Ride-Hailing Networks with Strategic Drivers: The Impact of Platform Control Capabilities on Performance.

Supplemental Materials

S1 Supplemental Lemmas and Proofs for Control Regime A

Under regime A, the lower capacity threshold n_2^A of zone (3)—moderate capacity *with repositioning*—and the optimal capacity allocation within this zone, do not have explicit expressions (see Proposition 4). Lemmas S-1 and S-2 fill in the remaining details.

Given a level of participating capacity n, Proposition 4 shows that the optimal capacity allocation in zone (3) has $r_{12} > 0, r_{21} = 0$ and $q = (q_1^*(s), 0)$. Therefore, the constraints of Problem A at a given capacity n, (12a)–(12c) and (22), simplify to

$$\overline{s} + \left(\frac{t_{12}}{t_{21}} s_{21} - s_{12}\right) + q_1^*(s) = n \tag{S.1}$$

and $0 \leq s \leq S$, $\frac{t_{12}}{t_{21}} s_{21} > s_{12}$. Note that s determines r_{12} by the second term and q by $q_1^*(s)$.

Lemma S-1 shows that there are three possible optimal capacity allocation *patterns* at any level of participating capacity in zone (3). These patterns differ in terms of whether demand is rejected at the low-demand location, and if so, for which route(s).

Lemma S-1. *Under control regime A, the optimal capacity allocation as a function of the participating capacity $n \in (n_2^A, n_3^A]$ (moderate capacity zone with repositioning) follows one of three patterns, denoted by $s_i(n)$, $i = 1, 2, 3$. Let $\overline{s_i}(n)$ denote the total service capacity under pattern i, and $\overline{s_i}^{-1}(\cdot)$ denote its inverse. The optimal pattern $i^*(n)$ is the one that attains the maximum service capacity, i.e., $i^*(n) = \operatorname{argmax}_{i \in \{1,2,3\}} \overline{s_i}(n)$.*

(1) <u>No demand rejection at the low-demand location</u>: only s_{21} is increasing in this zone.

$$s_1(n) = (S_{11}, S_{12}, s_{21}, S_{22}) \text{ subject to (S.1)}, \quad n \in (\overline{s_1}^{-1}(n_1^A), n_3^A].$$

(2) <u>Rejecting only cross-traffic demand at the low-demand location</u>: for small n, s_{21} is increasing while $s_{12} = 0$; for large n, $s_{21} = S_{21}$ and s_{12} is increasing.

$$s_2(n) = (S_{11}, s_{12}, s_{21}, S_{22}) \text{ subject to } (S_{21} - s_{21}) s_{12} = 0 \text{ and (S.1)}, \quad n \in (\overline{s_2}^{-1}(n_1^A), n_3^A].$$

(3) <u>Rejecting local and cross-traffic demand at the low-demand location</u>: for small n, s_{21} is increasing while $s_{11} = s_{12} = 0$; for medium n, $s_{21} = S_{21}$, s_{11} is increasing and $s_{12} = 0$; for large n, $s_{21} = S_{21}, s_{11} = S_{11}$ and s_{12} is increasing.

$$s_3(n) = (s_{11}, s_{12}, s_{21}, S_{22}) \text{ subject to } (S_{21} - s_{21}) s_{11} = (S_{11} - s_{11}) s_{12} = 0 \text{ and (S.1)}, \quad n \in (\overline{s_3}^{-1}(n_1^A), n_3^A].$$

Proof. By Proposition 4, the optimal capacity allocation of given participating capacity $n \in (n_2^A, n_3^A]$ has $r_{12} > 0, r_{21} = 0$ and $q = (q_1^*(s), 0)$. Therefore, for fixed n, Problem A reduces to $\max_{s,r,q}\{\mathit{\Pi}(s) :$ (12a) –

(12c), (22)}, and it can be reformulated as maximizing the total service capacity over feasible service capacity vector s:

$$\max_{s} \quad \bar{s} \tag{S.2a}$$

$$\text{s.t.} \quad g(s) := \bar{s} + \left(\frac{t_{12}}{t_{21}} s_{21} - s_{12}\right) + q_1^*(s) \le n \tag{S.2b}$$

$$0 \le s_{ij} \le S_{ij},\ \forall i, j. \tag{S.2c}$$

Note that $g(s)$ is the total capacity expressed as a function of s. Relaxing the equality constraint (12b) to the inequality constraint (S.2b) does not matter since positive q_2 and $q_1 = q_1^*(s) + k(s)q_2$ are feasible by (22) (but not optimal by Proposition 4). Constraint $r_{12} = \frac{t_{12}}{t_{21}} s_{21} - s_{12} > 0$ is omitted since a violation results in $s_{21} \le \frac{t_{21}}{t_{12}} S_{12} \Rightarrow \bar{s} \le n_1^A$, clearly suboptimal in zone (3).

Let $\alpha, \overline{\beta}_{ij}, \underline{\beta}_{ij}$ be the dual variables associated with the capacity constraint (S.2b), the upper and lower bound constraints (S.2c), respectively. The KKT conditions are

$$\text{(stationarity)} \quad \alpha \frac{\partial g(s)}{\partial s_{ij}} + \overline{\beta}_{ij} - \underline{\beta}_{ij} = 1,\ \forall i, j, \tag{S.3a}$$

$$\text{(complementary slackness)} \quad \alpha(n - g(s)) = \overline{\beta}_{ij}(S_{ij} - s_{ij}) = \underline{\beta}_{ij} s_{ij} = 0,\ \forall i, j, \tag{S.3b}$$

$$\text{(dual feasibility)} \quad \alpha, \overline{\beta}_{ij}, \underline{\beta}_{ij} \ge 0,\ \forall i, j, \tag{S.3c}$$

$$\text{(primal feasibility)} \quad g(s) \le n, \tag{S.3d}$$

$$\text{(primal feasibility)} \quad 0 \le s_{ij} \le S_{ij}, \forall i, j. \tag{S.3e}$$

The complementary slackness constraints (S.3b) and dual feasibility constraints (S.3c) establish the relationship between primal and dual variables: $\overline{\beta}_{ij} = 0$ ($\underline{\beta}_{ij} = 0$) when s_{ij} is not at its upper (lower) bound; s_{ij} must be at its upper (lower) bound when $\overline{\beta}_{ij} > 0$ ($\underline{\beta}_{ij} > 0$); $\alpha = 0$ when $g(s) < n$ and $g(s) = n$ when $\alpha > 0$. Moreover, $\overline{\beta}_{ij} \cdot \underline{\beta}_{ij} = 0$. We omit explicit references to the primal and dual feasibility constraints (S.3c)–(S.3e) in the following proof.

To prove the lemma, we will use the above KKT conditions to establish that any optimal solution to problem (S.2a)–(S.2c) for $n \in (n_2^A, n_3^A]$ must satisfy four necessary conditions (a)–(d) stated below. Before that, we calculate some first and second partial derivatives of $g(s)$ that will also be used to prove the four conditions:

$$\frac{\partial g(s)}{\partial s_{11}} = \frac{s_{21}(1 + \frac{t_{12}}{t_{21}}) + s_{22}}{(s_{21} + s_{22})\bar{\gamma}p - \left(s_{21} + s_{22} + s_{21}\frac{t_{12}}{t_{21}}\right)c}(\bar{\gamma}p - c) > 1, \tag{S.4}$$

$$\frac{\partial g(s)}{\partial s_{12}} = \frac{s_{21}(1 + \frac{t_{12}}{t_{21}}) + s_{22}}{(s_{21} + s_{22})\bar{\gamma}p - \left(s_{21} + s_{22} + s_{21}\frac{t_{12}}{t_{21}}\right)c}\bar{\gamma}p > \frac{\partial g(s)}{\partial s_{11}} > 1, \tag{S.5}$$

$$\frac{\partial g(s)}{\partial s_{21}} = 1 + \frac{t_{12}}{t_{21}} + \frac{(s_{11}(\bar{\gamma}p - c) + s_{12}\bar{\gamma}p)s_{22}\frac{t_{12}}{t_{21}}}{\left[(s_{21} + s_{22})\bar{\gamma}p - \left(s_{21} + s_{22} + s_{21}\frac{t_{12}}{t_{21}}\right)c\right]^2}\bar{\gamma}p > 1, \tag{S.6}$$

$$\frac{\partial g(s)}{\partial s_{22}} = 1 - \frac{(s_{11}(\bar{\gamma}p - c) + s_{12}\bar{\gamma}p)s_{21}\frac{t_{12}}{t_{21}}}{\left[(s_{21} + s_{22})\bar{\gamma}p - \left(s_{21} + s_{22} + s_{21}\frac{t_{12}}{t_{21}}\right)c\right]^2}\bar{\gamma}p < 1. \tag{S.7}$$

It follows that

$$\frac{\partial^2 g(s)}{\partial s_{11}^2} = \frac{\partial^2 g(s)}{\partial s_{12}^2} = \frac{\partial^2 g(s)}{\partial s_{11}\partial s_{12}} = 0, \quad \frac{\partial^2 g(s)}{\partial s_{11}\partial s_{21}}, \frac{\partial^2 g(s)}{\partial s_{12}\partial s_{21}} > 0, \quad \frac{\partial^2 g(s)}{\partial s_{21}^2} \le 0, \tag{S.8}$$

where the last inequality is strict when $s_{11} + s_{12} > 0$.

Now we are ready to state and prove the four necessary conditions.

(a) All capacity is used within this zone and s_{21} has a lower bound:

$$g(s) = n, \tag{S.9}$$

$$S_{12}\frac{t_{21}}{t_{12}} \leq s_{21}. \tag{S.10}$$

To prove (S.9), note that when $s \neq S$, pick any $s_{ij} < S_{ij}$, then $\overline{\beta}_{ij} = 0$ by (S.3b). By (S.3a) this implies that $\alpha \neq 0$ and hence $g(s) = n$ by (S.3b). When $s = S$, $g(s) = n = n_3^A$. For (S.10), $s_{21} \geq S_{12}\frac{t_{21}}{t_{12}}$ follows directly from $\overline{s} > n_1^A$ in zone (3). By (S.3b), $s_{21} > 0$ also implies $\underline{\beta}_{21} = 0$.

(b) Rejecting local demand at the high-demand location (s_{22}) is suboptimal:

$$s_{22} = S_{22}. \tag{S.11}$$

Using $\underline{\beta}_{21} = 0$ from part (a) and $\frac{\partial g(s)}{\partial s_{21}} > 1$, stationarity constraints (S.3a) imply that $\alpha < 1$. Putting this and $\frac{\partial g(s)}{\partial s_{22}} < 1$ back to (S.3a), we obtain $\overline{\beta}_{22} > 0$. Therefore it follows from (S.3b) that $s_{22} = S_{22}$ and $\underline{\beta}_{22} = 0$.

(c) Rejecting cross-traffic demand (s_{12}) is more profitable than rejecting local demand (s_{11}) at the low-demand location:

$$(S_{11} - s_{11})s_{12} = 0. \tag{S.12}$$

We prove this by contradiction using (S.3a) and (S.3b). Suppose on the contrary $(S_{11} - s_{11})s_{12} > 0$ for some $s_{11} < S_{11}$ and $s_{12} > 0$, then (S.3b) require $\overline{\beta}_{11} = \underline{\beta}_{12} = 0$ and hence (S.3a) yield

$$\alpha\frac{\partial g(s)}{\partial s_{11}} - \underline{\beta}_{11} = \alpha\frac{\partial g(s)}{\partial s_{12}} + \overline{\beta}_{12} = 1.$$

This cannot happen due to $\frac{\partial g(s)}{\partial s_{11}} < \frac{\partial g(s)}{\partial s_{12}}$ and (S.3c). Therefore we must have $(S_{11} - s_{11})s_{12} = 0$.

(d) Neither demand stream at the low-demand location is partially served unless s_{21} is fully served:

$$s_{12}(S_{12} - s_{12})(S_{21} - s_{21}) = 0, \tag{S.13}$$

$$s_{11}(S_{11} - s_{11})(S_{21} - s_{21}) = 0. \tag{S.14}$$

We prove the two equations in similar ways by showing that any violation will lead to suboptimality. For (S.13), suppose on the contrary $s_{12}(S_{12} - s_{12})(S_{21} - s_{21}) > 0$ for some $0 < s_{12} < S_{12}$ and $s_{21} < S_{21}$, then (S.3b) and $\underline{\beta}_{21} = 0$ from part (a) require $\overline{\beta}_{12} = \underline{\beta}_{12} = \overline{\beta}_{21} = \underline{\beta}_{21} = 0$. It hence follows from (S.3a) that

$$\alpha\frac{\partial g(s)}{\partial s_{12}} = \alpha\frac{\partial g(s)}{\partial s_{21}} = 1,$$

thus $\alpha > 0$ and $1 < \frac{\partial g(s)}{\partial s_{11}} < \frac{\partial g(s)}{\partial s_{12}} = \frac{\partial g(s)}{\partial s_{21}}$. By the second derivatives in (S.8), increasing s_{12} and decreasing s_{21} will always maintain the inequality

$$1 < \frac{\partial g(s)}{\partial s_{11}} < \frac{\partial g(s)}{\partial s_{12}} < \frac{\partial g(s)}{\partial s_{21}}. \tag{S.15}$$

Therefore we can keep increasing s_{12} ($\Delta s_{12} > 0$) and decreasing s_{21} ($\Delta s_{21} < 0$) simultaneously such that the following equality holds at any subsequent s_{12} and s_{21}:

$$\Delta s_{12}\frac{\partial g(s)}{\partial s_{12}} + \Delta s_{21}\frac{\partial g(s)}{\partial s_{21}} = 0. \tag{S.16}$$

In this way we can maintain

$$\Delta g(s)=\sum_{i,j}\frac{\partial g(s)}{\partial s_{ij}}\Delta s_{ij}=\Delta s_{12}\frac{\partial g(s)}{\partial s_{12}}+\Delta s_{21}\frac{\partial g(s)}{\partial s_{21}}=0,$$

i.e., keep $g(s)$ constant, while improving the objective function (service capacity) by

$$\Delta\overline{s}=\Delta s_{12}+\Delta s_{21}=\Delta s_{12}\left(1-\frac{\partial g(s)/\partial s_{12}}{\partial g(s)/\partial s_{21}}\right)>0,$$

which follows from (S.15) and (S.16), until $s_{12}(S_{12}-s_{12})(S_{21}-s_{21})=0$ is satisfied.

Similarly, for (S.14), suppose on the contrary $s_{11}(S_{11}-s_{11})(S_{21}-s_{21})>0$ for some $0<s_{11}<S_{11}$ and $s_{21}<S_{21}$, then (S.3b) and $\underline{\beta}_{21}=0$ from part (a) require $\overline{\beta}_{11}=\underline{\beta}_{11}=\overline{\beta}_{21}=\underline{\beta}_{21}=0$. It hence follows from (S.3a) that

$$\alpha\frac{\partial g(s)}{\partial s_{11}}=\alpha\frac{\partial g(s)}{\partial s_{21}}=1,$$

thus $\alpha>0$ and $1<\frac{\partial g(s)}{\partial s_{21}}=\frac{\partial g(s)}{\partial s_{11}}<\frac{\partial g(s)}{\partial s_{12}}$. By the second derivatives in (S.8), increasing s_{21} and decreasing s_{11} will always maintain the inequality

$$1<\frac{\partial g(s)}{\partial s_{21}}<\frac{\partial g(s)}{\partial s_{11}}<\frac{\partial g(s)}{\partial s_{12}}. \tag{S.17}$$

Therefore we can keep increasing s_{21} ($\Delta s_{21}>0$) and decreasing s_{11} ($\Delta s_{11}<0$) simultaneously such that the following equality holds at any subsequent s_{11} and s_{21}:

$$\Delta s_{11}\frac{\partial g(s)}{\partial s_{11}}+\Delta s_{21}\frac{\partial g(s)}{\partial s_{21}}=0. \tag{S.18}$$

In this way we can maintain

$$\Delta g(s)=\sum_{i,j}\frac{\partial g(s)}{\partial s_{ij}}\Delta s_{ij}=\Delta s_{11}\frac{\partial g(s)}{\partial s_{11}}+\Delta s_{21}\frac{\partial g(s)}{\partial s_{21}}=0,$$

i.e., keep $g(s)$ constant, while improving the objective function (service capacity) by

$$\Delta\overline{s}=\Delta s_{11}+\Delta s_{21}=\Delta s_{21}\left(1-\frac{\partial g(s)/\partial s_{21}}{\partial g(s)/\partial s_{11}}\right)>0,$$

which follows from (S.17) and (S.18), until $s_{11}(S_{11}-s_{11})(S_{21}-s_{21})=0$ is satisfied.

It is then easy to verify that the above necessary conditions (a)–(d) directly imply the three patterns in Lemma S-1 as how service capacity varies with an increasing n in the moderate capacity zone. By condition (b), s_{22} stays constant at S_{22}. By condition (a), $s_{21}>0$ and there are two cases:

(1) $0<s_{21}<S_{21}$: it follows from condition (d) that there are four possible values of s_{11} and s_{12}: (i) $s_{12}=S_{12}, s_{11}=S_{11}$, leading to pattern (1); (ii) $s_{12}=S_{12}, s_{11}=0$, contradicting condition (c); (iii) $s_{12}=0, s_{11}=S_{11}$, leading to pattern (2) with small n; (iv) $s_{12}=0, s_{11}=0$, leading to pattern (3) with small n.

(2) $s_{21}=S_{21}$: condition (d) is satisfied for any values of feasible s_{11} and s_{12}. It follows from condition (c) that (i) $s_{12}>0\Rightarrow s_{11}=S_{11}$, leading to pattern (2) with large n; (ii) $s_{11}<S_{11}\Rightarrow s_{12}=0$, leading to pattern (3) with medium n; or (iii) $s_{11}=S_{11}$ and $s_{12}>0$, leading to pattern (3) with large n.

We have thus shown all the three possible patterns.

Note that for each pattern i, the service capacity $\overline{s_i}(n)$ increases with n from $n=\overline{s_i}^{-1}(n_1^A)$, where $\overline{s}=n_1^A$ is equal to the constant service capacity in zone (2), and up to $n=n_3^A$, the right end of zone (3). This also implies that $n_2^A = \min_i\{\overline{s_i}^{-1}(n_1^A)\}$. $\square$

Next we prove the monotonicity of the per-driver profit rate with respect to n in zone (3).

Lemma S-2. *Per-driver profit rate under control regime A, $\pi_A(n)$, is decreasing in n for $n \in (n_2^A, n_3^A]$.*

Proof. Substituting $\overline{s}$ and $\overline{r}$ from Proposition 4 into (12d) yields

$$\pi_A(n) = \frac{(\bar{\gamma}p - c)\overline{s} - c\overline{r}}{n} = \begin{cases} \bar{\gamma}p - c & \text{zone (1) } (n \le n_1^A), \\ \frac{n_1^A}{n}(\bar{\gamma}p - c) & \text{zone (2) } (n_1^A < n \le n_2^A), \\ \frac{1}{n}[(\bar{\gamma}p - c)\overline{s}^* - c\overline{r}^*] & \text{zone (3) } (n_2^A < n \le n_3^A), \\ \frac{1}{n}(\bar{\gamma}p\overline{S} - cn_2^C) & \text{zone (4) } (n > n_3^A). \end{cases} \tag{S.19}$$

Lemma S-1 shows that for participating capacity $n \in (n_2^A, n_3^A]$, the optimal capacity allocation *may alternate* among three patterns characterized by $s_i(n)$, with service capacity $\overline{s_i}(n)$ for $i = 1, 2, 3$. To prove this lemma, we show that the per-driver profit rate is decreasing for n varying within each of the 3 patterns or at *feasible* transitions between patterns. First, note that $n = g(s)$ in zone (3) (see (S.9)) from the proof of Lemma S-1. Hence $\partial n/\partial s_{ij} = \partial g(s)/\partial s_{ij}$. Then:

(i) Within pattern (1): only s_{21} is increasing,

$$\begin{aligned} \pi'(n) &= \frac{\left[(\bar{\gamma}p - c) - c\frac{t_{12}}{t_{21}}\right]\left(\frac{\partial g(s)}{\partial s_{21}}\right)^{-1} n - [(\bar{\gamma}p - c)\overline{s} - cr_{12}]}{n^2} \\ &= -\frac{S_{22}\bar{\gamma}p}{n^2}\frac{t_{12}}{t_{21}}\left(\frac{\partial g(s)}{\partial s_{21}}\right)^{-1}\left(1 + \frac{S_{11}(\bar{\gamma}p - c) + S_{12}\bar{\gamma}p}{(s_{21} + S_{22})\bar{\gamma}p - \left(s_{21} + S_{22} + s_{21}\frac{t_{12}}{t_{21}}\right)c}\right)^2 < 0. \end{aligned} \tag{S.20}$$

(ii) Within pattern (2): for small n, s_{21} is increasing while $s_{12} = 0$,

$$\begin{aligned} \pi'(n) &= \frac{\left[(\bar{\gamma}p - c) - c\frac{t_{12}}{t_{21}}\right]\left(\frac{\partial g(s)}{\partial s_{21}}\right)^{-1} n - [(\bar{\gamma}p - c)\overline{s} - cr_{12}]}{n^2} \\ &= -\frac{S_{22}\bar{\gamma}p}{n^2}\frac{t_{12}}{t_{21}}\left(\frac{\partial g(s)}{\partial s_{21}}\right)^{-1}\left(1 + \frac{S_{11}(\bar{\gamma}p - c)}{(s_{21} + S_{22})\bar{\gamma}p - \left(s_{21} + S_{22} + s_{21}\frac{t_{12}}{t_{21}}\right)c}\right)^2 < 0. \end{aligned} \tag{S.21}$$

For large n, $s_{21} = S_{21}$ and s_{12} is increasing,

$$\pi'(n) = \frac{\bar{\gamma}p\left(\frac{\partial g(s)}{\partial s_{12}}\right)^{-1} n - [(\bar{\gamma}p - c)\overline{s} - cr_{12}]}{n^2} = 0. \tag{S.22}$$

Note that $\pi(n)$ is continuous at the turning (non-differentiable) point where $s = (S_{11}, 0, S_{21}, S_{22})$.

(iii) Within pattern (3): for small n, s_{21} is increasing while $s_{11} = s_{12} = 0$,

$$\pi'(n) = \frac{\left[(\bar{\gamma}p - c) - c\frac{t_{12}}{t_{21}}\right]\left(\frac{\partial g(s)}{\partial s_{21}}\right)^{-1} n - [(\bar{\gamma}p - c)\overline{s} - cr_{12}]}{n^2} = -\frac{S_{22}\bar{\gamma}p}{n^2}\frac{t_{12}}{t_{21}}\left(\frac{\partial g(s)}{\partial s_{21}}\right)^{-1} < 0. \tag{S.23}$$

For medium n, $s_{21} = S_{21}$, s_{11} is increasing and $s_{12} = 0$,

$$\pi'(n) = \frac{(\bar{\gamma}p - c)\left(\frac{\partial g(s)}{\partial s_{11}}\right)^{-1} n - [(\bar{\gamma}p - c)\overline{s} - cr_{12}]}{n^2} = 0. \tag{S.24}$$

For large n, $s_{21} = S_{21}, s_{11} = S_{11}$ and s_{12} is increasing, we have the same (S.22). Note that $\pi(n)$ is continuous at the two turning (non-differentiable) points where $s = (0, 0, S_{21}, S_{22})$ and $s = (S_{11}, 0, S_{21}, S_{22})$.

As n increases, an *optimal transition* from pattern i to j at n must satisfy

$$\overline{s_i}(n) = \overline{s_j}(n) \quad \text{and} \quad \overline{s_i}'(n^-) < \overline{s_j}'(n^+). \tag{S.25}$$

Namely, patterns i and j have the same service capacity at transition n, and the service capacity increases faster after the transition. We then discuss all three possible transitions.

(i) Between patterns (1) and (2), it is only optimal to transit from (1) to (2). Recall pattern (2) given in Lemma S-1, n can be small or large at the transition. If n is *small* such that s_{21} is increasing while $s_{12} = 0$ at the transition, we have $s_1 = (S_{11}, S_{12}, s_{21}^{(1)}, S_{22}), s_2 = (S_{11}, 0, s_{21}^{(2)}, S_{22})$.[1] By $\overline{s_1} = \overline{s_2}$ in (S.25), there must be $s_{21}^{(1)} < s_{21}^{(2)}$ and hence

$$\overline{s_1}'(n) = \left(\frac{\partial g(s)}{\partial s_{21}^{(1)}}\right)^{-1} < \left(\frac{\partial g(s)}{\partial s_{21}^{(2)}}\right)^{-1} = \overline{s_2}'(n),$$

where the inequality follows from $\partial^2 g(s)/\partial s_{21}^2 < 0$ given in (S.8). Therefore, (S.25) implies that the transition must be from pattern (1) to (2): $(S_{11}, S_{12}, s_{21}^{(1)}, S_{22}) \to (S_{11}, 0, s_{21}^{(2)}, S_{22})$. Obviously r_{12} jumps up and $\pi(n)$ jumps down at the transition.

If n is *large* such that $s_{21} = S_{21}$ and s_{12} is increasing (or just starts increasing from 0) at the transition, we have $s_1 = (S_{11}, S_{12}, s_{21}, S_{22}), s_2 = (S_{11}, s_{12}, S_{21}, S_{22})$. There must be $\overline{s_1}'(n^-) < \overline{s_2}'(n^+)$ since otherwise

$$\begin{aligned}\overline{s_1}(n_3^A) = \overline{s_1}(n) + \int_n^{n_3^A} \overline{s_1}'(x)dx > \overline{s_1}(n) + \int_n^{n_3^A} \overline{s_1}'(n)dx \\ \geq \overline{s_1}(n) + \int_n^{n_3^A} \overline{s_2}'(n^+)dx = \overline{s_2}(n) + \int_n^{n_3^A} \overline{s_2}'(x)dx = \overline{s_2}(n_3^A),\end{aligned}$$

where the first inequality follows from $\overline{s_1}'(x) = (\partial g(s)/\partial s_{21})^{-1}$ with $g(s) = x$ and $\partial^2 g(s)/\partial s_{21}^2 < 0$ given in (S.8), the second inequality is by the opposite assumption that $\overline{s_1}'(n) = \overline{s_1}'(n^-) \geq \overline{s_2}'(n^+)$, and the next equality follows from $\overline{s_1}(n) = \overline{s_2}(n)$ (at transition), $\overline{s_2}'(x) = (\partial g(s)/\partial s_{12})^{-1}$ with $g(s) = x$, and $\partial^2 g(s)/\partial s_{12}^2 = 0$ given in (S.8). Therefore, (S.25) implies that the transition must be from pattern (1) to (2): $(S_{11}, S_{12}, s_{21}, S_{22}) \to (S_{11}, s_{12}, S_{21}, S_{22})$. Similarly, r_{12} jumps up and $\pi(n)$ jumps down at the transition.

(ii) Between patterns (1) and (3), it is only optimal to transit from (1) to (3). Recall pattern (3) given in Lemma S-1, n can be small, medium or large at the transition. The case where n is small (such that s_{21} is increasing) or large (such that s_{12} is increasing) can be shown identically as above. For the case where n is medium, such that s_{11} is increasing (or just starts increasing from 0) while $s_{12} = 0, s_{21} = S_{21}$ at the transition, we have $s_1 = (S_{11}, S_{12}, s_{21}, S_{22}), s_3 = (s_{11}, 0, S_{21}, S_{22})$. There must be $\overline{s_1}'(n^-) < \overline{s_3}'(n^+)$ since otherwise

$$\begin{aligned}\overline{s_1}(n_3^A) = \overline{s_1}(n) + \int_n^{n_3^A} \overline{s_1}'(x)dx > \overline{s_1}(n) + \int_n^{n_3^A} \overline{s_1}'(n)dx \\ \geq \overline{s_1}(n) + \int_n^{n_3^A} \overline{s_3}'(n^+)dx = \overline{s_3}(n) + \int_n^{n_3^A} \overline{s_3}'(x)dx = \overline{s_3}(n_3^A),\end{aligned}$$

[1] We use superscript numbers to distinguish patterns under consideration.

where $\int_n^{n_3^A} \overline{s_3}'(x)dx$ is an integration from n to $\overline{s_3}^{-1}((S_{11}, 0, S_{21}, S_{22}))$ and then to n_3^A. The reasoning for the (in)equalities is analogous to that in above part (i). Therefore, (S.25) implies that the transition must be from pattern (1) to (3): $(S_{11}, S_{12}, s_{21}, S_{22}) \to (s_{11}, 0, S_{21}, S_{22})$. Apparently r_{12} jumps up and $\pi(n)$ jumps down at the transition.

(iii) Between patterns (2) and (3), it is only optimal to transit from (2) to (3). Recall patterns (2) and (3) given in Lemma S-1, the capacity allocations are the *same* in the case where n is large (such that $s_{11} = S_{11}, s_{21} = S_{21}$ while s_{12} is increasing), thus transitions must happen when n is *not* large, where $s_{12} \equiv 0$ is in common. Furthermore, when n is not large, there is only one case under pattern (2) where s_{21} is increasing, and there are two cases under pattern (3): small n where $s_{11} = 0$ and s_{21} is increasing, and medium n where $s_{21} = S_{21}$ and s_{11} is increasing. This shares a similar structure as the discussion between patterns (1) and (2) in part (i) above, so we omit the details here. Similarly, note that the transitions, if any, are always from pattern (2) to (3).

□

Proof of Proposition 5. From Lemma S-1, it is optimal to reject rider requests at the low-demand location for some n in zone (3) if and only if pattern (2) or (3) provides the *largest* service capacity at some $n \in (n_2^A, n_3^A)$, i.e., $\exists n \in (n_2^A, n_3^A)$ such that $\overline{s_1}(n) < \max_{i\in\{2,3\}} \overline{s_i}(n)$. We need to compare the three patterns in terms of their service capacity $\overline{s_i}(n)$, $i = 1, 2, 3$. We have the following three observations.

(i) At the right end of zone (3), $\overline{s_1}(n_3^A) = \overline{s_2}(n_3^A) = \overline{s_3}(n_3^A) = \overline{S}$.

(ii) For n close to n_3^A ($n \to n_3^{A-}$), it follows from Lemma S-1 that pattern (1) has $s_1(n) = (S_{11}, S_{12}, s_{21}, S_{22})$ with s_{21} varying, while pattern (2) and (3) both have $s_2(n) = s_3(n) = (S_{11}, s_{12}, S_{21}, S_{22})$ with s_{12} varying. Therefore we have

$$\overline{s_1}'_-(n_3^A) = \frac{\partial \overline{s}/\partial s_{21}}{\partial g(s)/\partial s_{21}}\bigg|_{s=S} = \left(\frac{\partial g(s)}{\partial s_{21}}\bigg|_{s=S} \right)^{-1} > 0,$$
$$\overline{s_1}''_-(n_3^A) = \frac{\partial \overline{s_1}'_-(n_3^A)/\partial s_{21}}{\partial g(s)/\partial s_{21}}\bigg|_{s=S} = -\frac{\partial^2 g(s)/\partial s_{21}^2}{(\partial g(s)/\partial s_{21})^3}\bigg|_{s=S} > 0,$$

i.e., $\overline{s_1}(n)$ is strictly convex and increasing in n near n_3^A. And

$$\overline{s_2}'_-(n_3^A) = \overline{s_3}'_-(n_3^A) = \frac{\partial \overline{s}/\partial s_{12}}{\partial g(s)/\partial s_{12}}\bigg|_{s=S} = \left(\frac{\partial g(s)}{\partial s_{12}}\bigg|_{s=S} \right)^{-1},$$
$$\overline{s_2}''_-(n_3^A) = \overline{s_3}''_-(n_3^A) = \frac{\partial \overline{s_2}'_-(n_3^A)/\partial s_{12}}{\partial g(s)/\partial s_{12}}\bigg|_{s=S} = -\frac{\partial^2 g(s)/\partial s_{12}^2}{(\partial g(s)/\partial s_{12})^3}\bigg|_{s=S} = 0,$$

i.e., $\overline{s_2}(n)$ and $\overline{s_3}(n)$ both increase linearly in n near n_3^A.

(iii) The proof of Lemma S-2 establishes that any optimal pattern transitions as n *increases* in zone (3) must be from pattern (1) to (2), from pattern (1) to (3), or from pattern (2) to (3)—not vice versa.

Using the above observations, the sufficient and necessary condition that pattern (2) or (3) provides the *largest* service capacity at some $n \in (n_2^A, n_3^A)$ is

$$\overline{s_2}'_-(n_3^A) = \overline{s_3}'_-(n_3^A) < \overline{s_1}'_-(n_3^A). \tag{S.26}$$

To see this, if (S.26) holds, observation (i) and (ii) immediately imply that $\overline{s_2}(n_3^{A-}) = \overline{s_3}(n_3^{A-}) > \overline{s_1}(n_3^{A-})$, hence patterns (2) and (3) provide the (same) largest service capacity close to n_3^A. On the other hand, if

(S.26) does not hold, observation (i) and (ii) imply that $\overline{s_2}(n_3^{A-}) = \overline{s_3}(n_3^{A-}) < \overline{s_1}(n_3^{A-})$, i.e., pattern (1) is optimal near n_3^A. It then follows from observation (iii) that there is no transition from pattern (2) or (3) to pattern (1) as n increases in zone (3), hence pattern (1) is optimal throughout zone (3). Therefore (S.26) is sufficient and necessary for pattern (2) or (3) to be optimal somewhere in zone (3).

We now translate (S.26) to condition (36) in the proposition. Putting the first derivatives in observation (ii) into (S.26), we get $\left(\frac{\partial g(s)}{\partial s_{12}}\Big|_{s=S}\right)^{-1} < \left(\frac{\partial g(s)}{\partial s_{21}}\Big|_{s=S}\right)^{-1}$. Using (S.5) and (S.6) in the proof of Lemma S-1, this becomes

$$\frac{S_{21}(1+\frac{t_{12}}{t_{21}})+S_{22}}{(S_{21}+S_{22})\bar{\gamma}p-\left(S_{21}+S_{22}+S_{21}\frac{t_{12}}{t_{21}}\right)c}\bar{\gamma}p > 1+\frac{t_{12}}{t_{21}}+\frac{(S_{11}(\bar{\gamma}p-c)+S_{12}\bar{\gamma}p)S_{22}\frac{t_{12}}{t_{21}}}{\left[(S_{21}+S_{22})\bar{\gamma}p-\left(S_{21}+S_{22}+S_{21}\frac{t_{12}}{t_{21}}\right)c\right]^2}\bar{\gamma}p.$$

Applying the ratios defined in (35) and with algebraic rearrangement, we get inequality (36). □

S2 Extension to General Networks

In this section we extend the model to general networks and provide additional numerical results for the three-location networks in Figure 5.

S2.1 Model Primitives

Consider a general L-location network with location (node) set $V = \{1, \ldots, L\}$ and route (arc) set $V \times V$. Denote by $t \in \mathbb{R}_+^{L\times L}$ the (constant) travel time matrix and by $\Lambda \in \mathbb{R}_+^{L\times L}$ the potential demand rate matrix (given rider price p per unit of travel time). Drivers' profit and participation have the same structure as described in §2.1. As direct extensions from the two-location network, denote by $\lambda \in \mathbb{R}_+^{L\times L}$ ($\lambda \le \Lambda$) the *effective* demand rate matrix, $\nu \in \mathbb{R}_+^{L\times L}$ (with zero diagonal elements) the repositioning rate matrix, and $w, q \in \mathbb{R}_+^{L}$ the steady-state waiting time and queue length vectors, respectively. Denoting $\eta(\lambda,\nu) \in H := \{\eta \in \mathbb{R}_+^{L\times L} : \eta\mathbf{1} = \mathbf{1}\}$ as the matrix of steady-state repositioning fractions resulting from λ and ν, we have

$$\eta_{ij}(\lambda,\nu) = \begin{cases} \frac{\sum_{k\in V}\lambda_{ik}}{\sum_{k\in V}(\lambda_{ik}+\nu_{ik})} & j=i \\ \frac{\nu_{ij}}{\sum_{k\in V}(\lambda_{ik}+\nu_{ik})} & j\ne i, \end{cases} \quad i,j\in V. \tag{S.27}$$

The steady-state system constraints include: (i) flow balance at each location (outflows equal inflows), $\sum_{i\in V}(\lambda_{ij}+\nu_{ij}) = \sum_{k\in V}(\lambda_{jk}+\nu_{jk}), \forall j \in V$, and (ii) total capacity $\sum_{i,j\in V}(\lambda_{ij}+\nu_{ij})t_{ij}+\sum_{i\in V}[w_i\sum_{j\in V}\lambda_{ij}] = n$.

Similar to the two-location network model, per-driver profit rate can be computed in two ways: (i) as the per-driver proportion of the cumulative driver profits:

$$\pi(\lambda,\nu,n) = \frac{(\bar{\gamma}p-c)\sum_{i,j\in V}\lambda_{ij}t_{ij} - c\sum_{i,j\in V}\nu_{ij}t_{ij}}{n},$$

where the number of participating drivers n satisfies the participation equilibrium $n = NF(\pi(\lambda,\nu,n))$, and (ii) from the perspective of an individual driver circulating through the network, $\widetilde{\pi}(\widetilde{\eta};\lambda,w)$, as a function of her own repositioning strategy (fractions) $\widetilde{\eta}$ given the routing probabilities implied by λ and the queueing delays w. In equilibrium, the repositioning fractions $\eta(\lambda,\nu)$ induced by the aggregate flow rates (λ,ν) through (S.27) maximizes the individual drivers' profit rate, i.e.,

$$\eta(\lambda,\nu) \in \operatorname*{argmax}_{\widetilde{\eta}\in H} \widetilde{\pi}(\widetilde{\eta};\lambda,w). \tag{S.28}$$

Since every driver chooses $\eta(\lambda, \nu)$, each earns the same profit rate, so that $\widetilde{\pi}(\eta(\lambda, \nu); \lambda, w) = \pi(\lambda, \nu, n)$ for all (λ, ν, w, n) tuples that satisfy the system flow constraints described above.

Reformulating the Repositioning Equilibrium Constraint (S.28). Let $\mathcal{E}^{ij}$ be the matrix operator that replaces a matrix's ith row by the basis vector e_j, e.g., $\mathcal{E}^{ij}\widetilde{\eta}$ denotes the altered repositioning strategy $\widetilde{\eta}$ where the driver chooses a *pure strategy* at location i that repositions only to j. At any location i, let $\mathcal{R}_i(\widetilde{\eta}) \coloneqq \{j \in V : \widetilde{\eta}_{ij} > 0\} \subseteq V$ denote the subset of locations that are assigned positive probabilities by repositioning strategy $\widetilde{\eta}$, so $|\mathcal{R}_i(\widetilde{\eta})| = 1$ is a pure strategy and $|\mathcal{R}_i(\widetilde{\eta})| > 1$ is a mixed strategy. Since all locations in $\mathcal{R}_i(\widetilde{\eta})$, if adopted as the only repositioning destination under pure strategies, must yield *equal* profit rate that is *higher* than locations outside $\mathcal{R}_i(\widetilde{\eta})$, we have

$$\widetilde{\pi}(\mathcal{E}^{ij}\widetilde{\eta}; \lambda, w) = \widetilde{\pi}(\mathcal{E}^{ik}\widetilde{\eta}; \lambda, w) \geq \widetilde{\pi}(\mathcal{E}^{il}\widetilde{\eta}; \lambda, w), \quad \forall j, k \in \mathcal{R}_i(\widetilde{\eta}),\ l \notin \mathcal{R}_i(\widetilde{\eta}).$$

Therefore, the repositioning equilibrium constraint (S.28) is equivalent to the following constraints (S.29)–(S.31), where η represents the repositioning fractions $\eta(\lambda, \nu)$ determined in (S.27).

$$\widetilde{\pi}(\mathcal{E}^{ij}\eta; \lambda, w) = \xi_i - \zeta_{ij},\ \forall i, j \in V, \tag{S.29}$$

$$\zeta_{ij}\eta_{ij} = 0,\ \forall i, j \in V, \tag{S.30}$$

$$\xi \in \mathbb{R}^L,\ \zeta \in \mathbb{R}_+^{L\times L},\ \eta \in \mathbb{R}_+^{L\times L}. \tag{S.31}$$

S2.2 Steady-State Per-Driver Profit Rate

Given the steady-state system characterized by (λ, w), we can formulate an individual driver's location visiting process as a Semi-Markov Process (SMP), where the *state* is the latest location (node) the driver has visited and a *transition* occurs when the driver arrives at a location, before making a repositioning decision on whether to join the (potential) queue or reposition to another location. Then the driver's cumulative profit process is a Markov Renewal-Reward Process $\{\widetilde{\pi}(t)\}$ described by a sequence $\{(Y_k, X_k, W_k)\}_{k\in\mathbb{N}}$ as $\widetilde{\pi}(\sum_{i=1}^k X_i) = \sum_{i=1}^k W_i$, where state Y_k is the location after the kth transition, X_k is the sojourn time between the $(k-1)$th and kth transition (which includes potential queueing delay and travel time in service or in repositioning), and reward W_k is the driver profit collected between the $(k-1)$th and kth transition (which includes potential service revenue and driving cost).

The transition probability matrix of the embedded DTMC $\{Y_k\}$ is a function of the driver's repositioning strategy $\widetilde{\eta}$ given the routing probabilities implied by λ, $P(\widetilde{\eta}; \lambda)$, with elements

$$P_{ij}(\widetilde{\eta}; \lambda) = \begin{cases} \widetilde{\eta}_{ii}\frac{\lambda_{ii}}{\sum_{k\in V}\lambda_{ik}} & \text{if } j = i \\ \widetilde{\eta}_{ij} + \widetilde{\eta}_{ii}\frac{\lambda_{ij}}{\sum_{k\in V}\lambda_{ik}} & \text{if } j \neq i, \end{cases} \quad i, j \in V,\ \sum_{k\in V}\lambda_{ik} > 0. \tag{S.32}$$

If $\sum_{k\in V}\lambda_{ik} = 0$ for some i, then $P_{ii}(\widetilde{\eta}; \lambda) = \mathbf{1}(\widetilde{\eta}_{ii} > 0)$ and $P_{ij}(\widetilde{\eta}; \lambda) = \widetilde{\eta}_{ij}\mathbf{1}(\widetilde{\eta}_{ii} = 0)$ for $j \neq i$. In the case where $\sum_{k\in V}\lambda_{ik} = 0$ and $\widetilde{\eta}_{ii} > 0$, state i is absorbing. We assume $\sum_{k\in V}\lambda_{ik} > 0$ for $i \in V$ so that the Markov chain is irreducible. Let $p(\widetilde{\eta}; \lambda) \in \Delta^{L-1}$ be the associated stationary distribution.

The expected reward (driver profit) after a transition into location i is then given by

$$R_i(\widetilde{\eta}; \lambda) \coloneqq \mathbb{E}(W_{k+1} \mid Y_k = i) = (\bar{\gamma}p - c)\widetilde{\eta}_{ii}\sum_{j\in V}\frac{\lambda_{ij}}{\sum_{k\in V}\lambda_{ik}}t_{ij} - c\sum_{j\in V\setminus\{i\}}\widetilde{\eta}_{ij}t_{ij}, \quad i \in V, \tag{S.33}$$

and the expected sojourn time after a transition into location i is

$$T_i(\widetilde{\eta}; \lambda, w) \coloneqq \mathbb{E}(X_{k+1} \mid Y_k = i) = \widetilde{\eta}_{ii}\left[w_i + \sum_{j\in V}\frac{\lambda_{ij}}{\sum_{k\in V}\lambda_{ik}}t_{ij}\right] + \sum_{j\in V\setminus\{i\}}\widetilde{\eta}_{ij}t_{ij}, \quad i \in V. \tag{S.34}$$

Let $R(\widetilde{\eta};\lambda) \in \mathbb{R}^L$ and $T(\widetilde{\eta};\lambda,w) \in \mathbb{R}^L_+$ be the corresponding vectors, respectively.[2] The following proposition gives the steady-state driver profit rate by the renewal reward theorem.

Proposition 1. *An individual driver's expected steady-state profit rate is a function of her repositioning strategy $\widetilde{\eta}$ for system state (λ, w), given by*

$$\widetilde{\pi}(\widetilde{\eta};\lambda,w) := \lim_{t\to\infty} \frac{\widetilde{\pi}(t)}{t} = \frac{\Pi_1}{\tau_1} = \frac{e_1^T[I - P^\circ(\widetilde{\eta};\lambda)]^{-1}R(\widetilde{\eta};\lambda)}{e_1^T[I - P^\circ(\widetilde{\eta};\lambda)]^{-1}T(\widetilde{\eta};\lambda,w)}, \tag{S.35}$$

where Π_1 and τ_1 are the expected cycle profit and cycle length, respectively, with cycles defined as consecutive arrivals at location 1 (before joining the queue or repositioning), and $P^\circ = [0, P_{*,2}, \cdots P_{*,L}]$ is the $P(\widetilde{\eta};\lambda)$ matrix with first column replaced by 0.

Proof. Let $\Pi_i, i \in V$ be the expected profit collected by a driver starting from location i (before joining the queue or repositioning) and ending at coming back to location 1, and Π the corresponding vector. We have

$$\Pi_1 = \widetilde{\eta}_{11} \sum_{j\in V} \frac{\lambda_{1j}}{\sum_{k\in V}\lambda_{1k}}[(\bar{\gamma}p - c)t_{1j} + \Pi_j \mathbf{1}(j \neq 1)] + \sum_{j\in V\backslash\{1\}} \widetilde{\eta}_{1j}(-ct_{1j} + \Pi_j),$$

$$\Pi_j = \widetilde{\eta}_{jj} \sum_{l\in V} \frac{\lambda_{jl}}{\sum_{k\in V}\lambda_{jk}}[(\bar{\gamma}p - c)t_{jl} + \Pi_l \mathbf{1}(l \neq 1)] + \sum_{l\in V\backslash\{j\}} \widetilde{\eta}_{jl}\left(-ct_{jl} + \Pi_l \mathbf{1}(l \neq 1)\right), \quad j \in V\backslash\{1\}.$$

In vector form and using (S.32) and (S.33), we can derive

$$\Pi = P^\circ(\widetilde{\eta};\lambda)\Pi + R(\widetilde{\eta};\lambda) \Rightarrow \Pi = [I - P^\circ(\widetilde{\eta};\lambda)]^{-1}R(\widetilde{\eta};\lambda).$$

Let $\tau_i, i \in V$ be the expected duration starting from location i (before joining the queue or repositioning) and ending at coming back to location 1, and τ the corresponding vector. We have

$$\tau_1 = \widetilde{\eta}_{11}\left[w_1 + \sum_{j\in V} \frac{\lambda_{1j}}{\sum_{k\in V}\lambda_{1k}}(t_{1j} + \tau_j \mathbf{1}(j \neq 1))\right] + \sum_{j\in V\backslash\{1\}} \widetilde{\eta}_{1j}(t_{1j} + \tau_j),$$

$$\tau_j = \widetilde{\eta}_{jj}\left[w_j + \sum_{l\in V} \frac{\lambda_{jl}}{\sum_{k\in V}\lambda_{jk}}(t_{jl} + \tau_l \mathbf{1}(l \neq 1))\right] + \sum_{l\in V\backslash\{j\}} \widetilde{\eta}_{jl}\left(t_{jl} + \tau_l \mathbf{1}(l \neq 1)\right), \quad j \in V\backslash\{1\}.$$

In vector form and using (S.32) and (S.34), we can derive

$$\tau = P^\circ(\widetilde{\eta};\lambda)\tau + T(\widetilde{\eta};\lambda,w) \Rightarrow \tau = [I - P^\circ(\widetilde{\eta};\lambda)]^{-1}T(\widetilde{\eta};\lambda,w).$$

It follows from renewal reward theory that

$$\widetilde{\pi}(\widetilde{\eta};\lambda,w) = \frac{\Pi_1}{\tau_1} = \frac{e_1^T[I - P^\circ(\widetilde{\eta};\lambda)]^{-1}R(\widetilde{\eta};\lambda)}{e_1^T[I - P^\circ(\widetilde{\eta};\lambda)]^{-1}T(\widetilde{\eta};\lambda,w)}.$$

□

S2.3 Three Control Regimes: Problem Formulations

We now formulate the platform's revenue maximization problem under the three control regimes.

[2] The above quantities can be expressed in compact matrix/vector form as

$$P(\widetilde{\eta};\lambda) = \mathbf{diag}(D(\widetilde{\eta}))\left[\text{diag}(\lambda\mathbf{1})^{-1}\lambda\right] + \widetilde{\eta} - \mathbf{diag}(D(\widetilde{\eta})),$$

$$R(\widetilde{\eta};\lambda) = (\bar{\gamma}p - c)\mathbf{diag}(D(\widetilde{\eta}))\left[\mathbf{diag}(\lambda\mathbf{1})^{-1}((\lambda \circ t)\mathbf{1})\right] - c\left[\left(\widetilde{\eta} - \mathbf{diag}(D(\widetilde{\eta}))\right) \circ t\right]\mathbf{1},$$

$$T(\widetilde{\eta};\lambda,w) = \mathbf{diag}(D(\widetilde{\eta}))\left[w + \mathbf{diag}(\lambda\mathbf{1})^{-1}((\lambda \circ t)\mathbf{1})\right] + \left[\left(\widetilde{\eta} - \mathbf{diag}(D(\widetilde{\eta}))\right) \circ t\right]\mathbf{1},$$

where $\mathbf{diag}(a)$ generates a diagonal matrix from vector a, and $D(A)$ generates a vector from the diagonal elements of matrix A.

Centralized Control (C). This benchmark regime can be formulated as the following optimization problem.

$$\text{(Problem C)} \quad \max_{\lambda,\nu,w,n} \quad \mathit{\Pi}(\lambda) := \gamma p \sum_{i,j \in V} \lambda_{ij} t_{it} \tag{S.36a}$$

$$\text{s.t.} \quad \sum_{i \in V} (\lambda_{ij} + \nu_{ij}) = \sum_{k \in V} (\lambda_{jk} + \nu_{jk}), \quad \forall j \in V, \tag{S.36b}$$

$$\sum_{i,j \in V} (\lambda_{ij} + \nu_{ij}) t_{ij} + \sum_{i \in V} [w_i \sum_{j \in V} \lambda_{ij}] = n, \tag{S.36c}$$

$$0 \le \lambda \le \mathit{\Lambda}, \ \nu \ge 0, \ w \ge 0, \tag{S.36d}$$

$$\pi(\lambda, \nu, n) = \frac{(\bar{\gamma} p - c) \sum_{i,j \in V} \lambda_{ij} t_{ij} - c \sum_{i,j \in V} \nu_{ij} t_{ij}}{n}, \tag{S.36e}$$

$$n = NF(\pi(\lambda, \nu, n)). \tag{S.36f}$$

For fixed capacity n the problem for regime C is a simple LP given by

$$\mathit{\Pi}_C(n) = \max_{\lambda,\nu,w} \{\mathit{\Pi}(\lambda) : \text{(S.36b)–(S.36d)}\}. \tag{S.37}$$

Admission Control (A). Representing the repositioning equilibrium constraint (S.28) using (S.29)–(S.31) together with (S.27) and (S.35), the platform's problem can be formulated as an MPEC (Mathematical Program with Equilibrium Constraints):

$$\text{(Problem A)} \quad \max_{\lambda,\nu,w,n,\xi,\zeta} \{\mathit{\Pi}(\lambda) : \text{(S.27), (S.29)–(S.31), (S.35), (S.36b)–(S.36f)}\}. \tag{S.38}$$

For fixed capacity n the problem for regime A is a nonlinear problem given by

$$\mathit{\Pi}_A(n) = \max_{\lambda,\nu,w,n,\xi,\zeta} \{\mathit{\Pi}(\lambda) : \text{(S.27), (S.29)–(S.31), (S.35), (S.36b)–(S.36d)}\}. \tag{S.39}$$

Minimal Control (M). Under pro-rata (FIFO) matching, we need the following additional admission constraints: The effective demand rates are proportional to the potential demand at each location:

$$\lambda_{ij} = k_i \mathit{\Lambda}_{ij}, \ 0 \le k_i \le 1, \quad \forall i, j \in V, \tag{S.40}$$

where k_i is the service rate at location i. Drivers cannot be repositioning out of location i if the potential rider demand at that location has not been *fully* served, i.e.,

$$(1 - k_i)\nu_{ij} = 0, \quad \forall i, j \in V, \tag{S.41}$$

and demand requests originating at location i can only be lost if this location has no supply buffer, so no drivers are waiting, i.e.,

$$(1 - k_i) w_i = 0, \quad \forall i \in V. \tag{S.42}$$

Under this regime, the platform's problem can be formulated as:

$$\text{(Problem M)} \quad \max_{\lambda,\nu,w,n,\xi,\zeta} \{\mathit{\Pi}(\lambda) : \text{(S.27), (S.29)–(S.31), (S.35), (S.36b)–(S.36f), (S.40)–(S.42)}\}. \tag{S.43}$$

For fixed capacity n the problem for regime M is the nonlinear problem

$$\mathit{\Pi}_M(n) = \max_{\lambda,\nu,w,n,\xi,\zeta} \{\mathit{\Pi}(\lambda) : \text{(S.27), (S.29)–(S.31), (S.35), (S.36b)–(S.36d), (S.40)–(S.42)}\}. \tag{S.44}$$

S2.4 Additional Numerical Results for Networks in Figure 5

In §6 we introduced several types of three-location ring network in Figure 5 and discussed main findings from network I under Admission Control. Here we present detailed capacity allocation for network I under Minimal Control (Table 1), as well as for networks II and III under Minimal Control and Admission Control (Tables 2 to 5). Note the common setting of unit travel times, rider price $p = 4$, commission rate $\gamma = 25\%$ and driving cost $c = 1$.

n	service	*k* 1	*k* 2	*k* 3	*r* 12	*r* 13	*q* 1	*q* 2	*q* 3	*η* 11	*η* 12	*η* 13
6	6.00	0.90	0.33	0.41	0.00	0.00	0.00	0.00	0.00	100%	0%	0%
7	6.64	1.00	0.36	0.45	0.00	0.00	0.36	0.00	0.00	100%	0%	0%
8	6.64	1.00	0.36	0.45	0.00	0.00	1.36	0.00	0.00	100%	0%	0%
9	6.64	1.00	0.36	0.45	0.00	0.00	2.36	0.00	0.00	100%	0%	0%
10	6.64	1.00	0.36	0.45	0.00	0.00	3.36	0.00	0.00	100%	0%	0%
11	6.64	1.00	0.36	0.45	0.00	0.00	4.36	0.00	0.00	100%	0%	0%
12	6.64	1.00	0.36	0.45	0.00	0.00	5.36	0.00	0.00	100%	0%	0%
13	6.64	1.00	0.36	0.45	0.00	0.00	6.36	0.00	0.00	100%	0%	0%
14	6.87	1.00	0.37	0.50	0.00	0.12	7.01	0.00	0.00	96%	0%	4%
15	7.49	1.00	0.40	0.62	0.00	0.45	7.06	0.00	0.00	87%	0%	13%
16	8.15	1.00	0.44	0.74	0.00	0.79	7.06	0.00	0.00	79%	0%	21%
17	8.80	1.00	0.47	0.87	0.00	1.13	7.06	0.00	0.00	73%	0%	27%
18	9.45	1.00	0.50	0.99	0.00	1.47	7.07	0.00	0.01	67%	0%	33%
19	9.85	1.00	0.68	0.87	0.83	0.94	7.30	0.00	0.06	63%	17%	20%
20	10.31	1.00	0.70	0.96	0.82	1.19	7.44	0.00	0.23	60%	16%	24%
21	10.66	1.00	0.74	0.99	0.97	1.23	7.68	0.00	0.46	58%	19%	24%
22	11.08	1.00	0.82	0.99	1.29	1.16	7.86	0.00	0.61	55%	24%	21%
23	11.50	1.00	0.90	1.00	1.60	1.09	8.04	0.00	0.76	53%	28%	19%
24	12.00	1.00	1.00	1.00	2.00	1.00	8.14	0.00	0.86	50%	33%	17%
25	12.00	1.00	1.00	1.00	2.00	1.00	8.52	0.33	1.14	50%	33%	17%

Table 1: Optimal capacity allocation for network I under Minimal Control (M)

n	service	*k* 1	*k* 2	*k* 3	*r* 21	*r* 23	*q* 1	*q* 2	*q* 3	*η* 21	*η* 22	*η* 23
7	7.00	0.67	1.00	0.33	0.00	0.00	0.00	0.00	0.00	0%	100%	0%
8	7.00	0.67	1.00	0.33	0.00	0.00	0.00	1.00	0.00	0%	100%	0%
9	7.00	0.67	1.00	0.33	0.00	0.00	0.00	2.00	0.00	0%	100%	0%
10	7.00	0.67	1.00	0.33	0.00	0.00	0.00	3.00	0.00	0%	100%	0%
11	7.00	0.67	1.00	0.33	0.00	0.00	0.00	4.00	0.00	0%	100%	0%
12	7.00	0.67	1.00	0.33	0.00	0.00	0.00	5.00	0.00	0%	100%	0%
13	7.19	0.71	1.00	0.34	0.08	0.00	0.00	5.72	0.00	3%	97%	0%
14	7.89	0.88	1.00	0.38	0.38	0.00	0.00	5.73	0.00	11%	89%	0%
15	8.40	1.00	1.00	0.40	0.60	0.00	0.11	5.89	0.00	17%	83%	0%
16	8.40	1.00	1.00	0.40	0.60	0.00	0.52	6.48	0.00	17%	83%	0%
17	8.40	1.00	1.00	0.40	0.60	0.00	0.93	7.07	0.00	17%	83%	0%
18	8.40	1.00	1.00	0.40	0.60	0.00	1.33	7.67	0.00	17%	83%	0%
19	8.40	1.00	1.00	0.40	0.60	0.00	1.74	8.26	0.00	17%	83%	0%
20	8.47	1.00	1.00	0.41	0.59	0.06	2.10	8.78	0.00	16%	82%	2%
21	8.86	1.00	1.00	0.48	0.52	0.38	2.24	8.99	0.00	13%	77%	10%
22	9.45	1.00	1.00	0.58	0.42	0.88	2.25	9.00	0.00	10%	70%	20%
23	10.05	1.00	1.00	0.67	0.33	1.37	2.25	9.00	0.00	7%	64%	29%
24	10.65	1.00	1.00	0.78	0.22	1.88	2.25	9.00	0.00	4%	59%	37%
25	11.25	1.00	1.00	0.88	0.12	2.38	2.25	9.00	0.00	2%	55%	43%
26	11.85	1.00	1.00	0.97	0.03	2.87	2.25	9.00	0.00	0%	51%	49%
27	12.00	1.00	1.00	1.00	0.00	3.00	2.43	9.29	0.29	0%	50%	50%
28	12.00	1.00	1.00	1.00		3.00	2.96	9.52	0.52	0%	50%	50%

Table 2: Optimal capacity allocation for network II under Minimal Control (M)

n	service	*s* 11	*s* 22	*s* 32	*r* 23	*q* 1	*q* 2	*q* 3	*η* 22	*η* 23
10	9.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00	100%	0%
11	9.00	1.00	1.00	1.00	0.00	0.00	2.00	0.00	100%	0%
12	9.00	1.00	1.00	1.00	0.00	0.00	3.00	0.00	100%	0%
13	9.00	1.00	1.00	1.00	0.00	0.00	4.00	0.00	100%	0%
14	9.00	1.00	1.00	1.00	0.00	0.00	5.00	0.00	100%	0%
15	9.12	1.00	1.00	1.12	0.12	0.14	5.62	0.00	96%	4%
16	9.32	1.00	1.00	1.32	0.32	0.37	5.98	0.00	90%	10%
17	9.54	1.00	1.00	1.54	0.54	0.58	6.34	0.00	85%	15%
18	9.77	1.00	1.00	1.77	0.77	0.80	6.67	0.00	80%	20%
19	10.00	1.00	1.00	2.00	1.00	1.00	7.00	0.00	75%	25%
20	10.25	1.00	1.00	2.25	1.25	1.20	7.31	0.00	71%	29%
21	10.50	1.00	1.00	2.50	1.50	1.38	7.62	0.00	67%	33%
22	10.76	1.00	1.00	2.76	1.76	1.57	7.90	0.00	63%	37%
23	11.04	1.00	1.00	3.04	2.04	1.74	8.18	0.00	60%	40%
24	11.32	1.00	1.00	3.32	2.32	1.90	8.45	0.00	56%	44%
25	11.62	1.00	1.00	3.62	2.62	2.06	8.70	0.00	53%	47%
26	11.92	1.00	1.00	3.92	2.92	2.21	8.94	0.00	51%	49%
27	12.00	1.00	1.00	4.00	3.00	2.65	9.17	0.17	50%	50%
28	12.00	1.00	1.00	4.00	3.00	2.96	9.52	0.52	50%	50%

Table 3: Optimal capacity allocation for network II under Admission Control (A)

S3 Driver Supply and Actual Gains in Platform Revenue and Per-Driver Profit

In this section we illustrate the impact of the driver supply characteristics, specifically, the outside opportunity cost distribution F, on the *actual* platform revenue and per-driver profit gains, compared to the upper bounds in Propositions 8 and 9, and on the tension between the drivers' and the platform's gains. For simplicity we focus on the gains from admission control, i.e., regime A over M. (Similar effects determine the actual gains from repositioning.)

Figure 1 illustrates these gains for two opportunity cost distributions. Panel (a) presents a case where admission control yields large benefits for the platform as a result of a large increase in driver participation, and consequently only small benefits for individual drivers. Specifically, the top chart in panel (a) shows for the three control regimes the per-driver profits that are non-increasing functions of the capacity, and the increasing marginal opportunity cost function $F^{-1}(n/N)$. Achieving the upper bound on platform revenue gains from admission control requires two conditions, namely, $n_M^* = n_2^M$ or equivalently, $F^{-1}(n_2^M/N) = \pi_M(n_2^M)$, and $n_A^* = n_3^A$. The first condition holds in the example, the second condition requires infinitely elastic supply around the profit level $\pi_M(n_2^M)$, i.e., that F grows sufficiently fast around this point such that $n_3^A - n_2^M$ additional drivers join if the per-driver profit is slightly larger, so that $F^{-1}(n_3^A/N) = \pi_M(n_3^A)$. The example depicted in Figure 1 (a) shows how the upper bound can be approached if the supply increases substantially for a moderate change in per-driver profit rate.

Panel (b), in contrast, presents a case where admission control (under regime A or C) yields the maximum achievable per-driver profit gains as a result of a small increase in driver participation, and consequently only modest platform revenue gains. As shown in the top chart of panel (b), in this case the marginal opportunity cost function yields the same equilibrium capacity under minimal control as in panel (a), i.e., $F^{-1}(n_2^M/N) = \pi_M(n_2^M)$; however, the driver supply is so inelastic that the number of drivers willing to participate at the maximum profit rate $(\overline{\gamma}p - c)$ is smaller than the minimum number required to serve all riders without repositioning, that is, $n_A^* \leq n_1^A$ where $F^{-1}(n_A^*/N) = \overline{\gamma}p - c$. The platform's commission is too high to entice more drivers to participate.

n	service	$k1$	$k2$	$k3$	$r21$	$r23$	$r31$	$q1$	$q2$	$q3$	$\eta 21$	$\eta 22$	$\eta 23$	$\eta 31$	$\eta 33$
7	7.00	0.33	0.89	0.78	0.00	0.00	0.00	0.00	0.00	0.00	0%	100%	0%	0%	100%
8	7.87	0.37	1.00	0.87	0.00	0.00	0.00	0.00	0.13	0.00	0%	100%	0%	0%	100%
9	7.88	0.38	1.00	0.88	0.00	0.00	0.00	0.00	1.12	0.00	0%	100%	0%	0%	100%
10	7.88	0.38	1.00	0.88	0.00	0.00	0.00	0.00	2.12	0.00	0%	100%	0%	0%	100%
11	7.88	0.38	1.00	0.88	0.00	0.00	0.00	0.00	3.12	0.00	0%	100%	0%	0%	100%
12	7.88	0.38	1.00	0.88	0.00	0.00	0.00	0.00	4.12	0.00	0%	100%	0%	0%	100%
13	7.88	0.38	1.00	0.88	0.00	0.00	0.00	0.00	5.12	0.00	0%	100%	0%	0%	100%
14	8.31	0.40	1.00	0.98	0.00	0.17	0.00	0.00	5.53	0.00	0%	95%	5%	0%	100%
15	8.43	0.41	1.00	1.00	0.03	0.19	0.00	0.00	6.00	0.36	1%	93%	6%	0%	100%
16	8.50	0.42	1.00	1.00	0.08	0.17	0.00	0.00	6.56	0.69	3%	92%	5%	0%	100%
17	8.87	0.48	1.00	1.00	0.39	0.04	0.00	0.00	6.82	0.88	11%	87%	1%	0%	100%
18	9.00	0.50	1.00	1.00	0.50	0.00	0.00	0.00	7.10	1.40	14%	86%	0%	0%	100%
19	9.00	0.50	1.00	1.00	0.50	0.00	0.00	0.00	7.30	2.20	14%	86%	0%	0%	100%
20	9.00	0.50	1.00	1.00	0.50	0.00	0.00	0.00	7.50	3.00	14%	86%	0%	0%	100%
21	9.01	0.50	1.00	1.00	0.50	0.00	0.00	0.00	7.70	3.78	14%	86%	0%	0%	100%
22	9.00	0.50	1.00	1.00	0.50	0.00	0.00	0.00	7.90	4.60	14%	86%	0%	0%	100%
23	9.02	0.50	1.00	1.00	0.51	0.00	0.00	0.00	8.10	5.37	14%	86%	0%	0%	100%
24	9.20	0.53	1.00	1.00	0.60	0.00	0.06	0.00	8.18	5.96	17%	83%	0%	2%	98%
25	9.38	0.56	1.00	1.00	0.70	0.00	0.12	0.00	8.30	6.50	19%	81%	0%	4%	96%
26	9.53	0.59	1.00	1.00	0.77	0.00	0.17	0.00	8.45	7.07	20%	80%	0%	5%	94%
27	9.75	0.62	1.00	1.00	0.88	0.00	0.25	0.00	8.55	7.58	23%	77%	0%	8%	92%
28	10.06	0.68	1.00	1.00	1.03	0.00	0.35	0.00	8.61	7.95	26%	74%	0%	10%	89%
29	10.46	0.74	1.00	1.00	1.23	0.00	0.48	0.00	8.66	8.17	29%	71%	0%	14%	86%
30	10.90	0.82	1.00	1.00	1.45	0.00	0.63	0.00	8.69	8.34	33%	67%	0%	17%	83%
31	11.18	0.86	1.00	1.00	1.59	0.00	0.73	0.00	8.83	8.66	35%	65%	0%	20%	80%
32	11.66	0.94	1.00	1.00	1.83	0.00	0.89	0.00	8.88	8.74	38%	62%	0%	23%	77%
33	12.00	1.00	1.00	1.00	2.00	0.00	1.00	0.00	9.00	9.00	40%	60%	0%	25%	75%
34	12.00	1.00	1.00	1.00	2.00	0.00	1.00	0.33	9.33	9.33	40%	60%	0%	25%	75%
35	12.00	1.00	1.00	1.00	2.00	0.00	1.00	0.67	9.67	9.67	40%	60%	0%	25%	75%

Table 4: Optimal capacity allocation for network III under Minimal Control (M)

n	service	$s12$	$s13$	$s22$	$s33$	$r21$	$r31$	$q1$	$q2$	$q3$	$\eta 21$	$\eta 22$	$\eta 31$	$\eta 33$
10	9.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	1.00	0.00	0%	100%	0%	100%
11	9.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	2.00	0.00	0%	100%	0%	100%
12	9.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	3.00	0.00	0%	100%	0%	100%
13	9.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	4.00	0.00	0%	100%	0%	100%
14	9.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	5.00	0.00	0%	100%	0%	100%
15	9.12	1.12	1.00	1.00	1.00	0.12	0.00	0.00	5.62	0.14	4%	96%	0%	100%
16	9.32	1.32	1.00	1.00	1.00	0.32	0.00	0.00	5.98	0.37	10%	90%	0%	100%
17	9.54	1.54	1.00	1.00	1.00	0.54	0.00	0.00	6.34	0.58	15%	85%	0%	100%
18	9.76	1.77	1.00	1.00	1.00	0.77	0.00	0.00	6.67	0.80	20%	80%	0%	100%
19	10.00	2.00	1.00	1.00	1.00	1.00	0.00	0.00	7.00	1.00	25%	75%	0%	100%
20	10.24	2.25	1.00	1.00	1.00	1.25	0.00	0.00	7.31	1.20	29%	71%	0%	100%
21	10.50	2.50	1.00	1.00	1.00	1.50	0.00	0.00	7.62	1.38	33%	67%	0%	100%
22	10.76	2.77	1.00	1.00	1.00	1.77	0.00	0.00	7.91	1.56	37%	63%	0%	100%
23	11.00	3.00	1.00	1.00	1.00	2.00	0.00	0.00	8.15	1.85	40%	60%	0%	100%
24	11.00	3.00	1.00	1.00	1.00	2.00	0.00	0.00	8.20	2.80	40%	60%	0%	100%
25	11.00	3.00	1.00	1.00	1.00	2.00	0.00	0.00	8.25	3.75	40%	60%	0%	100%
26	11.00	3.00	1.00	1.00	1.00	2.00	0.00	0.01	8.31	4.68	40%	60%	0%	100%
27	11.00	3.00	1.00	1.00	1.00	2.00	0.00	0.02	8.37	5.60	40%	60%	0%	100%
28	11.00	3.00	1.00	1.00	1.00	2.00	0.00	0.05	8.45	6.51	40%	60%	0%	100%
29	11.00	3.00	1.00	1.00	1.00	2.00	0.00	0.05	8.50	7.44	40%	60%	0%	100%
30	11.04	2.97	1.07	1.00	1.00	1.97	0.07	0.00	8.51	8.41	40%	60%	2%	98%
31	11.33	2.81	1.69	0.89	0.94	1.81	0.69	0.00	8.53	8.64	39%	61%	19%	81%
32	11.66	2.89	1.86	0.95	0.97	1.89	0.86	0.00	8.77	8.82	39%	61%	22%	78%
33	12.00	3.00	2.00	1.00	1.00	2.00	1.00	0.00	9.00	9.00	40%	60%	25%	75%
34	12.00	3.00	2.00	1.00	1.00	2.00	1.00	0.33	9.33	9.33	40%	60%	25%	75%
35	12.00	3.00	2.00	1.00	1.00	2.00	1.00	0.67	9.67	9.67	40%	60%	25%	75%

Table 5: Optimal capacity allocation for network III under Admission Control (A)

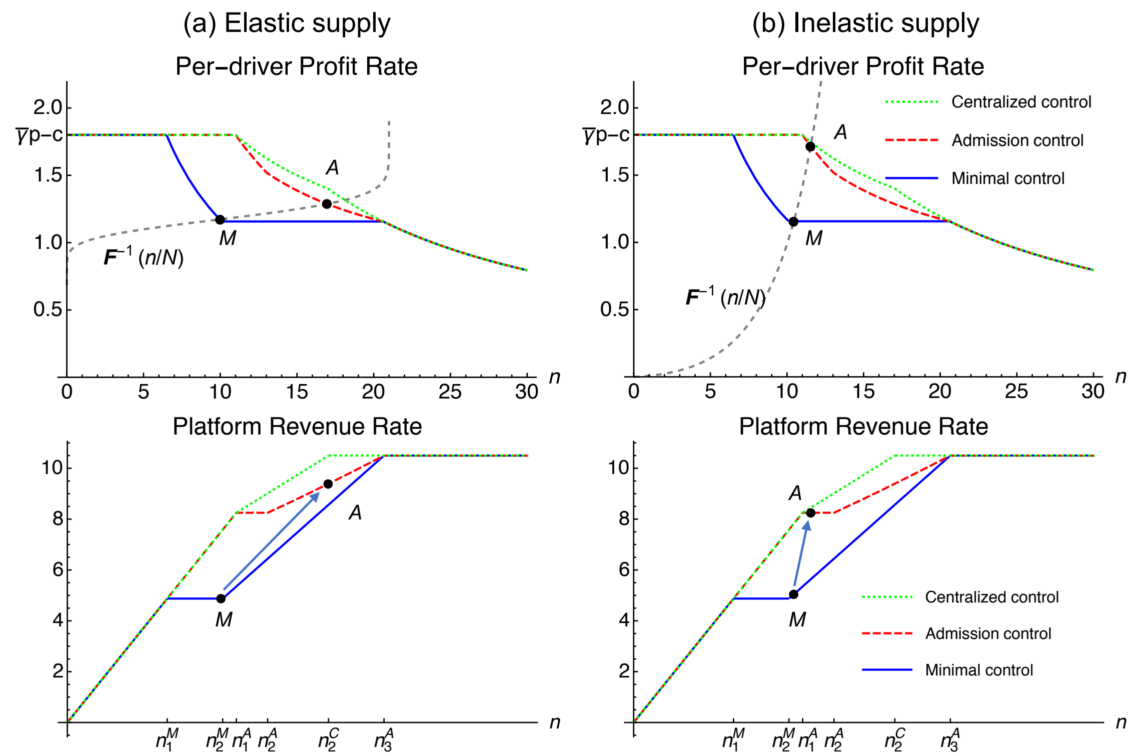

Figure 1: Impact of admission control on the equilibrium capacity, per-driver profit, and platform revenue ($S = (3, 1, 4, 6), \overline{S} = 14, N = 21, t = 1, \gamma = 0.25, p = 3, c = 0.45$)