Notation and proofs

In this e-companion we present the notation used throughout this paper, and we complete the proofs of some propositions.

EC.1. Notation

Sets

I: set of demand nodes;

J: set of candidate facility locations (leader and competition);

J_c: set of competition's facilities;

J_1: set of leader's candidate sites;

$J_1^* \subseteq J_1$: set of leader's open facilities

$J^* \subseteq J$: set of open facilities (leader and competitor).

Parameters

d_i: demand originating from node $i \in I$;

t_{ij}: travel time between nodes $i \in I$ and $j \in J$;

α: coefficient of the waiting time in the disutility formula;

β: coefficient of the balking probability in the disutility formula;

B: available budget (for opening new facilities and associated service rates);

c_f: fixed cost associated with opening a new facility;

c_μ : cost per unit of service;

$\bar{\mu}$: maximum service rate allowed by the budget;

p: number of facilities to open.

Basic decision variables

y_j: binary variable set to 1 if a facility is open at site j, and to 0 otherwise;

μ_j: service rate at open facilities.

Additional variables

x_{ij}: arrival rate at at facility $j \in J$ originating from demand node $i \in I$;

λ_j: arrival rate at node $j \in J$;

ρ_j: utilization rate of facility $j \in J$;

$\bar{\lambda}_j$: throughput rate (customers accessing service) at node $j \in J$;

w_j: mean queueing time at facility j.

EC.2. Proofs of Propositions 1, 2, 4, 5, 6, 7, 9, 10, 11 and Theorem 1

Proposition 1. *The waiting time w_j is increasing in λ_j.*

Proof. The derivative of w_j with respect to λ_j (see Equation (10)) is

$$\frac{\partial w_j}{\partial \lambda_j} = \frac{\partial w_j}{\partial \rho_j}\frac{\partial \rho_j}{\partial \lambda_j} = \frac{\partial w_j}{\partial \rho_j}\frac{1}{\mu_j}.$$

To show that $\partial w_j/\partial \rho_j$ is nonnegative for all $\rho_j \neq 1$, let us consider

$$\frac{\partial w_j}{\partial \rho_j} = \frac{1}{\mu_j}\left(-\frac{K^2\rho_j^{K-1}}{\left(\rho_j^K - 1\right)^2} + \frac{1}{(\rho_j - 1)^2}\right).$$

Basic algebraic manipulation yields

$$\frac{1}{(\rho_j - 1)^2} \geq \frac{K^2\rho_j^{K-1}}{\left(\rho_j^K - 1\right)^2} \Longleftrightarrow \sum_{i=0}^{K-1} \rho_j^i \geq K\rho_j^{(K-1)/2}. \tag{EC.1}$$

To prove that the right-hand inequality holds true, we consider two cases.

If K is odd:

$$\sum_{i=0}^{K-1} \rho_j^i = \sum_{i=0}^{(K-1)/2-1} \left(\rho_j^i + \rho_j^{K-1-i}\right) + \rho_j^{(K-1)/2}$$

$$\geq 2\sum_{i=0}^{(K-1)/2-1} \rho_j^{(K-1)/2} + \rho_j^{(K-1)/2} = K\rho_j^{(K-1)/2}.$$

If K is even:

$$\sum_{i=0}^{K-1} \rho_j^i = \sum_{i=0}^{(K-2)/2} \left(\rho_j^i + \rho_j^{K-1-i}\right) \geq 2 \sum_{i=0}^{(K-2)/2} \rho_j^{(K-1)/2} = K\rho_j^{(K-1)/2}.$$

It follows that w_j is an increasing function of λ_j.

PROPOSITION 2. *The probability of balking p_{Kj} is increasing in λ_j.*

Proof. The derivative of p_{Kj} with respect to λ_j is

$$\begin{aligned} p'_{Kj} &= \frac{\lambda_j^{K-1}\mu_j}{\left(\lambda_j^{K+1} - \mu_j^{K+1}\right)^2} \left[\lambda_j^{K+1} - (K+1)\lambda_j\mu_j^K + K\mu_j^{K+1}\right] \\ &= \sigma[x^{K+1} - (K+1)\,x + K], \end{aligned}$$

where σ is a positive number and $x = \lambda_j/\mu_j$. By differentiating with respect to x, we find that the right-hand-side achieves its minimum value 0 at $x = 1$, which concludes the proof.

PROPOSITION 4. *When $K = \infty$, i.e., balking does not occur (in this case, the model admits a solution only if the total service rate exceeds the total demand rate), the lower level objective function is convex jointly in λ and μ.*

Proof. If $K = \infty$, the probability of balking can be removed from the objective, since it is equal to 0. Moreover, $w_j = 1/(\mu_j - \lambda_j)$, and the lower level objective takes the form

$$\sum_{i\in I}\sum_{j\in J^*}\left[\frac{1}{\theta}x_{ij}\ln x_{ij} + x_{ij}t_{ij}\right] - \alpha\sum_{j\in J^*}\ln(\mu_j - \lambda_j).$$

Basic algebra shows that its Hessian is positive semidefinite, hence the function is convex.

PROPOSITION 5. *The integral of the waiting time, $W_j(\lambda_j,\mu_j)$ is pseudoconvex.*

Proof. Let $x = (\lambda_x, \mu_x)$ and $y = (\lambda_y, \mu_y)$. Assume that $\nabla W(x)(y - x) \geq 0$. Then we have:

$$\begin{aligned} &\left(w_j(x), -\frac{\lambda_x}{\mu_x}w_j(x)\right)(\lambda_y - \lambda_x, \mu_y - \mu_x) \geq 0 \\ \Rightarrow \quad & (\rho_y - \rho_x)w_j(x) \geq 0 \\ \Rightarrow \quad & \rho_y \geq \rho_x, \end{aligned} \tag{EC.2}$$

since w_j is nonnegative. On the other hand, $\partial W_j/\partial\rho = \mu_j w_j$ is nonnegative, we have that W_j is increasing in ρ, so $\rho_y \geq \rho_x \Rightarrow W_j(y) \geq W_j(x)$. From Eq (EC.2) it follows that if $\nabla W(x)(y-x) \geq 0$ then $W_j(y) \geq W_j(x)$, hence W_j is pseudoconvex.

PROPOSITION 6. *G_1 is strongly monotone in x of modulus $\theta \cdot d_{MAX}$.*

Proof. Gilbert et al. (2015) have already argued that G_1 is strongly monotone. Indeed, the associated Jacobian is a positive definite diagonal matrix over D, with the smallest possible eigenvalue $1/(\theta \cdot d_{\mathrm{MAX}})$. It follows that G_1 is strongly monotone with modulus $\theta \cdot d_{\mathrm{MAX}}$.

PROPOSITION 7. *G_2 is monotone in x.*

Proof.

$$
\begin{aligned}
\langle G_2(\mu,x) - G_2(\mu,y), x-y\rangle &= \sum_{i\in I}\sum_{j\in J^*}\left(\frac{1}{\mu_j - \sum\limits_{l\in I} x_{l,j}} - \frac{1}{\mu_j - \sum\limits_{l\in I} y_{l,j}}\right)\cdot(x_{ij}-y_{ij}) \\
&= \sum_{j\in J^*}\left[\frac{\mu_j - \sum\limits_{l\in I} y_{l,j} - \mu_j + \sum\limits_{l\in I} x_{l,j}}{\left(\mu_j - \sum\limits_{l\in I} x_{l,j}\right)\cdot\left(\mu_j - \sum\limits_{l\in I} y_{l,j}\right)}\cdot\sum_{i\in I}(x_{ij}-y_{ij})\right] \\
&= \sum_{j\in J^*}\left[\frac{\sum\limits_{l\in I}(x_{l,j}-y_{l,j})\sum\limits_{l\in I}(x_{l,j}-y_{l,j})}{\left(\mu_j - \sum\limits_{l\in I} x_{l,j}\right)\cdot\left(\mu_j - \sum\limits_{l\in I} y_{l,j}\right)}\right] \\
&\geq 0
\end{aligned}
$$

PROPOSITION 9. *If $K=\infty$ and there are no fixed costs, the surrogate model is convex.*

Proof. According to Proposition 4, the objective is jointly convex in μ and λ. Moreover one can, without loss of generality, open all facilities and hence dispense with the binary vector y. Notwithstanding, a facility can be closed by setting its service level to zero.

PROPOSITION 10. *At the optimum of (PH*), if $K=\infty$, queueing delays are equal for all leader's facilities.*

Proof. For fixed y variables, Equation (66) can be rewritten as

$$\sum_{j\in J^*}\mu_j \le \bar{\mu}, \tag{EC.3}$$

where $\bar{\mu}$ is the maximum possible total service rate allowed by the budget. But $K=\infty$, so $w_j(\lambda_j,\mu_j)=1/(\mu_j-\lambda_j)$ and $p_{Kj}(\lambda_j,\mu_j)=0$, which yields the mathematical program

$$\begin{aligned}(\text{PHY*})\qquad &\min_{\mu,x}\ \sum_{i\in I}\sum_{j\in J^*} x_{ij}t_{ij}-\alpha\sum_{j\in J^*}\ln(\mu_j-(\sum_{i\in I}x_{ij}))\\ &\text{s.t.}\quad \text{constraints (65), (68), (69), (EC.3)}\end{aligned}$$

Let δ_i, π_{ij} and γ be the Lagrange multipliers associated with Equations (65), (69) and (EC.3), respectively. Variables δ_i are free, while γ and π_{ij} are restricted to be nonnegative. The stationarity conditions of the above program are:

$$\frac{\partial L}{\partial x_{ij}}=0 \ \Rightarrow \qquad t_{ij}+\alpha w_j(\lambda_j,\mu_j)-\delta_i-\pi_{ij}=0,\ \forall i\in I,\forall j\in J^* \tag{EC.4}$$

$$\frac{\partial L}{\partial \mu_j}=0 \ \Rightarrow \qquad -\alpha w_j(\lambda_j,\mu_j)+\gamma=0,\ \forall j\in J^*\cap J_1, \tag{EC.5}$$

and the conclusion follows from Equation (EC.5).

We observe, after plugging $\alpha w_j(\lambda_j,\mu_j)$ from Equation (EC.5) into Equation (EC.4) for a given demand node i, that only one flow x_{ij} is nonzero, provided that transportation times to the leader's facilities are distinct.

Proposition 11. *There exists a value of ξ^* for which (PHY*(ξ^*)) yields an optimal solution for (P*).*

Proof. Let y^* and μ^* be optimal for (P*). Without loss of generality (there are no fixed costs) we assume that all facilities are open. At equilibrium, let c_i^* be the cost associated with demand node i and optimal service rate μ^*. Let x^*, $w_j(x^*,\mu_j^*)$ and c_i^* satisfy Equation (24) and (25). If x_{ij} is positive, we have:

$$t_{ij}+\alpha w_j(x^*,\mu^*)=c_i^*,\quad \forall j\in J,\forall i\in I. \tag{EC.6}$$

Let $C = \max_{i\in I}\{c_i^*\}$ in the initial formulation. For $j \in J$, we let $\xi_j = c_i^* - t_{ij} - C$ and select and index i corresponding to a positive flow x_{ij}^*. If no such i exists, then $\mu_j^* = 0$, otherwise the leader would waste monetary resources. We then set $\xi_j = -C$.

Now, let δ_i, π_{ij} and γ be the Lagrange multipliers associated with Equations (65), (69) and (EC.3), respectively. Variables δ_i and γ are free, while π_{ij} are restricted to be nonnegative. The stationarity conditions of the program above take the form

$$\frac{\partial \mathcal{L}}{\partial x_{ij}} = 0 \;\Rightarrow\qquad t_{ij} + \alpha w_j(x,\mu_j) - \delta_i = 0, \;\text{ if } x_{ij} > 0,\; \forall i \in I, \forall j \in J \tag{EC.7}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_j} = 0 \;\Rightarrow\qquad -\alpha w_j(x,\mu_j) + \gamma + \xi_j = 0,\; \forall j \in J_1. \tag{EC.8}$$

Note that the derivative of the Lagrangian with respect to x_{ij} is left unchanged, i.e., Equation (EC.7) is equivalent to Equation (EC.4). If $\gamma = C$, we derive from Equation (EC.8) that $\alpha w_j(x,\mu_j) = c_i^* - t_{ij}$, which is equivalent to Equation (EC.7). This completes the proof, since for the given values of ξ, variables x and μ match the optimal solution of (P*).

Theorem 1. *The error of the upper-level objective function is $O(1/N_1 + 1/N_2)$, where N_1 and N_2 are the number of samples for the linearization of g_1 and g_2, respectively.*

Proof. Let $\bar{G}$ be an approximation of G. We denote by $\bar{x}$ the solution of IV$(\bar{G}(\mu,\cdot), D)$, and by x the solution of IV$(G(\mu,\cdot), D)$. Then the following inequalities hold:

$$\langle G(\mu,x), \bar{x} - x\rangle \geq 0$$

$$\left\langle \bar{G}(\mu,\bar{x}), x - \bar{x}\right\rangle \geq 0$$

$$\Rightarrow \left\langle \bar{G}(\mu,\bar{x}) - G(\mu,x), x - \bar{x}\right\rangle \geq 0 \tag{EC.9}$$

From the strong monotonicity of G and Eq. (EC.9) it follows that

$$\left\langle \bar{G}(\mu,\bar{x}) - G(\mu,\bar{x}), x - \bar{x}\right\rangle \geq \frac{1}{\theta \cdot d_{\mathrm{MAX}}} ||x - \bar{x}||^2. \tag{EC.10}$$

Applying the Cauchy-Schwarz inequality, we obtain

$$\theta \cdot d_{\mathrm{MAX}} \cdot ||\bar{G}(\mu,\bar{x}) - G(\mu,\bar{x})|| \geq ||x - \bar{x}||. \tag{EC.11}$$

It follows that

$$
\begin{aligned}
|f(x)-f(\bar{x})| &= |\sum_{i\in I}\sum_{j\in J_1^*}(x_{ij}-\bar{x}_{ij})| \le \sqrt{|I|\cdot|J|}||x-\bar{x}|| \quad \text{(Cauchy-Schwarz inequality)} \\
&\le \sqrt{|I|\cdot|J|}\;\theta\cdot d_{\text{MAX}}\cdot||\bar{G}(\mu,\bar{x})-G(\mu,\bar{x})||. \qquad \text{(EC.12)}
\end{aligned}
$$

We perform two separate linear approximations on g_1 and g_2, respectively. Then the mappings $\bar{G}_1(x)$ and $\bar{G}_2(x)$ are piecewise constant approximations, that we detail separately.

A. $\bar{G}_1$: Each component (i,j) of this vector is a piecewise constant approximation of $\log(x_{ij})$, satisfying:

i) there are N_1 total samples on x_{ij}, starting from $r_{\min}$ to d_{MAX};

ii) the sampling points are chosen so that the segments are vertically equidistant;

iii) the vertical positions of the segments are the slopes of the tangents to $x\log(x)$, evaluated at the sampling points.

Let Δ_1 be the difference between two consecutive slope values:

$$\Delta_1 = \frac{\log(d_{\text{MAX}})-\log(r_{\min})}{N_1-1}.$$

Then $|\bar{G}_{1(i,j)}-G_{1(i,j)}|\le\Delta_1$, which yields:

$$||\bar{G}_1(x)-G_1(x)|| = \sqrt{\sum_{i\in I}\sum_{j\in J^*}|\bar{G}_{1(i,j)}-G_{1(i,j)}|^2} \le \frac{(\log(d_{\text{MAX}})-\log(r_{\min}))\sqrt{|I|\cdot|J|}}{N_1-1} \qquad \text{(EC.13)}$$

B. $\bar{G}_2$: is a mapping whose (i,j)-component is a constant piecewise approximation of $1/q_j$, where $q_j=\mu_j-\sum_{i\in I}x_{ij}$. Similar to $\bar{G}_1$, this linearization satisfies the following:

i) there are N_2 total samples, starting from ψ to μ_{MAX};

ii) the sampling points are chosen so that the segments are vertically equidistant;

iii) the vertical positions of the segments are the slopes of the tangents $-\log(q)$ evaluated at the sampling points.

We note by Δ_2 the difference between two consecutive slope values:

$$\Delta_2 = \frac{\dfrac{1}{\psi} - \dfrac{1}{\mu_{\text{MAX}}}}{N_2 - 1} /$$

Then $|\bar{G}_{2(i,j)} - G_{2(i,j)}| \leq \Delta_2$, which yields

$$||\bar{G}_2(x) - G_2(x)|| = \sqrt{\sum_{i \in I} \sum_{j \in J^*} |\bar{G}_2(x_{ij}) - G_2(x_{ij})|^2} \leq \frac{\left(\dfrac{1}{\psi} - \dfrac{1}{\mu_{\text{MAX}}}\right) \sqrt{|I| \cdot |J|}}{N_2 - 1} \tag{EC.14}$$

From Eq. (EC.12) it follows that, given y and μ:

$$|f(x) - f(\bar{x})| \leq \theta \cdot d_{\text{MAX}} |I| \cdot |J| \left[\frac{(\log(d_{\text{MAX}}) - \log(r_{\min}))}{N_1 - 1} + \alpha \frac{\dfrac{1}{\psi} - \dfrac{1}{\mu_{\text{MAX}}}}{N_2 - 1} \right] \in O(\frac{1}{N_1} + \frac{1}{N_2}). \tag{EC.15}$$

Theorem 1 has several implications.

- For a given set of open facilities, the absolute difference between the optimal and the approximated objective value is bounded by the right-hand-side of inequality (EC.15). For large values of N_1 and N_2, the two values are very close.
- If the optimal solution is unique in terms of the location vector y, and the absolute difference between the objective and other solutions objectives are lower than the right hand side of inequality (EC.15), the approximation algorithm will find the optimum locations.

EC.3. Linearization of optimality conditions

EC.3.1. Complementarity constraints for Program (P2-lin)

Let γ_i, δ_j, ν_{ij}^n, π_j^{rp}, η_j^{rp}, and ϕ_{ij} be the dual variables associated with constraints (47), (48), (49), (50), (51) and (52), respectively. Then the complementarity constraints for program (P2-lin) can be written as:

$$\gamma_i \left(\sum_{j \in J^*} x_{ij} - d_i \right) = 0 \qquad \forall i \in I \tag{EC.16}$$

$$\delta_j \left(\lambda_j - \sum_{i \in I} x_{ij} \right) = 0 \qquad \forall j \in J^* \tag{EC.17}$$

$$\nu_{ij}^n \left(v_{ij} - a_f^n x_{ij} - b_f^n \right) = 0 \qquad \forall i \in I; \quad \forall j \in J^*; \quad \forall n \in N \tag{EC.18}$$

$$\pi_j^{rp}\left(u_j - a_g^{rp}\lambda_j - b_g^{rp}\mu_j - c_g^{rp}\right) = 0 \qquad \forall j \in J^*;\quad \forall r \in R;\quad \forall p \in P \tag{EC.19}$$

$$\eta_j^{rp}\left(z_j - a_h^{rp}\lambda_j - b_h^{rp}\mu_j - c_h^{rp}\right) = 0 \qquad \forall j \in J^*;\quad \forall r \in R;\quad \forall p \in P \tag{EC.20}$$

$$\phi_{ij}x_{ij} = 0 \qquad \forall i \in I;\quad \forall j \in J^*, \tag{EC.21}$$

and can be linearized in the standard fashion, through the introduction of binary variables and big-M constants. For instance, the last constraint is replaced by the inequalities

$$\phi_{ij} \le M u_{ij}$$

$$x_{ij} \le M(1 - u_{ij}),$$

where $u_{ij} \in \{0,1\}$. Although is possible to find a valid upper bound for the variable ϕ_{ij}, a large value of M is required, which leads to a poor relaxation and consequently an ill-behaved branch-and-bound algorithm.

EC.3.2. Equality between primal and dual objectives

Alternatively, constraints (EC.16) – (EC.21) can be replaced with constraint (EC.22), which represents the equality between between the primal and dual objective of (P2-lin). Then the optimality constraints of (P2-lin) are

$$\sum_{i\in I}\gamma_i d_i + \sum_{n\in N}\sum_{i\in I}\sum_{j\in J}\nu_{ij}^n b_f^n + \sum_{r\in R}\sum_{p\in P}\sum_{j\in J}\left(b_g^{rp}\mu_j\pi_j^{rp} + b_h^{rp}\mu_j\eta_j^{rp} + c_g^{rp}\pi_j^{rp} + c_h^{rp}\eta_j^{rp}\right)$$
$$= \sum_{i\in I}\sum_{j\in J}\left[\frac{1}{\theta}v_{ij} + x_{ij}t_{ij}\right] + \alpha\sum_{j\in J}u_j + \beta\sum_{j\in J}z_j, \tag{EC.22}$$

$$\sum_{j\in J} x_{ij} = d_i, \qquad \forall i \in I$$

$$\lambda_j = \sum_{i\in I} x_{ij}, \qquad \forall j \in J$$

$$v_{ij} - a_f^n x_{ij} \ge b_f^n, \qquad \forall i \in I; \forall j \in J; \forall n \in N$$

$$u_j - a_g^{rp}\lambda_j - b_g^{rp}\mu_j \ge c_g^{rp}, \qquad \forall j \in J; \forall r \in R; \forall p \in P$$

$$z_j - a_h^{rp}\lambda_j - b_h^{rp}\mu_j \ge c_h^{rp}, \qquad \forall j \in J; \forall r \in R; \forall p \in P$$

$$\gamma_i + \delta_j - \sum_{n \in N} a_f^n \nu_{ij}^n \le t_{ij}, \qquad \forall i \in I;\ \forall j \in J$$

$$-\delta_j - \sum_{r \in R} \sum_{p \in P} \left(a_g^{rp} \pi_j^{rp} + a_h^{rp} \eta_j^{rp} \right) = 0, \qquad \forall j \in J$$

$$\sum_{n \in N} \nu_{ij}^n = \frac{1}{\theta}, \qquad \forall i \in I;\ \forall j \in J$$

$$\sum_{r \in R} \sum_{p \in P} \pi_j^{rp} = \alpha, \qquad \forall j \in J$$

$$\sum_{r \in R} \sum_{p \in P} \eta_j^{rp} = \beta, \qquad \forall j \in J$$

$$\pi_j^{rp}, \eta_j^{rp} \ge 0, \qquad \forall j \in J; \forall r \in R; \forall p \in P$$

$$x_{ij} \ge 0,\ \nu_{ij}^n \ge 0, \qquad \forall i \in I; \forall j \in J; \forall n \in N.$$

To obtain a MILP formulation, we linearize the nonlinear terms $\mu_j \pi_j^{rp}$ and $\mu_j \eta_j^{rp}$ via the triangle method described in D'Ambrosio et al. (2010). For each term $\mu_j \pi_j^{kq}$ we introduce $2(R-1)(P-1)$ binary variables $\bar{l}_{jrpkq}^{\pi}$ and $\underline{l}_{jrpkq}^{\pi}$ associated with the upper and lower triangles, respectively, of the rectangle defined by the intervals $[\pi^r, \pi^{r+1})$ and $[\mu^p, \mu^{p+1})$. Note that the values of π and η are upper bounded by α and β, respectively. Additionally, μ is bounded by the maximum value allowed by the leader's budget, $\bar{\mu}$. Next, we introduce $J_1 RP$ continuous variables $s_{jrpkq} \in [0,1]$ which will be used to express the couple (π_j^{kq}, μ_j) as a convex combination of triangle vertices. We introduce a similar linearization for the term $\mu_j \eta_j^{kh}$. The approximation of $\mu_j \pi_j^{kq}$ and $\mu_j \eta_j^{kq}$ is then

$$\sum_{r=1}^{R-1} \sum_{p=1}^{P-1} \left(\bar{l}_{jrpkq}^{\pi} + \underline{l}_{jrpkq}^{\pi} \right) = 1, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.23}$$

$$s_{jrpkq}^{\pi} \le \bar{l}_{jrpkq}^{\pi} + \underline{l}_{jrpkq}^{\pi} + \bar{l}_{jrp-1kq}^{\pi} + \underline{l}_{jr-1p-1kq}^{\pi} + \bar{l}_{jr-1p-1kq}^{\pi} + \underline{l}_{jr-1pkq}^{\pi},$$
$$\forall j \in J_1; \forall r \in R; \forall p \in P; \forall k \in R; \forall q \in P \tag{EC.24}$$

$$\sum_{r=1}^{R} \sum_{p=1}^{P} s_{jrpkq}^{\pi} = 1, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.25}$$

$$\pi_j^{kq} = \sum_{r=1}^{R} \sum_{p=1}^{P} s_{jrpkq}^{\pi} \pi^r, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.26}$$

$$\mu_j = \sum_{r=1}^{R} \sum_{p=1}^{P} s_{jrpkq}^{\pi} \mu^p, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.27}$$

$$e^{\pi}_{jkq} = \sum_{r=1}^{R}\sum_{p=1}^{P} s^{\pi}_{jrpkq}\pi^{r}\mu^{p}, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.28}$$

$$\sum_{r=1}^{R-1}\sum_{p=1}^{P-1}\left(\bar{l}^{\eta}_{jrpkq} + \underline{l}^{\eta}_{jrpkq}\right) = 1, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.29}$$

$$s^{\eta}_{jrpkq} \le \bar{l}^{\eta}_{jrpkq} + \underline{l}^{\eta}_{jrpkq} + \bar{l}^{\eta}_{jrp-1kq} + \underline{l}^{\eta}_{jr-1p-1kq} + \bar{l}^{\eta}_{jr-1p-1kq} + \underline{l}^{\eta}_{jr-1pkq},$$
$$\forall j \in J_1; \forall r \in R; \forall p \in P; \forall k \in R; \forall q \in P \tag{EC.30}$$

$$\sum_{r=1}^{R}\sum_{p=1}^{P} s^{\eta}_{jrpkq} = 1, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.31}$$

$$\eta^{kq}_{j} = \sum_{r=1}^{R}\sum_{p=1}^{P} s^{\eta}_{jrpkq}\eta^{r}, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.32}$$

$$\mu_j = \sum_{r=1}^{R}\sum_{p=1}^{P} s^{\eta}_{jrpkq}\mu^{p}, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.33}$$

$$e^{\eta}_{jkq} = \sum_{r=1}^{R}\sum_{p=1}^{P} s^{\eta}_{jrpkq}\eta^{r}\mu^{p}, \qquad \forall j \in J_1; \forall k \in R; \forall q \in P \tag{EC.34}$$

The complete MILP formulation is presented below. It involves variables associated with the original fixed point (or bilevel) formulation (y, μ, x), together with variables issued from the linearizations and primal-dual optimality conditions.

$$\text{(P-lin)} \qquad \max_{\substack{x,y,\mu,\lambda,u,v,z,\\ e,\pi,\eta,\nu,\gamma,\delta,\bar{l}^{\pi},\underline{l}^{\pi},\\ s^{\pi},e^{\pi},s^{\eta},\bar{l}^{\eta},\underline{l}^{\eta}\\ s,\bar{l},\underline{l}}} \sum_{j\in J_1} e_j$$

$$\sum_{j\in J_1} y_j c_f + \sum_{j\in J_1} c_{\mu}\mu_j \le B,$$

$$\mu_j \le \bar{\mu} y_j, \qquad \forall j \in J_1$$

$$\sum_{r\in R}\sum_{p\in P}\sum_{j\in J_c}\left(b^{rp}_{g}\pi^{rp}_{j}\mu_j + b^{rp}_{h}\eta^{rp}_{j}\mu_j + c^{rp}_{g}\pi^{rp}_{j} + c^{rp}_{h}\eta^{rp}_{j}\right) + \sum_{n\in N}\sum_{i\in I}\sum_{j\in J}\nu^{n}_{ij}b^{n}_{f}$$
$$+ \sum_{r\in R}\sum_{p\in P}\sum_{j\in J_1}\left(b^{rp}_{g}e^{\pi}_{jrp} + b^{rp}_{h}e^{\eta}_{jrp} + c^{mp}_{g}\pi^{rp}_{j} + c^{rp}_{h}\eta^{rp}_{j}\right) + \sum_{i\in I}\gamma_i d_i$$
$$= \sum_{i\in I}\sum_{j\in J}\left[\frac{1}{\theta}v_{ij} + x_{ij}t_{ij}\right] + \alpha\sum_{j\in J} u_j + \beta\sum_{j\in J} z_j,$$

$$\sum_{j\in J} x_{ij} = d_i, \qquad \forall i \in I$$

$$\lambda_j = \sum_{i\in I} x_{ij}, \qquad \forall j \in J$$

$$v_{ij} - a_f^n x_{ij} \geq b_f^n, \qquad \forall i \in I; \forall j \in J; \forall n \in N$$

$$u_j - a_g^{rp}\lambda_j - b_g^{rp}\mu_j \geq c_g^{rp}, \qquad \forall j \in J; \forall r \in R; \forall p \in P$$

$$z_j - a_h^{rp}\lambda_j - b_h^{rp}\mu_j \geq c_h^{rp}, \qquad \forall j \in J; \forall r \in R; \forall p \in P$$

$$\gamma_i + \delta_j - \sum_{n\in N} a_f^n \nu_{ij}^n \leq t_{ij}, \qquad \forall i \in I;\ \forall j \in J$$

$$-\delta_j - \sum_{r\in R}\sum_{p\in P}\left(a_g^{rp}\pi_j^{rp} + a_h^{rp}\eta_j^{rp}\right) = 0, \qquad \forall j \in J$$

$$\sum_{n\in N} \nu_{ij}^n = \frac{1}{\theta}, \qquad \forall i \in I;\ \forall j \in J$$

$$\sum_{r\in R}\sum_{p\in P} \pi_j^{rp} = \alpha, \qquad \forall j \in J$$

$$\sum_{r\in R}\sum_{p\in P} \eta_j^{rp} = \beta, \qquad \forall j \in J$$

constraints (EC.23) – (EC.34) and (53) – (58),

$$y_j \in \{0,1\},\ \mu_j, \pi_j^{rp}, \eta_j^{rp} \geq 0, \qquad \forall j \in J; \forall r \in R; \forall p \in P$$

$$x_{ij} \geq 0,\ \nu_{ij}^n \geq 0, \qquad \forall i \in I; \forall j \in J; \forall n \in N.$$

EC.3.3. Example of lower level linearization when $K = \infty$

Recall that, according to Proposition 4, the function is convex if the buffer zone is infinite (no balking). In that situation, the maximum of the linear approximations is consistent with the original function, give or take the approximation error. Proceeding as before, we obtain

$$g^{rp}(\lambda,\mu) = a_g^{rp}\lambda + b_g^{rp}\mu + c_g^{rp} = \frac{\alpha}{\mu^p - \lambda^r}\lambda - \frac{\alpha}{\mu^p - \lambda^r}\mu - \alpha(\ln(\mu^p - \lambda^r) - 1). \tag{EC.35}$$

This yields the linearized lower level program

$$(\text{P2}^\infty) \qquad \min_{x,v,u,\lambda} \quad \sum_{i\in I}\sum_{j\in J^*}\left[\frac{1}{\theta}v_{ij} + x_{ij}t_{ij}\right] + \alpha\sum_{j\in J^*} u_j \tag{EC.36}$$

$$\text{s.t.} \quad \sum_{j\in J^*} x_{ij} = d_i, \qquad \forall i \in I \tag{EC.37}$$

$$\lambda_j = \sum_{i \in I} x_{ij}, \qquad \forall j \in J^* \tag{EC.38}$$

$$v_{ij} - a_f^n x_{ij} \geq b_f^n, \qquad \forall i \in I; \forall j \in J^*; \forall n \in N \tag{EC.39}$$

$$u_j - a_g^{rp} \lambda_j - b_g^{rp} \mu_j \geq c_g^{rp}, \qquad \forall j \in J^*; \forall r \in R; \forall p \in P \tag{EC.40}$$

$$x_{ij} \geq 0, \qquad \forall i \in I; \forall j \in J^*. \tag{EC.41}$$

EC.3.4. Taxonomy

This section provides a taxonomy of the models most relevant to our research, with respect to four features: (i) user choice environment (yes or no), (ii) stochastic (or not), (iii) inclusion of congestion (or not) at facilities, (iv) inclusion (or not) of competition. The relevant information is displayed in Table EC.1.

Authors	user choice	stochastic	congestion	competition
Abouee-Mehrizi et al. (2011)	×	×	×	
Averbakh et al. (2007)	×			
Berman and Drezner (2006)	×		×	
Camacho-Vallejo et al. (2014)	×			
Castillo et al. (2009)		×	×	
Desrochers et al. (1995)			×	
Kim (2013)			×	
Küçükaydin et al. (2011)	×	×		×
Labbé and Hakimi (1991)				×
Marianov and Serra (2001)			×	
Marianov (2003)		×	×	
Marianov et al. (2008)	×	×	×	×
Marić et al. (2012)	×			
Rahmati et al. (2014)		×	×	
Vidyarthi and Jayaswal (2014)		×	×	
Zhang et al. (2010)	×		×	

Table EC.1 Taxonomy of congested facility location models