathsf{H}}}}\left(\theta\lambda_{\mathsf{L}}\left(\sum_{k=0}^{m-1}\lambda^k u_{\mathsf{L}}(k) + a\lambda^m u_{\mathsf{L}}(m)\right)\right. \\
&\qquad \left. +(1-\theta)\left(\lambda_{\mathsf{H}}\sum_{k=0}^{m}\lambda^k u_{\mathsf{H}}(k) + (a\lambda_{\mathsf{L}}+\lambda_{\mathsf{H}})\sum_{k>m}\lambda^m\lambda_{\mathsf{H}}^{k-1-m}u_{\mathsf{H}}(k)\right)\right).
\end{aligned}$$

Again, this is of the form $\frac{\alpha+\beta a}{\gamma+\delta a}$, where $\alpha, \beta, \gamma, \delta$ are independent of $a \in [0,1)$. Thus, we obtain that $W(m+a,\theta)$ is monotone in $a \in [0,1)$, and hence $W(x,\theta)$ is monotone between consecutive integers. Since $W(x,\theta)$ is continuous in x, the maximum of $W(x,\theta)$ is attained at an integer. □

Lemma EC.4 **(Properties of the** (LEAVE) **function)**. *For $x \in \mathbb{R}_+$, the function $L(x)$ is strictly decreasing as long as it is non-negative, subsequent to which it stays negative. Formally, we have $L(x) \leq \max\{\inf_{0\leq u\leq x} L(u), 0\}$.*

Proof of Lemma EC.4. Consider a threshold policy $x = m + a$, where $m \in \mathbb{N}_0$ and $a \in [0,1)$. We have

$$\begin{aligned}
L(x) &= \sum_{k=0}^{\infty}(\lambda\pi_k - \pi_{k+1})u_{\mathsf{L}}(k) \\
&= \lambda_{\mathsf{L}}\pi_m(1-a)u_{\mathsf{L}}(m) + \lambda_{\mathsf{L}}\sum_{k=1}^{\infty}\pi_{m+k}u_{\mathsf{L}}(m+k) \\
&= \lambda_{\mathsf{L}} \cdot \frac{\lambda^m(1-a)u_{\mathsf{L}}(m) + \lambda^m(\lambda_{\mathsf{H}}+\lambda_{\mathsf{L}}a)\sum_{k=1}^{\infty}\lambda_{\mathsf{H}}^{k-1}u_{\mathsf{L}}(m+k)}{\sum_{k=0}^{m}\lambda^k + \frac{\lambda^m(\lambda_{\mathsf{H}}+a\lambda_{\mathsf{L}})}{1-\lambda_{\mathsf{H}}}}.
\end{aligned}$$

Since this is a ratio of two linear functions of a, we obtain that it is monotone in a, and hence, it suffices to analyze $L(x)$ as a function over integers. After some algebra, we have

$$\begin{aligned}
\frac{1}{\lambda_{\mathsf{L}}}L(m) &= \frac{\lambda^m\sum_{k=0}^{\infty}\lambda_{\mathsf{H}}^k u_{\mathsf{L}}(m+k)}{\sum_{k=0}^{m}\lambda^k + \frac{\lambda^m\lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}}} \\
&= \frac{\lambda^m\sum_{k=0}^{\infty}\lambda_{\mathsf{H}}^k}{\sum_{k=0}^{m}\lambda^k + \frac{\lambda^m\lambda_{\mathsf{H}}}{1-\lambda_{\mathsf{H}}}} \cdot \frac{\sum_{k=0}^{\infty}\lambda_{\mathsf{H}}^k u_{\mathsf{L}}(m+k)}{\sum_{k=0}^{\infty}\lambda_{\mathsf{H}}^k}.
\end{aligned}$$

Now, both factors on the right-hand side are strictly decreasing in m. Further, the first factor is positive. If $L(m)$ is non-negative, then the second factor is non-negative, and hence $L(m+1) -$

$L(m) < 0$. On the other hand, if $L(m) < 0$, then the second factor is negative, and since it is decreasing, we obtain $L(m+1) < 0$ as well. Thus, we conclude that $L(x)$ is strictly decreasing as long as it is non-negative, subsequent to which it stays negative. Formally, we have $L(x) \leq \max\{\inf_{0 \leq u \leq x} L(y), 0\}$. $\square$

EC.3. Proofs from Section 4

Proof of Proposition 1. First note that $0 < W_{\mathsf{L}}(\mathsf{fi}) \leq W_{\mathsf{L}}(\mathsf{sm})$ simply follows from the observation that fi is a feasible signaling mechanism. Thus its welfare is a lower bound on that achieved by the optimal signaling mechanism sm.

Next, we prove $W_{\mathsf{L}}(\mathsf{fi}) \geq \beta_{\mathsf{fi}} W_{\mathsf{L}}(\mathsf{sm})$. The proof consists of two steps. In the first step, we show $x_{\mathsf{sm}} \geq m_{\mathsf{fi}} - 1$, where $x_{\mathsf{sm}} \in \mathbb{R}_+$ is the threshold of the sm mechanism. We prove this by showing that the second obedience constraint, (LEAVE), will not be satisfied if the threshold is below $m_{\mathsf{fi}} - 1$. More precisely, let π^{sm} denote the steady-state distribution corresponding to sm mechanism, and let $x_{\mathsf{sm}} = m + a$ where $m \in \mathbb{N}_0$ and $a \in [0, 1)$. Then

$$L(\pi^{\mathsf{sm}}) = \begin{cases} \lambda_{\mathsf{L}} \pi_m (1-a) \cdot u_{\mathsf{L}}(m) + \lambda_{\mathsf{L}} \pi_{m+1} \cdot u_{\mathsf{L}}(m+1), & \text{if } a > 0; \\ \lambda_{\mathsf{L}} \pi_m \cdot u_{\mathsf{L}}(m), & \text{if } a = 0. \end{cases}$$

Here, the first case follows from the fact that, under the optimal signaling mechanism sm, a user is asked to leave with probability $1 - a$ if the queue length equals m, which occurs with probability π_m, and asked to leave with probability 1 if the queue-length equals $m + 1$, which occurs with probability π_{m+1}. The second case follows analogously.

Since $\pi^{\mathsf{sm}} \in \Pi_{\mathsf{SM}}$, we have $L(\pi^{\mathsf{sm}}) \leq 0$. This condition, along with the fact that $u_{\mathsf{L}}(\cdot)$ is strictly decreasing, forces $u_{\mathsf{L}}(m+1) < 0$ if $a > 0$, and $u_{\mathsf{L}}(m) \leq 0$ and $u_{\mathsf{L}}(m+1) < 0$ if $a = 0$. In both cases, we have $m + 1 \geq m_{\mathsf{fi}}$, and hence $x_{\mathsf{sm}} = m + a \geq m_{\mathsf{fi}} - 1 + a \geq m_{\mathsf{fi}} - 1$. Further, from Theorem 1, we have $x_{\mathsf{sm}} \leq m_{\mathsf{fi}}$. Putting these two together, we obtain

$$x_{\mathsf{sm}} \in [m_{\mathsf{fi}} - 1, m_{\mathsf{fi}}].$$

Since $W_{\mathsf{L}}(x)$ is monotone between integers (as established in Lemma EC.3), we thus obtain that $W_{\mathsf{L}}(\mathsf{fi}) = W_{\mathsf{L}}(\mathsf{sm})$ if and only if $W_{\mathsf{L}}(m_{\mathsf{fi}}) \geq W_{\mathsf{L}}(m_{\mathsf{fi}} - 1)$. Furthermore, we have $W_{\mathsf{L}}(\mathsf{sm}) \leq \max\{W_{\mathsf{L}}(m_{\mathsf{fi}} - 1), W_{\mathsf{L}}(m_{\mathsf{fi}})\}$. Now,

$$\begin{aligned} W_{\mathsf{L}}(m_{\mathsf{fi}} - 1) &= \frac{\sum_{n=0}^{m_{\mathsf{fi}}-2} \lambda_{\mathsf{L}}^n u_{\mathsf{L}}(n)}{\sum_{n=0}^{m_{\mathsf{fi}}-1} \lambda_{\mathsf{L}}^n} \\ &\leq \frac{\sum_{n=0}^{m_{\mathsf{fi}}-1} \lambda_{\mathsf{L}}^n u_{\mathsf{L}}(n)}{\sum_{n=0}^{m_{\mathsf{fi}}-1} \lambda_{\mathsf{L}}^n} = \left(\frac{\sum_{n=0}^{m_{\mathsf{fi}}} \lambda_{\mathsf{L}}^n}{\sum_{n=0}^{m_{\mathsf{fi}}-1} \lambda_{\mathsf{L}}^n} \right) \cdot \frac{\sum_{n=0}^{m_{\mathsf{fi}}-1} \lambda_{\mathsf{L}}^n u_{\mathsf{L}}(n)}{\sum_{n=0}^{m_{\mathsf{fi}}} \lambda_{\mathsf{L}}^n} = \frac{1}{\beta_{\mathsf{fi}}} \cdot W_{\mathsf{L}}(\mathsf{fi}). \end{aligned}$$

Here, in the inequality follows from the fact that $u_{\mathsf{L}}(m_{\mathsf{fi}} - 1) \geq 0$, and the first and the second equalities follow from the definition of a threshold mechanism. In the final equality, we have used the definition of β_{fi}. Thus, taken together, we obtain $W_{\mathsf{L}}(\mathsf{fi}) \geq \beta_{\mathsf{fi}} W_{\mathsf{L}}(\mathsf{sm})$. The statement of the proposition follows after noting that $\beta_{\mathsf{fi}} \geq 1 - \frac{1}{m_{\mathsf{fi}}+1}$ for all $\lambda_{\mathsf{L}} \leq 1$. $\square$

Proof of Proposition 2. Recall that p^{ni} denotes the probability with which a L-type user joins under the no-information mechanism. Since $\lambda_{\mathsf{H}} = 0$, we note that $p^{\mathsf{ni}} > 0$, as a L-type user will find it optimal to join the queue if no other such user does so. Consequently, we consider the cases $p^{\mathsf{ni}} = 1$ and $p^{\mathsf{ni}} \in (0,1)$. Suppose $p^{\mathsf{ni}} = 1$. Then we have

$$
\begin{aligned}
W_{\mathsf{L}}(\mathsf{ni}) &= \sum_{n \in \mathbb{N}_0} (1-\lambda_{\mathsf{L}}) \lambda_{\mathsf{L}}^n u_{\mathsf{L}}(n) \\
&\leq \sum_{n < m_{\mathsf{fi}}} (1-\lambda_{\mathsf{L}}) \lambda_{\mathsf{L}}^n u_{\mathsf{L}}(n) + u_{\mathsf{L}}(m_{\mathsf{fi}}) \lambda_{\mathsf{L}}^{m_{\mathsf{fi}}} \\
&= \left(\sum_{n=0}^{m_{\mathsf{fi}}} (1-\lambda_{\mathsf{L}}) \lambda_{\mathsf{L}}^n \right) \cdot \frac{\sum_{n=0}^{m_{\mathsf{fi}}-1} \lambda_{\mathsf{L}}^n u_{\mathsf{L}}(n)}{\sum_{n=0}^{m_{\mathsf{fi}}} \lambda_{\mathsf{L}}^n} + u_{\mathsf{L}}(m_{\mathsf{fi}}) \lambda_{\mathsf{L}}^{m_{\mathsf{fi}}} \\
&= \left(1 - \lambda_{\mathsf{L}}^{m_{\mathsf{fi}}+1}\right) \cdot W_{\mathsf{L}}(\mathsf{fi}) + u_{\mathsf{L}}(m_{\mathsf{fi}}) \lambda_{\mathsf{L}}^{m_{\mathsf{fi}}} \\
&< \left(1 - \lambda_{\mathsf{L}}^{m_{\mathsf{fi}}+1}\right).
\end{aligned}
$$

Here, we use the fact that $u_{\mathsf{L}}(k)$ is decreasing k in the first inequality, and the final inequality follows from $u_{\mathsf{L}}(m_{\mathsf{fi}}) < 0$. On the other hand, if $p^{\mathsf{ni}} \in (0,1)$, we have $W_{\mathsf{L}}(\mathsf{ni}) = 0 < (1 - \lambda_{\mathsf{L}}^{m_{\mathsf{fi}}+1}) W_{\mathsf{L}}(\mathsf{fi})$. $\square$

Proof of Proposition 3. Recall from Theorem 4 that the full-information mechanism is Pareto-efficient if and only if $W_{\mathsf{L}}(m_{\mathsf{fi}}) - W_{\mathsf{L}}(m_{\mathsf{fi}} - 1) > 0$. By a little algebra, this condition can be shown to be equivalent to

$$
f(\lambda_{\mathsf{L}}, \lambda_{\mathsf{H}}) \triangleq \lambda_{\mathsf{L}} u_{\mathsf{L}}(0) - \lambda_{\mathsf{L}} \sum_{k=1}^{m_{\mathsf{fi}}-1} (\lambda_{\mathsf{H}} + \lambda_{\mathsf{L}})^k (u_{\mathsf{L}}(k-1) - u_{\mathsf{L}}(k)) - (1-\lambda_{\mathsf{H}}) u_{\mathsf{L}}(m_{\mathsf{fi}} - 1) < 0.
$$

It is straightforward to verify that $f(0, \lambda_{\mathsf{H}}) < 0$, $f(1-\lambda_{\mathsf{H}}, \lambda_{\mathsf{H}}) = 0$, $\partial_{\mathsf{L}} f(0, \lambda_{\mathsf{H}}) = u_{\mathsf{L}}(0) - \sum_{k=1}^{m_{\mathsf{fi}}-1} \lambda_{\mathsf{H}}^k (u_{\mathsf{L}}(k-1) - u_{\mathsf{L}}(k)) \geq u_{\mathsf{L}}(m_{\mathsf{fi}} - 1) > 0$, and $\partial_{\mathsf{L}}^2 f < 0$ for $\lambda_{\mathsf{L}} \in [0, 1-\lambda_{\mathsf{H}}]$, where ∂_{L} denotes the partial derivative with respect to λ_{L}. These facts imply that for any fixed $\lambda_{\mathsf{H}} \in (0,1)$, the function $f(\cdot, \lambda_{\mathsf{H}})$ has a root $\bar{\Lambda}_{\mathsf{L}}(\lambda_{\mathsf{H}}) \in (0, 1-\lambda_{\mathsf{H}}]$ satisfying $f(\lambda_{\mathsf{L}}, \lambda_{\mathsf{H}}) < 0$ for $\lambda_{\mathsf{L}} < \bar{\Lambda}_{\mathsf{L}}(\lambda_{\mathsf{H}})$ and $f(\lambda_{\mathsf{L}}, \lambda_{\mathsf{H}}) > 0$ for $\bar{\Lambda}_{\mathsf{L}}(\lambda_{\mathsf{H}}) < \lambda_{\mathsf{L}} < 1 - \lambda_{\mathsf{H}}$. Thus, we obtain that the full-information mechanism is Pareto-efficient if and only if $\lambda_{\mathsf{L}} < \bar{\Lambda}_{\mathsf{L}}(\lambda_{\mathsf{H}})$. Finally, the definition of f, along with some straightforward algebra, yields the following lower-bound:

$$
\bar{\Lambda}_{\mathsf{L}}(\lambda_{\mathsf{H}}) \geq \frac{u_{\mathsf{L}}(m_{\mathsf{fi}} - 1)}{u_{\mathsf{L}}(0) - \sum_{k=1}^{m_{\mathsf{fi}}-1} \lambda_{\mathsf{H}}^k (u_{\mathsf{L}}(k-1) - u_{\mathsf{L}}(k))} \cdot (1 - \lambda_{\mathsf{H}}).
$$

In order to prove the second part, we note that the proof of the first part implies that if $\lambda_{\mathsf{L}} < \bar{\Lambda}_{\mathsf{L}}(\lambda_{\mathsf{H}})$, we have $J(m_{\mathsf{fi}} - 1) = W_{\mathsf{L}}(m_{\mathsf{fi}} - 1) \geq W_{\mathsf{L}}(m_{\mathsf{fi}}) > 0$. Furthermore, the assumption $L(m_{\mathsf{fi}} - 1) \leq 0$ implies that the threshold mechanism with threshold of $m_{\mathsf{fi}} - 1$ is an obedient mechanism. This further implies that the efficient signaling mechanism has a threshold of at most $m_{\mathsf{fi}} - 1$. To see why, suppose $x_{\mathsf{sm}} > m_{\mathsf{fi}} - 1$. As established in Lemma EC.3, $W_{\mathsf{H}}(m_{\mathsf{fi}} - 1) > W_{\mathsf{H}}(x_{\mathsf{sm}})$ and $W_{\mathsf{L}}(m_{\mathsf{fi}} - 1) \geq W_{\mathsf{L}}(x_{\mathsf{sm}})$ implying that the threshold mechanism with threshold $m_{\mathsf{fi}} - 1$ Pareto dominates the one with threshold x_{sm} which is a contradiction.

In light of the above observations, we have $W_i(\mathsf{sm}) \geq W_i(m_{\mathsf{fi}} - 1)$, for $i \in \{\mathsf{L}, \mathsf{H}\}$. Thus in the following, we establish a multiplicative gap between $W_i(m_{\mathsf{fi}} - 1)$ and $W_i(m_{\mathsf{fi}})$ for each $i \in \{\mathsf{L}, \mathsf{H}\}$.

We start with the L-type users. Observe that, for any threshold mechanism with threshold $m \leq m_{\mathsf{fi}}$, we have

$$W_{\mathsf{L}}(m) = \frac{\lambda_{\mathsf{L}}}{Z_m} \sum_{k=0}^{m-1} \lambda^k u_{\mathsf{L}}(k),$$

where $Z_m = \sum_{k=0}^{m-1} \lambda^k + \sum_{k=m}^{\infty} \lambda^m \lambda_{\mathsf{H}}^{k-m} = \frac{1-\lambda^m}{1-\lambda} + \frac{\lambda^m}{1-\lambda_{\mathsf{H}}}$. Thus, we obtain

$$\begin{aligned}
W_{\mathsf{L}}(m-1) - W_{\mathsf{L}}(m) &= \frac{\lambda_{\mathsf{L}}}{Z_{m-1}} \sum_{k=0}^{m-2} \lambda^k u_{\mathsf{L}}(k) - \frac{\lambda_{\mathsf{L}}}{Z_m} \sum_{k=0}^{m-1} \lambda^k u_{\mathsf{L}}(k) \\
&= \frac{\lambda_{\mathsf{L}}}{Z_{m-1}} \sum_{k=0}^{m-2} \lambda^k u_{\mathsf{L}}(k) - \frac{\lambda_{\mathsf{L}}}{Z_{m-1}} \sum_{k=0}^{m-1} \lambda^k u_{\mathsf{L}}(k) + \frac{\lambda_{\mathsf{L}}}{Z_{m-1}} \sum_{k=0}^{m-1} \lambda^k u_{\mathsf{L}}(k) - \frac{\lambda_{\mathsf{L}}}{Z_m} \sum_{k=0}^{m-1} \lambda^k u_{\mathsf{L}}(k) \\
&= \frac{\lambda_{\mathsf{L}} \lambda^{m-1}}{(1-\lambda_{\mathsf{H}}) Z_{m-1}} \left(W_{\mathsf{L}}(m) - (1-\lambda_{\mathsf{H}}) u_{\mathsf{L}}(m-1) \right) \\
&= \left(\frac{(1-\lambda)\lambda_{\mathsf{L}} \lambda^{m-1}}{1-\lambda_{\mathsf{H}} - \lambda_{\mathsf{L}} \lambda^{m-1}} \right) \left(W_{\mathsf{L}}(m) - (1-\lambda_{\mathsf{H}}) u_{\mathsf{L}}(m-1) \right).
\end{aligned}$$

Equivalently, we have

$$W_{\mathsf{L}}(m-1) = \left(1 + \frac{(1-\lambda)\lambda_{\mathsf{L}} \lambda^{m-1}}{1-\lambda_{\mathsf{H}} - \lambda_{\mathsf{L}} \lambda^{m-1}} \right) W_{\mathsf{L}}(m) - \left(\frac{(1-\lambda_{\mathsf{H}})(1-\lambda)\lambda_{\mathsf{L}} \lambda^{m-1}}{1-\lambda_{\mathsf{H}} - \lambda_{\mathsf{L}} \lambda^{m-1}} \right) u_{\mathsf{L}}(m-1).$$

Letting $m = m_{\mathsf{fi}}$ and using the assumption that $u_{\mathsf{L}}(m_{\mathsf{fi}} - 1) \leq W_{\mathsf{L}}(\mathsf{fi})$, we get

$$W_{\mathsf{L}}(m_{\mathsf{fi}} - 1) \geq \left(1 + \frac{\lambda_{\mathsf{H}}(1-\lambda)\lambda_{\mathsf{L}} \lambda^{m_{\mathsf{fi}}-1}}{1-\lambda_{\mathsf{H}} - \lambda_{\mathsf{L}} \lambda^{m_{\mathsf{fi}}-1}} \right) W_{\mathsf{L}}(m_{\mathsf{fi}}) = \beta_{\mathsf{L},\mathsf{sm}} \cdot W_{\mathsf{L}}(m_{\mathsf{fi}}).$$

Note that by definition, $1 - \lambda_{\mathsf{H}} - \lambda_{\mathsf{L}} \lambda^{m_{\mathsf{fi}}-1} > 1 - \lambda_{\mathsf{H}} - \lambda_{\mathsf{L}} > 0$ for $\lambda < 1$, implying that $\beta_{\mathsf{L},\mathsf{sm}} > 1$.

Next, we proceed to the H type. Similar to $W_{\mathsf{L}}(m)$, we have

$$W_{\mathsf{H}}(m) = \frac{\lambda_{\mathsf{H}}}{Z_m} \left(\sum_{k=0}^{m-1} \lambda^k u_{\mathsf{H}}(k) + \lambda^m \sum_{k=m}^{\infty} \lambda_{\mathsf{H}}^{k-m} u_{\mathsf{H}}(k) \right) \triangleq \lambda_{\mathsf{H}} \frac{F_m}{Z_m}.$$

Thus, we get

$$W_{\mathsf{H}}(m-1) - W_{\mathsf{H}}(m) = \frac{\lambda_{\mathsf{H}}}{Z_{m-1}} \left(F_{m-1} - F_m \right) + \frac{\lambda_{\mathsf{L}} \lambda^{m-1}}{(1-\lambda_{\mathsf{H}}) Z_{m-1}} W_{\mathsf{H}}(m).$$

Furthermore,

$$F_{m-1} - F_m = -\lambda_{\mathsf{L}} \lambda^{m-1} \sum_{k=m}^{\infty} \lambda_{\mathsf{H}}^{k-m} u_{\mathsf{H}}(k).$$

Thus,

$$
\begin{aligned}
W_{\mathsf{H}}(m-1)-W_{\mathsf{H}}(m) &= \frac{\lambda_{\mathsf{L}}\lambda^{m-1}}{(1-\lambda_{\mathsf{H}})Z_{m-1}}\left(W_{\mathsf{H}}(m)-\lambda_{\mathsf{H}}\left(\sum_{k=m}^{\infty}(1-\lambda_{\mathsf{H}})\lambda_{\mathsf{H}}^{k-m}u_{\mathsf{H}}(k)\right)\right)\\
&= \left(\frac{(1-\lambda)\lambda_{\mathsf{L}}\lambda^{m-1}}{1-\lambda_{\mathsf{H}}-\lambda_{\mathsf{L}}\lambda^{m-1}}\right)\left(W_{\mathsf{H}}(m)-\lambda_{\mathsf{H}}\left(\sum_{k=m}^{\infty}(1-\lambda_{\mathsf{H}})\lambda_{\mathsf{H}}^{k-m}u_{\mathsf{H}}(k)\right)\right).
\end{aligned}
$$

This implies that

$$
\begin{aligned}
W_{\mathsf{H}}(m-1) &= \left(1+\frac{(1-\lambda)\lambda_{\mathsf{L}}\lambda^{m-1}}{1-\lambda_{\mathsf{H}}-\lambda_{\mathsf{L}}\lambda^{m-1}}\right)W_{\mathsf{H}}(m)-\left(\frac{\lambda_{\mathsf{H}}(1-\lambda)\lambda_{\mathsf{L}}\lambda^{m-1}}{1-\lambda_{\mathsf{H}}-\lambda_{\mathsf{L}}\lambda^{m-1}}\right)\left(\sum_{k=m}^{\infty}(1-\lambda_{\mathsf{H}})\lambda_{\mathsf{H}}^{k-m}u_{\mathsf{H}}(k)\right)\\
&\ge \left(1+\frac{(1-\lambda)\lambda_{\mathsf{L}}\lambda^{m-1}}{1-\lambda_{\mathsf{H}}-\lambda_{\mathsf{L}}\lambda^{m-1}}\right)W_{\mathsf{H}}(m)-\left(\frac{\lambda_{\mathsf{H}}(1-\lambda)\lambda_{\mathsf{L}}\lambda^{m-1}}{1-\lambda_{\mathsf{H}}-\lambda_{\mathsf{L}}\lambda^{m-1}}\right)u_{\mathsf{H}}(m).
\end{aligned}
$$

Again, letting $m=m_{\mathsf{fi}}$ and using the assumption that $u_{\mathsf{H}}(m_{\mathsf{fi}})\le W_{\mathsf{H}}(\mathsf{fi})$, we get

$$
W_{\mathsf{H}}(m_{\mathsf{fi}}-1)\ge\left(1+\frac{(1-\lambda_{\mathsf{H}})(1-\lambda)\lambda_{\mathsf{L}}\lambda^{m_{\mathsf{fi}}-1}}{1-\lambda_{\mathsf{H}}-\lambda_{\mathsf{L}}\lambda^{m_{\mathsf{fi}}-1}}\right)W_{\mathsf{H}}(m_{\mathsf{fi}})=\beta_{\mathsf{H,sm}}\cdot W_{\mathsf{H}}(m_{\mathsf{fi}}).
$$

Before proving the third part of the proposition, we note that using Assumption 1, along with a stochastic dominance argument, we obtain that function $g(x)\triangleq\sum_{k\in\mathbb{N}_0}(1-x)x^k u_{\mathsf{L}}(k)$ is strictly decreasing in $x\in[0,1)$. Furthermore, $g(0)=u_{\mathsf{L}}(0)>0$ and $\lim_{x\to 1^-}g(x)<0$. Thus, there exists a unique $\bar{\Lambda}_{\mathsf{H}}\in(0,1)$ such that $g(\bar{\Lambda}_{\mathsf{H}})=0$. To prove the second part of the proposition, we show that if $\lambda_{\mathsf{H}}\ge\bar{\Lambda}_{\mathsf{H}}$, then no L-type user joins under the no-information mechanism, i.e., $p^{\mathsf{ni}}=0$. The result then follows from Theorem 5. Thus, suppose $\lambda_{\mathsf{H}}\ge\bar{\Lambda}_{\mathsf{H}}$, and no (other) L-type user joins the queue under no-information mechanism. The steady-state distribution π of the queue is then that of an $M/M/1$ queue with arrival rate λ_{H}, and hence we have $\pi_n=(1-\lambda_{\mathsf{H}})\lambda_{\mathsf{H}}^n$ for $n\ge 0$. This implies that the expected utility (in steady-state) of a L-type user for joining is given by $\sum_{k\in\mathbb{N}_0}\pi_n u_{\mathsf{L}}(n)=\sum_{k\in\mathbb{N}_0}(1-\lambda_{\mathsf{H}})\lambda_{\mathsf{H}}^k u_{\mathsf{L}}(n)=g(\lambda_{\mathsf{H}})\le 0$. The inequality follows from the fact that $g(x)$ is decreasing in x and equals zero when $x=\bar{\Lambda}_{\mathsf{H}}$. Thus, we obtain that the optimal action for a L-type user is indeed not to join, and hence $p^{\mathsf{ni}}=0$. This completes the proof. $\square$

Proof of Theorem 4. Recall that under the full-information mechanism fi, the L-type users receive the "join" signal if and only if the queue-length is strictly less than the full-information threshold m_{fi}. Thus, conditional on receiving the "leave" signal, the queue-length is at least m_{fi}, and the expected utility of the L-type users for joining the queue is given by

$$
U_{\mathsf{L}}(0,\mathsf{join})\le u_{\mathsf{L}}(m_{\mathsf{fi}})<0,
$$

where we have used the definition of m_{fi} and the fact that $u_{\mathsf{L}}(k)$ is strictly decreasing in k. Together with the fact that the probability of receiving a "leave" signal is positive under the full-information mechanism, we obtain that $L(m_{\mathsf{fi}})<0$, and hence the (LEAVE) condition does not bind. Hence, from Theorem 3, we conclude that the full-information mechanism is Pareto-efficient within the class Π_{AP} if and only if it is so within the class Π_{SM}. Thus, we obtain the dichotomy in the theorem statement.

To show the final part of the theorem, suppose the full-information mechanism is Pareto-efficient within the class of admission policies Π_{AP}. Consider the admission policy with threshold $m_{\mathsf{fi}} - 1$. In Lemma EC.3 (see Appendix EC.2), we show that $W_{\mathsf{H}}(x)$ is strictly decreasing in the threshold x. Hence, we have $W_{\mathsf{H}}(m_{\mathsf{fi}} - 1) > W_{\mathsf{H}}(m_{\mathsf{fi}})$. Since the full-information mechanism is Pareto-efficient within Π_{AP}, this implies $W_{\mathsf{L}}(m_{\mathsf{fi}}) > W_{\mathsf{L}}(m_{\mathsf{fi}} - 1)$.

Conversely, suppose $W_{\mathsf{L}}(m_{\mathsf{fi}}) > W_{\mathsf{L}}(m_{\mathsf{fi}} - 1)$. In Lemma EC.3, we also show that $W_{\mathsf{L}}(x)$ is unimodal, i.e, $W_{\mathsf{L}}(x)$ is increasing for small $x \in \mathbb{R}_+$ and decreasing otherwise. The unimodality then implies that for all $0 \le \hat{x} < m_{\mathsf{fi}}$, we have $W_{\mathsf{L}}(\hat{x}) < W_{\mathsf{L}}(m_{\mathsf{fi}})$. Thus, no admission policy with threshold $\hat{x} < m_{\mathsf{fi}}$ Pareto dominates the full-information mechanism. Since any admission policy that is not Pareto-efficient is dominated by some threshold policy with threshold less than or equal to m_{fi}, we obtain that the full-information mechanism, with threshold m_{fi}, is Pareto-efficient within the class Π_{AP} of admission policies. $\square$

Proof of Theorem 5. Recall that under the no-information mechanism, a L-type users joins the queue with a fixed probability $p^{\mathsf{ni}} \in [0,1]$ irrespective of the queue-length upon arrival.

For $p^{\mathsf{ni}} \in (0,1)$ it is straightforward to verify that the resulting steady-state distribution does not have a threshold structure, and hence by Theorem 1, the no-information mechanism is not Pareto-efficient. For $p^{\mathsf{ni}} = 1$, the resulting steady-state distribution has a threshold structure with threshold equal to infinity. In this case, Theorem 1 implies that the no-information mechanism is not Pareto-efficient. Thus, if $p^{\mathsf{ni}} \in (0,1]$, then the no-information mechanism is Pareto-dominated within the class Π_{SM} of signaling mechanisms.

Finally, suppose $p^{\mathsf{ni}} = 0$. Then, the steady-state distribution π^{ni} is given by $\pi^{\mathsf{ni}}_n = (1 - \lambda_{\mathsf{H}})\lambda_{\mathsf{H}}^n$ for $n \ge 0$. Now, consider any other admission policy $\hat{\pi} \in \Pi_{\mathsf{AP}}$, where at least some fraction of L-type users are admitted into the queue. Using a coupling argument, it is straightforward to show that $\hat{\pi}$ stochastically dominates π^{ni}. Since $u_{\mathsf{H}}(n)$ is strictly decreasing in n, this further implies that $W_{\mathsf{H}}(\hat{\pi}) < W_{\mathsf{H}}(\pi^{\mathsf{ni}})$. Hence, it follows that the no-information mechanism ni is Pareto-efficient within the class Π_{AP} of admission policies. $\square$

EC.4. Proofs from Section 5

Proof of Theorem 6. First, suppose $\lambda_{\mathsf{H}} \in [\bar{\Lambda}_{\mathsf{H}}, 1]$, and fix a $\theta \in [0,1]$. From Theorem 5, we obtain that the no-information mechanism ni is Pareto-efficient, and furthermore, under ni, all L-type users choose their outside option. Consider the admission policy $\mathsf{ap}(\theta)$. If $\mathsf{ap}(\theta)$ makes some L-type users join the queue, then the welfare of H-type users can only be lower than that in ni:

$W_{\mathsf{H}}(\mathsf{ni}) \geq W_{\mathsf{H}}(\pi)$. Thus, for $\mathsf{ap}(\theta)$ to be Pareto-efficient, we must have $W_{\mathsf{L}}(\mathsf{ap}(\theta)) > W_{\mathsf{L}}(\mathsf{ni}) = 0$. Thus, we have $J(\mathsf{ap}(\theta)) = W_{\mathsf{L}}(\mathsf{ap}(\theta)) > 0$, and hence the obedience constraint (JOIN) holds. Furthermore, we have

$$J(\mathsf{ap}(\theta)) + L(\mathsf{ap}(\theta)) = \lambda_{\mathsf{L}} \sum_{n \in \mathbb{N}_0} \pi_n(\mathsf{ap}(\theta)) u_{\mathsf{L}}(n) \leq 0,$$

where $\pi(\mathsf{ap}(\theta))$ denotes the steady-state distribution under $\mathsf{ap}(\theta)$. This is because $\pi(\mathsf{ap}(\theta))$ stochastically dominates the steady-state distribution under ni, and hence the right-hand side expression is less than $\lambda_{\mathsf{L}} \sum_{n \in \mathbb{N}_0} (1 - \lambda_{\mathsf{H}}) \lambda_{\mathsf{H}}^n u_{\mathsf{L}}(n)$, which is non-positive as $\lambda_{\mathsf{H}} \geq \bar{\Lambda}_{\mathsf{H}}$. Since $J(\mathsf{ap}(\theta)) \geq 0$, this implies that $L(\mathsf{ap}(\theta)) \leq 0$, and hence $\mathsf{ap}(\theta)$ also satisfies the obedience constraint LEAVE. Taken together, we obtain that $\mathsf{ap}(\theta) \in \Pi_{\mathsf{SM}}$, and hence $\mathsf{ap}(\theta) = \mathsf{sm}(\theta)$.

Next, let $\lambda_{\mathsf{H}} < \bar{\Lambda}_{\mathsf{H}}$. Fix $\theta_1, \theta_2 \in [0,1]$ with $\theta_2 > \theta_1$, and let x_i denote the threshold of the Pareto-efficient admission policy $\mathsf{ap}(\theta_i)$. In the following, we first show that $x_1 \leq x_2$. By the definition of $W(\pi, \theta)$ and $\mathsf{ap}(\theta)$, we have

$$\begin{aligned}
\theta_1 W_{\mathsf{L}}(x_1) + (1-\theta_1) W_{\mathsf{H}}(x_1) &\geq \theta_1 W_{\mathsf{L}}(x_2) + (1-\theta_1) W_{\mathsf{H}}(x_2), \\
\theta_2 W_{\mathsf{L}}(x_2) + (1-\theta_2) W_{\mathsf{H}}(x_2) &\geq \theta_2 W_{\mathsf{L}}(x_1) + (1-\theta_2) W_{\mathsf{H}}(x_1).
\end{aligned}$$

After some algebra, we obtain

$$W_{\mathsf{L}}(x_2) - W_{\mathsf{L}}(x_1) \geq W_{\mathsf{H}}(x_2) - W_{\mathsf{H}}(x_1).$$

Now, if $x_1 > x_2$, then from Lemma EC.3 in Appendix EC.2, we obtain $W_{\mathsf{H}}(x_1) < W_{\mathsf{H}}(x_2)$. The preceding inequality would then imply $W_{\mathsf{L}}(x_1) < W_{\mathsf{L}}(x_2)$. However, this would imply that the admission policy $\mathsf{ap}(\theta_1)$ is Pareto-dominated by the policy $\mathsf{ap}(\theta_2)$, a contradiction. Thus, we obtain that $x_1 \leq x_2$.

Next, suppose the admission policy $\mathsf{ap}(\theta_1)$ satisfies the obedience constraints, and hence $L(x_1) \leq 0$. In Lemma EC.4 (stated and proven in Appendix EC.2), we establish that if $L(x) \leq 0$ then $L(u) \leq 0$ for all $u \geq x$. Since $x_1 \leq x_2$, Lemma EC.4 implies that $L(x_2) \leq 0$, and hence the (LEAVE) condition holds for $\mathsf{ap}(\theta_2)$. Further, by Theorem 2 we have $x_2 \leq m_{\mathsf{fi}}$, which implies $J(x_2) \geq 0$ and hence the (JOIN) condition holds for $\mathsf{ap}(\theta_2)$. Together, we obtain that $\mathsf{ap}(\theta_2)$ also satisfies the obedience constraints.

Thus, we conclude that if for some $\theta_1 \in [0,1]$ the admission policy $\mathsf{ap}(\theta_1)$ satisfies the obedience constraints, then so does the admission policy $\mathsf{ap}(\theta_2)$ for all $\theta_2 > \theta_1$. This implies the existence of (a smallest such) $\theta(\lambda_{\mathsf{L}}, \lambda_{\mathsf{H}}) \in [0,1]$ such that for all $\theta > \theta(\lambda_{\mathsf{L}}, \lambda_{\mathsf{H}})$ we have $\mathsf{sm}(\theta) = \mathsf{ap}(\theta)$.[18]

[18] Note that in this case, the threshold of the admission policy $\mathsf{ap}(\theta)$ (or equivalently the signaling mechanism $\mathsf{sm}(\theta)$) can be positive. For numerical examples, see Figure 4 and its related discussion in Section 6.

(Note that we allow the possibility that $\theta(\lambda_{\mathsf{L}},\lambda_{\mathsf{H}})=1$.) Further, we have $\theta(\lambda_{\mathsf{L}},\lambda_{\mathsf{H}})>0$, since for $\theta=0$, the admission policy $\mathsf{ap}(0)$ makes all L-type users take the outside option. However, the obedience condition (LEAVE) does not hold for $\mathsf{ap}(0)$ since $\lambda_{\mathsf{H}}<\bar{\Lambda}_{\mathsf{H}}$.

Finally, for $\theta<\theta(\lambda_{\mathsf{L}},\lambda_{\mathsf{H}})$, the admission policy $\mathsf{ap}(\theta)$ does not satisfy the obedience constraints, and hence $\mathsf{sm}(\theta)\neq\mathsf{ap}(\theta)$. Theorem 3 then implies that the (LEAVE) condition binds for all such θ, i.e., $L(\mathsf{sm}(\theta))=0$. In Lemma EC.4, we also prove that $L(x)$ is strictly decreasing as long as it is non-negative, and remains negative subsequently. Thus, there exists a unique threshold $\bar{x}\leq m_{\mathsf{fi}}$ (independent of θ) with $L(\bar{x})=0$. From this, we conclude that $\mathsf{sm}(\theta)$ is the threshold mechanism with threshold $\bar{x}$ for all $\theta<\theta(\lambda_{\mathsf{L}},\lambda_{\mathsf{H}})$. $\square$

EC.5. Proofs from Section 7.1

Proof of Proposition 4. *Proof of Part 1:* Let π denote the full-information mechanism. Suppose π is Pareto dominated by an admission policy $\tilde{\pi}$, i.e., $W_{\mathsf{H}}(\tilde{\pi})\geq W_{\mathsf{H}}(\pi)$ and $W_{\mathsf{L}}(\tilde{\pi})\geq W_{\mathsf{L}}(\pi)$, with at least one inequality strict. Since under the full-information mechanism π, none of the obedience constraints bind, we obtain that for small enough $\delta>0$, the admission policy $\pi_\delta=(1-\delta)\pi+\delta\tilde{\pi}$ satisfies all the obedience constraints, and hence can be implemented as a signaling mechanism. Using the linearity of the welfare functions, we conclude that π_δ Pareto dominates the full-information mechanism π.

Finally, using Lemma EC.5 stated later in this section, we obtain $W_{\mathsf{L}}(m_{\mathsf{L}},m_{\mathsf{H}}-1)>W_{\mathsf{L}}(m_{\mathsf{L}},m_{\mathsf{H}})$ and $W_{\mathsf{H}}(m_{\mathsf{L}}-1,m_{\mathsf{H}})\geq W_{\mathsf{H}}(m_{\mathsf{L}},m_{\mathsf{H}})$. Thus, if $W_{\mathsf{H}}(m_{\mathsf{L}},m_{\mathsf{H}}-1)\geq W_{\mathsf{H}}(m_{\mathsf{L}},m_{\mathsf{H}})$, then we obtain that the threshold mechanism $\mathsf{Th}(m_{\mathsf{L}},m_{\mathsf{H}}-1)$ Pareto-dominates the full information mechanism. On the other hand, if $W_{\mathsf{L}}(m_{\mathsf{L}}-1,m_{\mathsf{H}})\geq W_{\mathsf{L}}(m_{\mathsf{L}},m_{\mathsf{H}})$, then the threshold mechanism $\mathsf{Th}(m_{\mathsf{L}}-1,m_{\mathsf{H}})$ Pareto dominates the full information. In either case, there exists a threshold signaling mechanism that Pareto dominates the full-information mechanism, since for any $\delta>0$, the signaling mechanisms $(1-\delta)\mathsf{Th}(m,n-1)+\delta\mathsf{Th}(m,n)$ and $(1-\delta)\mathsf{Th}(m-1,n)+\delta\mathsf{Th}(m,n)$ both have a threshold structure.

Proof of Part 2: We begin by showing that the no-information mechanism is Pareto dominated in the class of admission policies. First, suppose under the no-information mechanism, the L-type users never join, while the H-type users join with some probability $p\in(0,1]$. In this case, the admission policy that never admits the L-type, and implements the admission rule that maximizes the H types' welfare Pareto dominates the no-information mechanism. Next, if the L-type users join with probability $p\in(0,1)$ under the no-information mechanism, then due to the assumption on the utilities, the H-type user always joins. Since $p\in(0,1)$, the welfare of the L-type in this case is zero. This implies that the admission policy that never admits L-type user and always admits the H-type user Pareto dominates the no-information mechanism. Finally, suppose both types join with probability 1 under the no-information mechanism. In this case, the admission policy that never admits any type above queue length m_{H} and always admits below this queue length achieves higher utility for both types, and hence Pareto dominates the no-information mechanism.

Next, suppose under the no-information mechanism, the L-type users join with positive probability. To show that the no-information mechanism is Pareto dominated by a signaling mechanism, we split the argument into two cases:

1. Suppose in equilibrium, both types join with probability 1. Consider the threshold mechanism $\mathsf{Th}(m_{\mathsf{H}}, m_{\mathsf{H}})$, i.e., the mechanism sends signal 2 up to queue length m_{H}, and sends signal 0 afterwards. From a straightforward argument, it follows that this mechanism is obedient, and achieves higher welfare for both types than the no-information mechanism.
2. Suppose in equilibrium, the H-type users join with probability 1 and the L-type users join with probability $p \in (0,1)$. Letting π denote the no-information mechanism, we have $\pi_{k,0} = 0$, $\pi_{k,1} > 0$ and $\pi_{k,2} > 0$ for all $k \geq 0$. Furthermore, we have $S_{\mathsf{H},1}(\pi) > S_{\mathsf{L},1}(\pi) = 0$. Thus, by Lemma EC.6 stated later in this section, we obtain that no-information mechanism is Pareto dominated by a signaling mechanism.

Taken together, we obtain the result. $\square$

The following lemmas are used in the proof of Proposition 4.

LEMMA EC.5. *Suppose $m \leq m_{\mathsf{L}}$. Then, for $n \geq m$ we have $W_{\mathsf{L}}(m, n-1) > W_{\mathsf{L}}(m,n)$ and $W_{\mathsf{H}}(m-1,n) > W_{\mathsf{H}}(m,n)$.*

Proof. First we define two auxiliary functions: $\Psi(m) \triangleq W_{\mathsf{L}}(m,n)/Z(m,n)$ and $\Phi(m,n) \triangleq W_{\mathsf{H}}(m,n)/Z(m,n)$ where $Z(m,n)$, $W_{\mathsf{L}}(m,n)$, $W_{\mathsf{H}}(m,n)$ are as defined in footnotes 15 and 16.

Let $m \leq m_{\mathsf{L}}$ and $n \geq m \in \mathbb{N}_0$. We have

$$\begin{aligned} W_{\mathsf{L}}(m,n-1) - W_{\mathsf{L}}(m,n) &= (Z(m,n-1) - Z(m,n))\Psi(m) \\ &= Z(m,n)Z(m,n-1)\left(\frac{1}{Z(m,n)} - \frac{1}{Z(m,n-1)}\right)\Psi(m) \\ &= Z(m,n-1)W_{\mathsf{L}}(m,n)\lambda^m \lambda_{\mathsf{H}}^{n-m}. \end{aligned}$$

Since $m \leq m_{\mathsf{L}}$, we have $W_{\mathsf{L}}(m,n) > 0$. Thus, we obtain $W_{\mathsf{L}}(m,n-1) > W_{\mathsf{L}}(m,n)$. Next, we have

$$\begin{aligned} & W_{\mathsf{H}}(m-1,n) - W_{\mathsf{H}}(m,n) \\ &= Z(m-1,n)\Phi(m-1,n) - Z(m,n)\Phi(m,n) \\ &= Z(m-1,n)\left(\Phi(m-1,n) - \Phi(m,n)\right) + \Phi(m,n)\left(Z(m-1,n) - Z(m,n)\right) \\ &= Z(m-1,n)Z(m,n)\lambda_{\mathsf{L}}\lambda^{m-1}\lambda_{\mathsf{H}} \cdot \\ & \quad \left(\left(\sum_{k=0}^{m-1}\lambda^k u_{\mathsf{H}}(k)\right)\left(\sum_{k=m}^{n}\lambda_{\mathsf{H}}^{k-m}\right) - \left(\sum_{k=0}^{m-1}\lambda^k\right)\left(\sum_{k=m}^{n-1}\lambda_{\mathsf{H}}^{k-m}u_{\mathsf{H}}(k)\right)\right) \end{aligned}$$

$$\geq Z(m-1,n)Z(m,n)\lambda_{\mathsf{L}}\lambda^{m-1}\lambda_{\mathsf{H}}\left(\sum_{k=0}^{m-1}\lambda^k\right)\left(\lambda_{\mathsf{H}}^{n-m}u_{\mathsf{H}}(m-1)+\left(\sum_{k=m}^{n-1}\lambda_{\mathsf{H}}^{k-m}\right)(u_{\mathsf{H}}(m-1)-u_{\mathsf{H}}(m))\right)$$
$$>0,$$

where the final inequality follows from the fact that u_{H} is strictly decreasing, and since $m\leq m_{\mathsf{L}}$, we have $u_{\mathsf{H}}(m-1)\geq u_{\mathsf{L}}(m-1)>0$. $\square$

LEMMA EC.6. *Consider a signaling mechanism $\pi=\{\pi_{k,j}:j=0,1,2;k\geq 0\}$ such that $\pi_{k,0}=0$ for all $k\geq 0$ and there exists an $m\geq 0$ with $\pi_{m,1}>0$ and $\pi_{m+1,2}>0$. If in addition $S_{\mathsf{H},1}(\pi)>0$, then π is Pareto dominated by a signaling mechanism.*

Proof. Suppose π is as stated in the lemma statement, and furthermore, there exists an $m\geq 0$ such that $\pi_{m,1}>0$ and $\pi_{m+1,2}>0$. Then, for small enough $\delta>0$, define $\tilde{\pi}$ as follows:

$$\tilde{\pi}_{k,2}=\begin{cases}\pi_{k,2} & \text{for } k<m;\\ \pi_{m,2}+\frac{\delta}{\lambda_{\mathsf{L}}}\sum_{n>m+1}\sum_j\pi_{n,j} & \text{for } k=m;\\ (1-\delta)\pi_{m+1,2}-\frac{\delta\lambda_{\mathsf{H}}}{\lambda_{\mathsf{L}}}\sum_{n\geq m+1}\sum_j\pi_{n,j} & \text{for } k=m+1;\\ (1-\delta)\pi_{k,2} & \text{for } k>m+1;\end{cases}$$
$$\tilde{\pi}_{k,1}=\begin{cases}\pi_{k,1} & \text{for } k<m;\\ \pi_{m,1}-\frac{\delta}{\lambda_{\mathsf{L}}}\sum_{n>m+1}\sum_j\pi_{n,j} & \text{for } k=m;\\ (1-\delta)\pi_{m+1,1}+\frac{\delta\lambda}{\lambda_{\mathsf{L}}}\sum_{n\geq m+1}\sum_j\pi_{n,j} & \text{for } k=m+1;\\ (1-\delta)\pi_{k,1} & \text{for } k>m+1;\end{cases}$$
$$\tilde{\pi}_{k,0}=\pi_{k,0}-\delta\pi_{k,0}\mathbf{I}\{k\geq m+1\}.$$

Then, it is straightforward to verify that $\tilde{\pi}$ satisfies the balance conditions, given by $\lambda\pi_{k,2}+\lambda_{\mathsf{H}}\pi_{k,1}=\sum_j\pi_{k+1,j}$ for all $k\geq 0$. Furthermore, we have, for each $i\in\{\mathsf{H},\mathsf{L}\}$,

$$\begin{aligned}S_{i,2}(\tilde{\pi})-S_{i,2}(\pi)&=\frac{\delta}{\lambda_{\mathsf{L}}}\sum_{n>m+1}\left(\textstyle\sum_j\pi_{n,j}\right)(u_i(m)-u_i(n-1)-\lambda_{\mathsf{H}}(u_i(m+1)-u_i(n)))\\&\quad-\frac{\delta\lambda_{\mathsf{H}}}{\lambda_{\mathsf{L}}}\sum_{n\geq m+1}\pi_{n,0}u_i(n)\\S_{i,1}(\tilde{\pi})-S_{i,1}(\pi)&=-\frac{\delta}{\lambda_{\mathsf{L}}}\sum_{n>m+1}\left(\textstyle\sum_j\pi_{n,j}\right)(u_i(m)-u_i(n-1)-\lambda(u_i(m+1)-u_i(n)))\\&\quad+\frac{\delta\lambda}{\lambda_{\mathsf{L}}}\sum_{n\geq m+1}\pi_{n,0}u_i(n)\\S_{i,0}(\tilde{\pi})-S_{i,0}(\pi)&=-\delta\sum_{n\geq m+1}\pi_{n,0}u_i(n).\end{aligned}$$

Now, for any $a\in[0,1)$, we have for all $n>m+1$,

$$\begin{aligned}u_i(m)-u_i(n-1)-a(u_i(m+1)-u_i(n))&>u_i(m)-u_i(n-1)-(u_i(m+1)-u_i(n))\\&=u_i(m)-u_i(m+1)-(u_i(n-1)-u_i(n))\geq 0,\end{aligned}$$

where the first inequality follows from the fact that $u_i(n)$ is strictly decreasing, and the second inequality follows from the fact that $u_i(k)-u_i(k+1)$ is non-increasing. Furthermore, since $\pi_{m+1,2}>0$, we must have $\sum_j \pi_{m+2,j}>0$. Coupled with the fact that $\pi_{k,0}=0$ for all $k\geq 0$, we obtain for all small enough $\delta>0$ and for $i\in\{\mathsf{H},\mathsf{L}\}$,

$$S_{\mathsf{L},2}(\tilde\pi)>S_{\mathsf{L},2}(\pi), \qquad S_{\mathsf{L},1}(\tilde\pi)<S_{\mathsf{L},1}(\pi), \qquad S_{i,0}(\tilde\pi)=S_{i,0}(\pi).$$

Since π is obedient with $S_{H,1}(\pi)>0$, we conclude that $\tilde\pi$ is obedient as well for small enough $\delta>0$. Finally, we have

$$\begin{aligned}
W_{\mathsf{L}}(\tilde\pi) &= \lambda_{\mathsf{L}} S_{\mathsf{L},2}(\tilde\pi) > \lambda_{\mathsf{L}} S_{\mathsf{L},2}(\pi) = W_{\mathsf{L}}(\pi)\\
W_{\mathsf{H}}(\tilde\pi) &= \lambda_{\mathsf{H}}(S_{\mathsf{H},1}(\tilde\pi)+S_{\mathsf{H},2}(\tilde\pi))\\
&= \lambda_{\mathsf{H}}\left(S_{\mathsf{H},1}(\pi)+S_{\mathsf{H},2}(\pi)\right)\\
&\quad + \delta \sum_{n>m+1}\left(\textstyle\sum_j \pi_{n,j}\right)(u_{\mathsf{H}}(m+1)-u_{\mathsf{H}}(n)) + \delta \sum_{n\geq m+1} \pi_{n,0} u_{\mathsf{H}}(n)\\
&> W_{\mathsf{H}}(\pi),
\end{aligned}$$

where the final inequality follows from the fact that u_{H} is strictly decreasing and $\sum_j \pi_{m+2,j}>0$. Thus, we obtain that π is Pareto dominated by $\tilde\pi$.

Thus, we obtain that any π with $\pi_{k,0}=0$, $S_{\mathsf{H},1}(\pi)>0$ and for which there exists an $m\geq 0$ such that $\pi_{m,1}>0$ and $\pi_{m+1,2}>0$ cannot be Pareto efficient. □

EC.6. Further Numerical Analysis for Section 7.1

In this section, we numerically examine the structure of the optimal signaling mechanism in the fully persuadable population setting introduced in Section 7.1. We start by presenting two examples which show that $\mathsf{sm}(\theta)$ may not have the structure of a threshold mechanism as defined in Section 7.1.

Examples: Suppose $u_{\mathsf{L}}(k)=1-c(k+1)$ and $u_{\mathsf{H}}(k)=1-c(k+1)-\ell_{\mathsf{H}}$ with $c=0.15$, and $\ell_{\mathsf{H}}=-0.7$. Further, let $\lambda_{\mathsf{H}}=0.7$ and $\lambda_{\mathsf{L}}=1-\lambda_{\mathsf{H}}=0.3$. Recall that $\sigma(n,s)\in[0,1]$ denotes the probability of sending signal $s\in\{0,1,2\}$ when the queue length is n. By solving the linear program introduced in Section 7.1, we obtain that:

1. The mechanism $\mathsf{sm}(0.7)$ is given by:

$$\sigma(n,s)=\begin{cases}
\mathbf{I}(s=2) & \text{for } 0\leq n<4;\\
\mathbf{I}(s=1) & \text{for } 4\leq n<10;\\
0.444\times\mathbf{I}(s=1)+0.556\times\mathbf{I}(s=0) & \text{for } n=10;\\
0.190\times\mathbf{I}(s=1)+0.810\times\mathbf{I}(s=0) & \text{for } n=11;\\
\mathbf{I}(s=0) & \text{otherwise,}
\end{cases}$$

2. The mechanism $\mathsf{sm}(0.8)$ is given by:

$$\sigma(n,s)=\begin{cases}\mathbf{I}(s=2) & \text{for } 0\le n<4;\\ \mathbf{I}(s=1) & \text{for } 4\le n<9;\\ 0.774\times\mathbf{I}(s=1)+0.226\times\mathbf{I}(s=0) & \text{for } n=9;\\ \mathbf{I}(s=1) & \text{for } n=10;\\ 0.199\times\mathbf{I}(s=1)+0.801\times\mathbf{I}(s=0) & \text{for } n=11;\\ \mathbf{I}(s=0) & \text{otherwise,}\end{cases}$$

The above examples show that while the signaling mechanism still follows a "monotone" structure by sending signal 2 (i.e., join for both types) for small queue length, and then signal 1 (i.e., leave for L-type and join for H-type) for medium queue length and then signal 0 (i.e., leave for both types) for sufficiently large queue lengths, the queue length at which the mechanism randomizes between the two signals does not necessarily follow the structure of the $\mathsf{Th}(x,y)$ defined at the beginning of Section 7.1.

Even though the above examples show that the optimal signaling can be "slightly" different from $\mathsf{Th}(x,y)$, our numerical analysis confirms that there will be little loss in limiting ourselves to the class of threshold signaling mechanism. As a representative example, for model primitives: $\lambda_{\mathsf{L}}=1-\lambda_{\mathsf{H}}$ with $\lambda_{\mathsf{H}}\in[0,1]$, $\ell_{\mathsf{H}}=-0.7$, $u_{\mathsf{L}}(k)=1-c(k+1)$, and $u_{\mathsf{H}}(k)=1-c(k+1)-\ell_{\mathsf{H}}$ with $c=0.15$, we compute, for each $(\theta,\lambda_{\mathsf{H}})$, the best threshold signaling mechanism (found through exhaustive search on a grid of two thresholds with 1/16 increments) which we denote by tsm. In Figure EC.1, we plot the heat map of $\frac{W(\mathsf{sm},\theta)-W(\mathsf{tsm},\theta)}{W(\mathsf{sm},\theta)-W(\mathsf{fi},\theta)}$. (Note that we use $W(\mathsf{sm},\theta)-W(\mathsf{fi},\theta)$ as the normalization factor to ensure that the ratio is in $[0,1]$.) We observe that the normalized gap is zero for most values of $(\theta,\lambda_{\mathsf{H}})$; in the regime of $(\theta,\lambda_{\mathsf{H}})$ where the gap is nonzero—which includes the examples presented above—it is very small and notably far from 1. Thus our numerical analysis suggests that if for practical reasons, using a threshold mechanism is more desirable, there exists a threshold mechanism which performs nearly as well as the optimal one, and better than the full-information mechanism.

Our further numerical analysis—which we omit for the sake of brevity—shows that in the linear utility case, such deviations from a threshold mechanism only occurs when $|\ell_{\mathsf{H}}|$ is small. For example, for model primitives used in Figure 5 where $\ell_{\mathsf{H}}\in\{-1,-5,-10\}$, for any $\theta\in\{1/12,2/12,\dots,11/12\}$ the optimal signaling mechanism has a threshold structure.

EC.7. Further Numerical Analysis for Section 7.2

In this section, we expand our numerical analysis for the model introduced in Section 7.2, where the two types have different service rates. First, using the linear program developed in Section 7.2 we verify the power of information design for a wide range of gap between the two service rates.

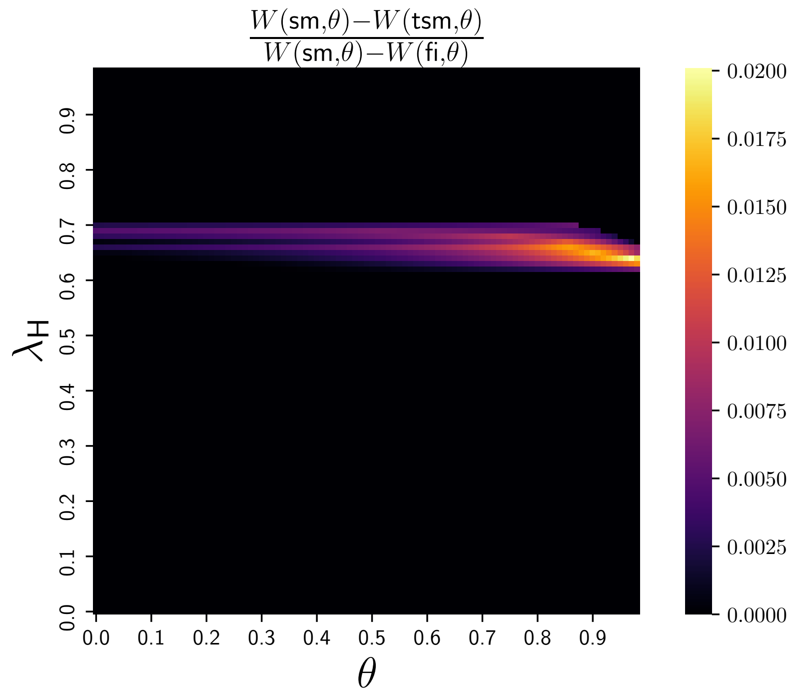


Figure EC.1 Heat map of the normalized welfare gap between optimal signaling mechanism and the best threshold mechanism (found through exhaustive search on a grid of $1/16$ increments). Model primitives: $\lambda_{\mathsf{L}} = 1 - \lambda_{\mathsf{H}}$ with $\lambda_{\mathsf{H}} \in [0,1]$, $\ell_{\mathsf{H}} = -0.7$, $u_{\mathsf{L}}(k) = 1 - c(k+1)$ and $u_{\mathsf{H}}(k) = 1 - c(k+1) - \ell_{\mathsf{H}}$ with $c = 0.15$.

Next, we illustrate the effectiveness of information design in a FCFS system, i.e., without a priority scheme.

In the left panel of Figure EC.2, we compare the welfare outcome of the Pareto-efficient signaling mechanism (sm) and the Pareto-efficient admission policy (ap) when we fix the service rate of H-type to be 1, but vary that of L-type from 0.8 to 1.25. (We recall that under the preemptive priority scheme, the welfare of H-type is unaffected by the signaling mechanism or the admission policy.) In particular, we plot the heat map of $W_{\mathsf{L}}(\mathsf{ap}) - W_{\mathsf{L}}(\mathsf{sm})$ on the plane $(\mu_{\mathsf{L}}, \lambda_{\mathsf{H}}) \in [0.8, 1.25] \times [0,1]$. We observe that for λ_{H} sufficiently large, $W_{\mathsf{L}}(\mathsf{ap}) = W_{\mathsf{L}}(\mathsf{sm})$ for any $\mu_{\mathsf{L}} \in [0.8, 1.25]$, implying that information design is as powerful as the first-best even when the service rate for the L-type users, μ_{L}, is considerably below or above its counterpart μ_{H} for the H-type users. To illustrate that information design remains effective even when λ_{H} is small, in the right panel of Figure EC.2, we plot the welfare of the L-type users under the Pareto-efficient signaling mechanism (sm) and the Pareto-efficient admission policy (ap) and the two benchmarks of full-information and no-information mechanisms when $\mu_{\mathsf{L}} \in [0.8, 1.25]$ and $\lambda_{\mathsf{H}} = 0.3$. We observe that for any μ_{L}, the welfare under sm remains close to that under ap and dominates that under the two benchmarks.

Next, we consider a setting where the two types differ in their service rates but not in their service priority, i.e., we revisit the FCFS queuing discipline when $\mu_{\mathsf{L}} \neq \mu_{\mathsf{H}}$. First, we remark that analyzing

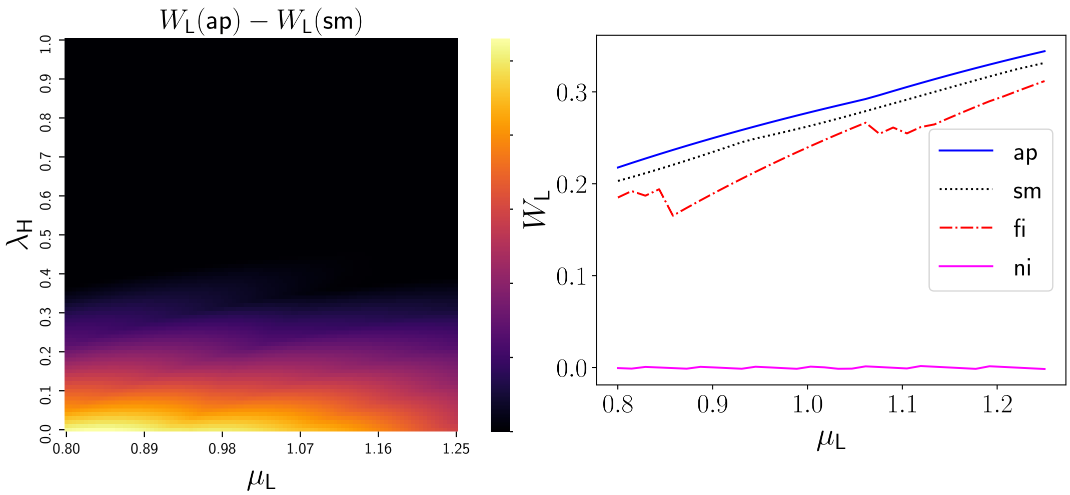


Figure EC.2 Left: Heat map of the welfare gap between the optimal admission policy and optimal public signaling mechanism, i.e., $W_L(\text{ap}) - W_L(\text{sm})$ when varying $(\mu_L, \lambda_H) \in [0.8, 1.25] \times [0, 1]$. Other model primitives are the same as in Figure 6. Right: Welfare outcomes for L-type under the optimal admission policy, optimal public signaling mechanism, full information, and no information when $\lambda_H = 0.3$ and $\mu_L \in [0.8, 1.25]$. Other model primitives are the same as in the left panel.

this setting under the FCFS service discipline is prohibitively challenging because of an explosion in the state space — it is no longer sufficient to keep track of the number of users in the queue (or even the number of users of different types). Instead, one must track the exact *sequence* of the types of users in the queue, as different type sequences (e.g., HHL vs HLH) imply different transitions in the underlying Markovian process. Because of this state space explosion, even numerically computing the Pareto efficient mechanisms under the FCFS discipline is challenging.

Nevertheless, to study the impact of information design, we restrict our attention to the class of threshold signaling mechanisms. As discussed before, this class of mechanisms are practically appealing due to their ease of implementation. To that end, we compute the Pareto-efficient threshold signaling mechanisms and compare its welfare outcomes with those of the two benchmarks of full- and no-information mechanisms as well as the welfare outcomes of the Pareto-efficient admission polices within the class of threshold policies. In Figure EC.3, we present our numerical results for the aforementioned setting and mechanisms. In particular, we consider a system with $\lambda_L = \lambda_H = 0.5$, $\mu_H = 1$, and $\mu_L \in \{1, 1.1, 1.2\}$. (The utility functions are the same as the ones described in Section 7.2.)[19] We observe that even when restricted to the class of threshold signaling mechanisms, information design results in Pareto-improvement compared to the two benchmarks of providing full or no information for all considered service rates.

[19] We note that due to lack of analytical tractability, we compute the Pareto-efficient threshold signaling mechanism and admission policies using discrete event simulations and exhaustive search over thresholds (on the expected wait

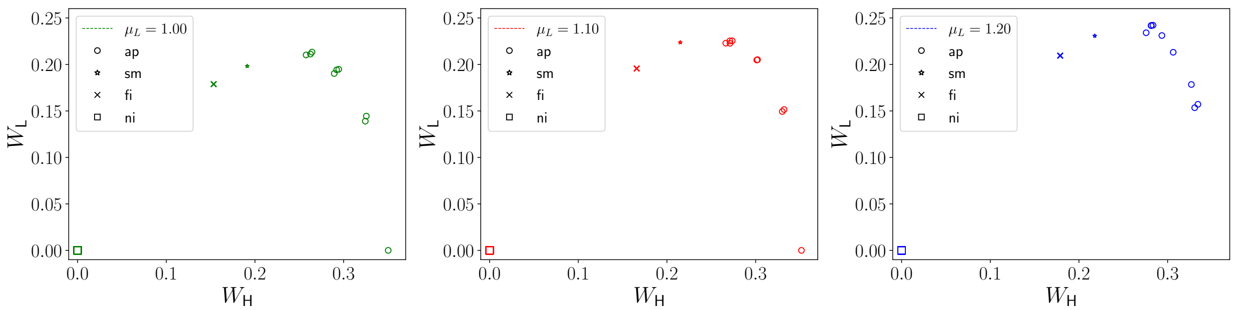


Figure EC.3 Welfare of Pareto-efficient threshold signaling mechanisms, admission policies, full-information, and no-information mechanisms for a FCFS system with $\mu_L = 1$ (left), $\mu_L = 1.1$ (center), and $\mu_L = 1.2$ (right), $\lambda_H = \lambda_L = 0.5$; Other model primitives are the same as in Figure 6.

EC.8. Exogenous Abandonment

In many situations, applicants of a social service may withdraw their request because they no longer need the service. For example, an individual seeking affordable housing may relocate to another city or move in with a partner. To include the possibility of such exogenous abandonment, in this section, we consider the same setting as introduced in Section 2, with one key modification: each arriving user has an independent deadline τ after which she no longer needs the service. More specifically, if she has not already received service by time τ after her arrival, her need for service disappears and she abandons the queue. We assume deadlines are i.i.d. and exponentially distributed with rate γ.

In the presence of exogenous abandonment, we let $u_i(k)$ denote the expected utility of a type-i user for joining when k users are already ahead in queue. Note that some of these users ahead in queue may abandon before completing their service, and the waiting time for a user is lower than than in the no-abandonment case. Furthermore, a user may obtain some utility subsequent to the abandonment. We assume all these aspects are incorporated into the utility function.

Given these modifications, we can follow the same steps as described in Section 2 and (i) establish a correspondence between signaling mechanisms and a set of all distributions satisfying obedience constraints, and (ii) characterize the Pareto frontier of the signaling mechanisms and that of the admission policies by formulating and solving linear optimization problems over feasible steady-state distributions. In particular, to obtain the Pareto frontier of signaling mechanisms, we solve the following linear program for each $\theta \in [0,1]$:

$$\max_{\pi} \quad \theta W_L(\pi) + (1-\theta) W_H(\pi)$$

$$\text{subject to, } J(\pi) \triangleq \sum_{n=0}^{\infty} \left((1+\gamma n)\pi_{n+1} - \lambda_H \pi_n\right) u_L(n) \geq 0, \tag{JOIN}$$

time) with granularity of 0.1.The apparent non-convexity of the Pareto-frontier for admission policies is due to unavoidable simulation noise.

$$L(\pi) \triangleq \sum_{n=0}^{\infty} (\lambda\pi_n - (1+\gamma n)\pi_{n+1})\, u_{\mathsf{L}}(n) \le 0 \tag{LEAVE}$$

$$\lambda_{\mathsf{H}}\pi_n \le (1+\gamma n)\pi_{n+1} \le (\lambda_{\mathsf{L}} + \lambda_{\mathsf{H}})\,\pi_n \quad \text{for all } n \ge 0 \tag{BALANCE}$$

$$\sum_{n=0}^{\infty} \pi_n = 1, \quad \pi_n \ge 0 \quad \text{for all } n \ge 0,$$

where $W_{\mathsf{H}}(\pi)$ is as defined in (3), and $W_{\mathsf{L}}(\pi)$ is given by $W_{\mathsf{L}}(\pi) = J(\pi)$ as defined above. The main difference is in the detailed-balance conditions (BALANCE), which capture the fact that the effective arrival rate into the queue is between λ_{H} and $\lambda = \lambda_{\mathsf{L}} + \lambda_{\mathsf{H}}$, and the effective departure rate equals $1+\gamma n$ when the queue-length equals n. Furthermore, the definitions of $J(\pi)$ and $L(\pi)$ reflect the fact that the joining rate of L-type users into the queue is proportional to $(1+\gamma n)\pi_{n+1} - \lambda_{\mathsf{H}}\pi_n$ when queue-length is n, and the rate of leaving of such users is given by $\lambda\pi_n - (1+\gamma n)\pi_{n+1}$. Finally, note that the Pareto frontier of admission policies can be obtained as before by not imposing the two obedience constraints (JOIN) and (LEAVE) in the preceding program.

For our numerical analysis of this model, we continue to focus on the setting of linear utilities with the same value for service and waiting costs across the two types. It follows from a straightforward analysis that, with n users already in queue, the probability that a joining user receives service (i.e., does not abandon before being served) is given by $\frac{1}{1+(n+1)\gamma}$, and the expected time until service completion or abandonment is given by $\frac{(n+1)}{1+(n+1)\gamma}$. Taken together, the utility function of type-i users is given by $u_i(n) = \frac{1}{1+(n+1)\gamma} \cdot (1 - c(n+1)) + \frac{(n+1)\gamma}{1+(n+1)\gamma} \cdot a_i$, where a_i denotes the utility obtained by a type-i user on abandonment. Note that when $\gamma = 0$ (i.e., with no abandonment), this utility function reduces to the one considered in Section 6.

In Figure EC.4, we plot the welfare of Pareto-efficient signaling mechanisms (stars) and admission policies (circles) for different values of $\lambda_{\mathsf{L}} \in \{0.13, 0.20, 0.30\}$, with $\gamma = 0.02$, $c = 0.15$, and $a_i = 0$ for $i \in \{\mathsf{L}, \mathsf{H}\}$. Similar to the setting in Figure 2, we fix $\lambda = 1$. For each value of λ_{L}, we also plot the full-information mechanism (fi, cross) and the no-information mechanism (ni, square). We observe that for all three values of λ_{L}, the full-information and no-information mechanisms are Pareto dominated by a signaling mechanism, illustrating the power of information design over these simple information sharing benchmarks. Further, the Pareto frontier of signaling mechanisms still overlaps with that of admission policies. Taken together, this numerical example shows that our qualitative insights continue to hold in the presence of exogenous abandonment. Finally, compared with Figure 2, we observe that the welfare of both types improves as the service is less congested due to abandonment.

EC.9. General User Heterogeneity

In our baseline model, we capture the extreme of user heterogeneity by considering two user types, one of which has no viable outside option and must join the service. However, in practice, it

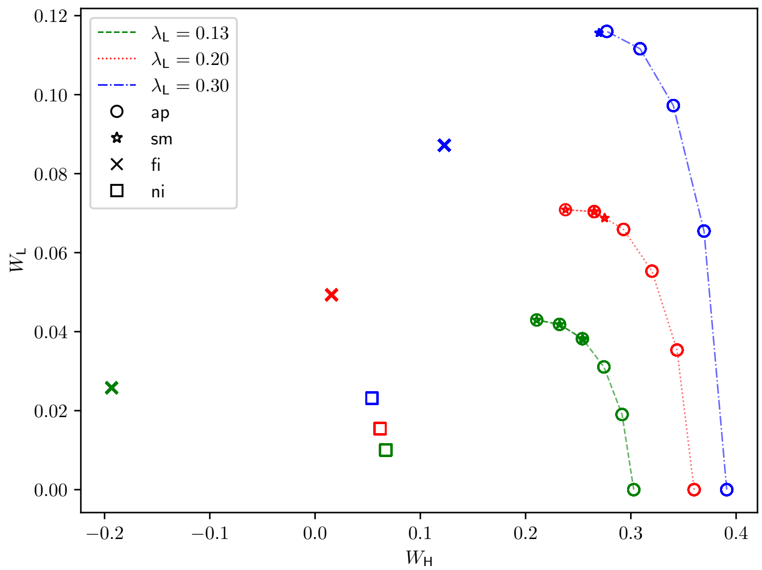


Figure EC.4 Welfare of Pareto-efficient signaling mechanisms and admission policies for $\lambda_L \in \{0.13, 0.20, 0.30\}$, $\lambda_H = 1 - \lambda_L$, $\gamma = 0.02$, $c = 0.15$, $a_L = a_H = 0$. Here, green (dashes) represents $\lambda_L = 0.13$, red (dots) represents $\lambda_L = 0.20$, and blue (dashdots) represents $\lambda_L = 0.30$. Further, circles ($\circ$) represent efficient admission policies (ap), stars ($\star$) represent efficient signaling mechanisms (sm), cross ($\times$) represents the full-information mechanism (fi), and square ($\square$) represents the no-information mechanism (ni).

is reasonable to expect a range of user types with varying levels of need for service and access to outside options. For instance, even among patients with less severe conditions who may be persuaded to avail the alternatives to an emergency room visit, the value of such alternatives might vary substantially based on the patients' symptoms. To incorporate such considerations, in this section, we extend our model to allow for multiple user types that differ in their outside options and value for service. We analyze this model numerically and show that our qualitative insights regarding the effectiveness of information design for welfare improvement continue to hold.

Suppose we have I user types, where a user of type $i \in [I]$ arrives at rate λ_i, gets utility $u_i(n)$ upon joining the queue with n users ahead of her, and has an outside option of $\ell_i \in \mathbb{R} \cup \{-\infty\}$. Here, $\ell_i = -\infty$ captures the case where type-i users have no viable outside option. (We assume no abandonment in this section.) Our baseline model corresponds to the case $I = 2$ with $\ell_1 = 0$ and $\ell_2 = -\infty$.

In practice, a social service provider may not always be able to observe the type of a user. Moreover, ethical concerns may limit a service provider from making information provision depend on the users' outside options. Such limitations may make private signaling infeasible, and due to such practical considerations, we focus on *public signaling mechanisms*. Note that in our baseline model, public and private signaling are the same because high-need users have no outside option and always join irrespective of the belief.

For public signals, using the revelation principle, one can show that it suffices to consider signaling mechanisms where signals correspond to subsets $S \subseteq [I]$, and which are obedient in the sense that when the signal is $S \subseteq [I]$ only users with type $i \in S$ find it optimal to join the queue, whereas users with type $j \notin S$ find it optimal to leave. Focusing on such signaling mechanisms, similar arguments as in Section 2 allow us to formulate a linear program to compute the Pareto-efficient (public) signaling mechanisms. To see this, for a public signaling mechanism, let $x_{n,S}$ denote the joint probability (in steady-state) that the queue-length upon arrival of a user is n and the user receives the signal $S \subseteq [I]$, and note that $\sum_{S\subseteq[I]} x_{n,S}$ denotes the probability that the queue-length is n in steady-state. The detailed-balance condition can then be written as $\sum_{S\subseteq[I]} \lambda_S x_{n,S} = \sum_{S\subseteq[I]} x_{n+1,S}$, where $\lambda_S = \sum_{i\in S} \lambda_i$ denotes the total arrival rate of the users with type $i \in S$. The welfare of type-i users can then be written as a function of $x = \{x_{n,S} : n \geq 0, S \subseteq [I]\}$ as follows:

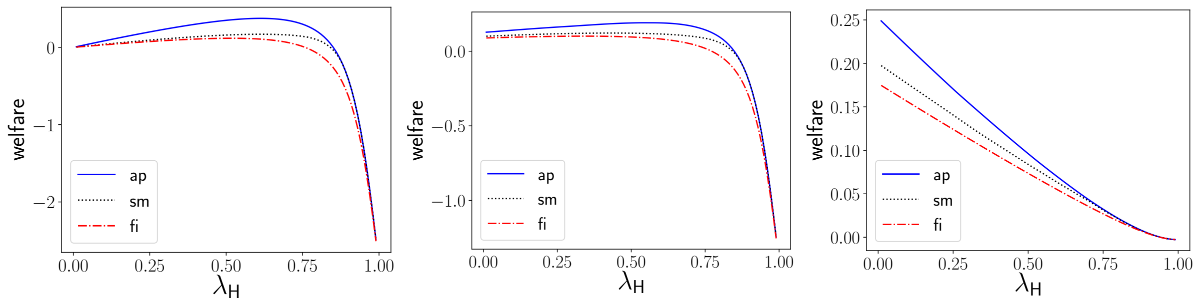
$$W_i(x) = \lambda_i \left(\sum_{n\geq 0} \sum_{S\ni i} x_{n,S} u_i(n) + \sum_{n\geq 0} \sum_{S\not\ni i} x_{n,S} \ell_i \right).$$

Further, upon receiving a signal $S \subseteq [I]$, since a user with type $i \in S$ finds it optimal to join the queue, this implies $\sum_{n\geq 0} x_{n,S}(u_i(n) - \ell_i) \geq 0$ for $i \in S$. Similarly, since a user with type $i \notin S$ finds it optimal to leave, we have $\sum_{n\geq 0} x_{n,S}(u_i(n) - \ell_i) \leq 0$ for $i \notin S$. Putting it all together, it follows that the Pareto-efficient public signaling mechanisms correspond to the optimal solutions of the following linear program for different choices of non-negative weights $\theta = (\theta_i : i \in [I])$:

$$\begin{aligned}
\max_x \quad & \sum_{i\in[I]} \theta_i W_i(x) \\
\text{subject to,} \quad & \sum_{n\geq 0} x_{n,S}(u_i(n) - \ell_i) \geq 0, \quad \text{for } i \in S \text{ and } S \subseteq [I], \\
& \sum_{n\geq 0} x_{n,S}(u_i(n) - \ell_i) \leq 0, \quad \text{for } i \notin S \text{ and } S \subseteq [I], \\
& \sum_{S\subseteq[I]} \lambda_S x_{n,S} = \sum_{S\subseteq[I]} x_{n+1,S}, \quad \text{for all } n \geq 0, \\
& \sum_{n\geq 0} \sum_{S\subseteq[I]} x_{n,S} = 1, \text{ and } x_{n,S} \geq 0 \text{ for all } n \geq 0, \text{ and } S \subseteq [I].
\end{aligned}$$

Using this linear program, we numerically investigate the effectiveness of information design by comparing it to the full-information mechanism and the Pareto-efficient admission polices.[20] In particular, we consider an example with three users types, all with the same linear utility function

[20] We compute the Pareto-efficient admission polices for any given weight $\theta = (\theta_i : i \in [I])$ by dropping the obedience constraints from the preceding linear program.



(a) $\theta = (0,0,1)$ (b) $\theta = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$ (c) $\theta = (\frac{1}{2}, \frac{1}{2}, 0)$

Figure EC.5 Welfare of the Pareto-efficient (public) signaling mechanism $\mathsf{sm}(\theta)$, the Pareto-efficient admission policy $\mathsf{ap}(\theta)$, and the full-information mechanism fi for three types $I = \{1,2,3\}$. Here, the arrival rates are given by $(\lambda_1, \lambda_2, \lambda_3) = (\lambda_{\mathsf{L}}, \lambda_{\mathsf{L}}, \lambda_{\mathsf{H}})$ with $\lambda_{\mathsf{L}} = (1-\lambda_{\mathsf{H}})/2$. The outside options equal $(\ell_1, \ell_2, \ell_3) = (0, -0.25, -\infty)$, and $u_i(k) = 1 - c(k+1)$ with $c = 0.15$.

$u(n) = 1 - c(n+1)$ with $c = 0.15$ but different outside options given by $(\ell_1, \ell_2, \ell_3) = (0, -0.25, -\infty)$. In particular, the first two types have viable outside options, while the third type has no viable outside option. In keeping with the terminology of our baseline model for easier comparison, we refer to the third-type as the H type, and the first two types as L types. The arrival rates are given by $(\lambda_1, \lambda_2, \lambda_3) = (\lambda_{\mathsf{L}}, \lambda_{\mathsf{L}}, \lambda_{\mathsf{H}})$ with $\lambda_{\mathsf{L}} = (1-\lambda_{\mathsf{H}})/2$.

In Figure EC.5, we plot the welfare $W(\pi,\theta)$ as a function of the arrival rate λ_{H} for the full-information mechanism, the optimal (public) signaling mechanism $\mathsf{sm}(\theta)$ and the optimal admission policy $\mathsf{ap}(\theta)$ for $\theta = (\theta_1, \theta_2, \theta_3) \in \left\{(0,0,1), \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}, 0\right)\right\}$. Similar to our observations in Figure 3, we observe that information design results in welfare improvement over the full-information mechanism when the user population is fairly balanced (given our parametrization of the arrival rates, this corresponds to λ_{H} being not too large). Further, for large enough λ_{H}, the optimal signaling mechanism $\mathsf{sm}(\theta)$ achieves the same welfare as that of the optimal admission policy $\mathsf{ap}(\theta)$. This point is further illustrated in Figure EC.6 (left panel), where, for the weight parametrization $\theta = (\frac{\theta_{\mathsf{L}}}{2}, \frac{\theta_{\mathsf{L}}}{2}, 1-\theta_{\mathsf{L}})$, we display the region in the $(\theta_{\mathsf{L}}, \lambda_{\mathsf{H}})$ plane where the optimal signaling mechanism achieves the same welfare as the optimal admission policy. Finally, Figure EC.6 (right panel) shows that even in regions where the two policies do not achieve the same welfare, the welfare gap is fairly small.

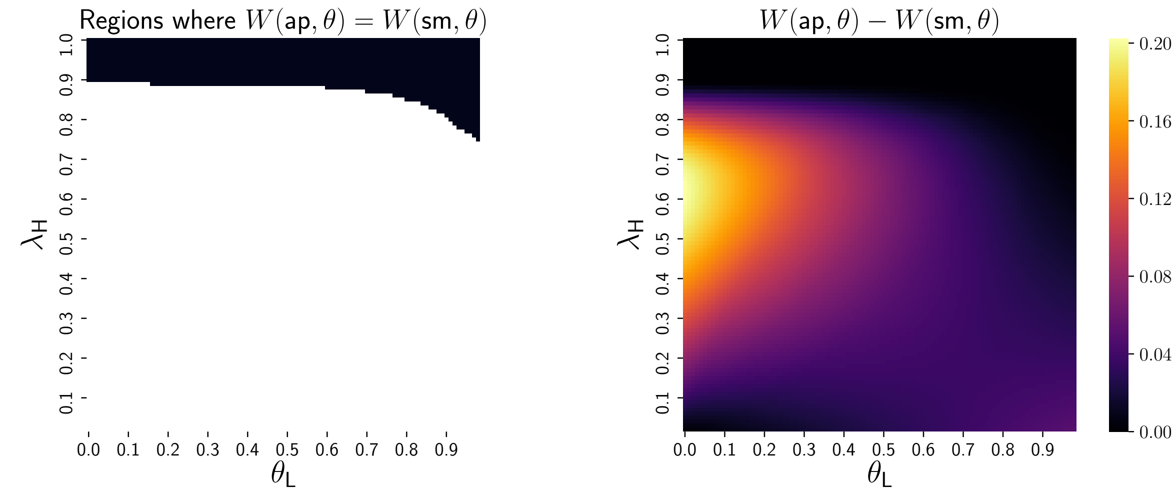


Figure EC.6 **Left: Regions of the $(\theta, \lambda_\mathsf{H})$ plane for which $\mathsf{sm}(\theta) = \mathsf{ap}(\theta)$, i.e., the signaling mechanism $\mathsf{sm}(\theta)$ is Pareto-efficient within Π_AP. Right: Heat map of the welfare gap between the optimal admission policy and optimal public signaling mechanism, i.e., $W(\mathsf{ap},\theta) - W(\mathsf{sm},\theta)$. Model primitives: The arrival rates are given by $(\lambda_1, \lambda_2, \lambda_3) = (\lambda_\mathsf{L}, \lambda_\mathsf{L}, \lambda_\mathsf{H})$ with $\lambda_\mathsf{L} = (1-\lambda_\mathsf{H})/2$. The outside options equal $(\ell_1, \ell_2, \ell_3) = (0, -0.25, -\infty)$, $u_i(k) = 1 - c(k+1)$ with $c = 0.15$, and the welfare weights are given by $\theta = (\theta_\mathsf{L}/2, \theta_\mathsf{L}/2, 1-\theta_\mathsf{L})$.**