Online Supplement

The online supplement is organized as follows. Appendix A provides proofs of lemmas, theorems, and propositions. Appendix B provides the formulation and details of a proposed deterministic problem in Section 3.1.1. Appendix C explains the optimization problem that needs to be solved for selective bidding problem. Appendix D explains the details of our proposed heuristic algorithm to solve the ad-exchanges selection problem for the *non-identical ad-exchanges* case. Appendix E illustrates the formulation and details of a proposed deterministic problem in Section 5.2.

Appendix A. Proofs of Lemmas, Theorems, and Propositions

Proof of Theorem 1.

To prove this theorem, we show that the NP-complete *Set Cover* problem can be reduced to $\mathcal{P}(\Omega)$. Consider an instance of the Set Cover problem:

DEFINITION 1 (SET COVER PROBLEM). Given a set of elements $U = \{u_1, u_2, ..., u_n\}$ and a collection of subsets $S_1, S_2, ..., S_m$ of U, the set cover problem is to tell whether there exist k of the subsets whose union is equals U.

Given an arbitrary instance of the Set Cover problem, we define a corresponding instance of problem $\mathcal{P}(\Omega)$ as follows. Set $T = 1$, i.e., the planning horizon only contains one slot, thus the decision variable $y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)$ can be simplified to $y^{\pi(W)}_{i,r}$. There is an ad-exchange i corresponding to each subset S_i, a location r corresponding to each element u_r. For all $y \in (0, 1]$, define $\psi_{i,r}(y) = 0$ (resp. $\psi_{i,r}(y) = n$) if and only if $u_r \in S_i$ (resp. $u_r \notin S_i$), $\forall i, r$. In addition, $\tilde{K}_i = 0$, $p_{i,r} = 1/(mn)$, $C_i = 1$, $\beta = 1/(mn)$, and $\xi_r = 1$, $\forall i, r$. The Set Cover problem is equivalent to deciding if there is a solution to $\mathcal{P}(\Omega)$ with cost no larger than k. We prove this in both directions. In one direction, selecting k ad-exchanges, say W', corresponding to sets in a Set Cover solution and setting $y^{\pi(W)}_{i,r} = 1$ (resp. $y^{\pi(W)}_{i,r} = 0$) if and only if $u_r \in S_i$ (resp. $u_r \notin S_i$), $\forall i \in W', r \in \{1, 2, ..., n\}$, results in obtaining at least one impression in each location r with probability $1/(mn)$. The cost of the above solution is no larger than k. In the other direction, if there exists a solution to problem $\mathcal{P}(\Omega)$ with cost no larger than k, then the number of ad-exchanges in that solution is at most k (this is because $C_i = 1, \forall i$). Moreover, for each location r, there exists at least one ad-exchange from that solution, say i, such that $u_r \in S_i$ (otherwise the cost is at least n due to $\psi_{i,r}(y) = n$) if and only $u_r \notin S_i$. Then we can select at most k subsets, corresponding to ad-exchanges in the solution to problem $\mathcal{P}(\xi)$, results in covering all elements in U. □

Proof of Proposition 1.

We first need to show that the optimal objective function value of the deterministic problem is a lower bound for the optimal objective function value of Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$. Then we need to show that there is a feasible solution for Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$ whose objective function value is same as the optimal objective function value of the deterministic problem.

For proving the first claim, we show for any feasible policy for Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$, there is a feasible solution for Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha}, W)$ whose objective function value is smaller. We consider an arbitrary feasible policy for Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$ as $\pi(W) := \big(y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\big)$. We define

$$\overline{y}^{\mathcal{D}}_{i,r} = \max\left\{\hat{y}^{\mathcal{D}}_{i,r}, \hat{y}^{\mathcal{D}}_{i,r} + \gamma\right\}, \quad \forall (i,r), \tag{A.1}$$

where

$$\hat{y}^{\mathcal{D}}_{i,r} = \frac{\sum_{t=1}^{T} \mathbb{E}_{\mathbf{S}^{\pi(W)}_t}\left[y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\right]}{T}, \quad \forall (i,r), \tag{A.2}$$

and $-\hat{y}^{\mathcal{D}}_{i,r} \le \gamma \le 1 - \hat{y}^{\mathcal{D}}_{i,r}$ and is obtained from the following equation

$$\sum_{r \in \mathcal{R}} p_{i,r} h_{i,r}\big(\hat{y}^{\mathcal{D}}_{i,r} + \gamma\big) = \frac{C_i}{T}. \tag{A.3}$$

To proceed the proof, there are two cases to consider.

• **Case 1**. $-\hat{y}^{\mathcal{D}}_{i,r} \le \gamma \le 0$.

Under this case, $\overline{y}^{\mathcal{D}}_{i,r} = \hat{y}^{\mathcal{D}}_{i,r}$. By multiplying both sides of (A.2) by $T p_{i,r}$ and summing over $i \in W$, we obtain $T \sum_{i \in W} p_{i,r} \overline{y}^{\mathcal{D}}_{i,r} = \sum_{t=1}^{T} \sum_{i \in W} p_{i,r} \mathbb{E}_{\mathbf{S}^{\pi(W)}_t}\left[y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\right]$. Since $y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)$ is a feasible solution for Problem $\mathcal{P}(\boldsymbol{\xi}, W)$, from (2), we know that $\sum_{t=1}^{T} \sum_{i \in W} p_{i,r} \mathbb{E}_{\mathbf{S}^{\pi(W)}_t}\left[y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\right] \ge \xi_r$. Thus, we have

$$T \sum_{i \in W} p_{i,r} \overline{y}^{\mathcal{D}}_{i,r} \ge \xi_r, \tag{A.4}$$

which is constraint $(E.1)$ of Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha}, W)$.

Since $-\hat{y}^{\mathcal{D}}_{i,r} \le \gamma \le 0$ and function $h_{i,r}(.)$ is strictly increasing, we have $h_{i,r}(\overline{y}^{\mathcal{D}}_{i,r}) \ge h_{i,r}(\overline{y}^{\mathcal{D}}_{i,r} + \gamma)$. We multiply both sides of above equation by $T p_{i,r}$ and sum over $r \in \mathcal{R}$ and get $T \sum_{r \in \mathcal{R}} p_{i,r} h_{i,r}(\overline{y}^{\mathcal{D}}_{i,r}) \ge T \sum_{r \in \mathcal{R}} p_{i,r} h_{i,r}(\overline{y}^{\mathcal{D}}_{i,r} + \gamma)$. We know from $(A.3)$ that $T \sum_{r \in \mathcal{R}} p_{i,r} h_{i,r}(\overline{y}^{\mathcal{D}}_{i,r} + \gamma) = C_i$. Thus, we have

$$max\left\{T \sum_{r \in \mathcal{R}} p_{i,r} h_{i,r}(\overline{y}^{\mathcal{D}}_{i,r}), C_i\right\} = T \sum_{r \in \mathcal{R}} p_{i,r} h_{i,r}(\overline{y}^{\mathcal{D}}_{i,r}). \tag{A.5}$$

Now we need to show that the objective function value of this solution in Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha}, W)$ is lower than that of $y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)$ in Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$. Since function $h_{i,r}(.)$ is convex, by Jensen's inequality, we have

$$\mathbb{E}_{\mathbf{S}^{\pi(W)}_t}\left[h_{i,r}\big(y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\big)\right] \ge h_{i,r}\Big(\mathbb{E}_{\mathbf{S}^{\pi(W)}_t}\left[y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\right]\Big), \quad \forall (i,r),$$

$$\frac{\sum_{t=1}^{T} \mathbb{E}_{\mathbf{S}^{\pi(W)}_t}\left[h_{i,r}\big(y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\big)\right]}{T} \ge h_{i,r}\left(\frac{\sum_{t=1}^{T} \mathbb{E}_{\mathbf{S}^{\pi(W)}_t}\left[y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\right]}{T}\right) = h_{i,r}\big(\overline{y}^{\mathcal{D}}_{i,r}\big), \quad \forall (i,r).$$

Multiplying both sides by $T p_{i,r}$ and summing over $r \in \mathcal{R}$, we get

$$\sum_{t=1}^{T} \sum_{r \in \mathcal{R}} p_{i,r} \mathbb{E}_{\mathbf{S}^{\pi(W)}_t}\left[h_{i,r}\big(y^{\pi(W)}_{t,i,r}(t, \mathbf{s}^{\pi(W)}_t)\big)\right] \ge T \sum_{r \in \mathcal{R}} p_{i,r} h_{i,r}\big(\overline{y}^{\mathcal{D}}_{i,r}\big). \tag{A.6}$$

Therefore, from $(A.5)$, $(A.6)$, and the fact that $max\Big\{\sum_{t=1}^{T}\sum_{r\in\mathcal{R}}p_{i,r}\mathbb{E}_{\mathbf{s}_t^{\pi(W)}}\Big[h_{i,r}\big(y_{t,i,r}^{\pi(W)}(t,\mathbf{s}_t^{\pi(W)})\big)\Big],C_i\Big\}\geq \sum_{t=1}^{T}\sum_{r\in\mathcal{R}}p_{i,r}\mathbb{E}_{\mathbf{s}_t^{\pi(W)}}\Big[h_{i,r}\big(y_{t,i,r}^{\pi(W)}(t,\mathbf{s}_t^{\pi(W)})\big)\Big]$, we can say

$$max\Big\{\sum_{t=1}^{T}\sum_{r\in\mathcal{R}}p_{i,r}\mathbb{E}_{\mathbf{s}_t^{\pi(W)}}\Big[h_{i,r}\big(y_{t,i,r}^{\pi(W)}(t,\mathbf{s}_t^{\pi(W)})\big)\Big],C_i\Big\}\geq max\Big\{T\sum_{r\in\mathcal{R}}p_{i,r}h_{i,r}\big(\overline{y}_{i,r}^{\mathcal{D}}\big),C_i\Big\}$$

By summing over $i\in W$, we have

$$\sum_{i\in W}max\Big\{\sum_{t=1}^{T}\sum_{r\in\mathcal{R}}p_{i,r}\mathbb{E}_{\mathbf{s}_t^{\pi(W)}}\Big[h_{i,r}\big(y_{t,i,r}^{\pi(W)}(t,\mathbf{s}_t^{\pi(W)})\big)\Big],C_i\Big\}\geq \sum_{i\in W}max\Big\{T\sum_{r\in\mathcal{R}}p_{i,r}h_{i,r}\big(\overline{y}_{i,r}^{\mathcal{D}}\big),C_i\Big\}$$

Hence the optimal objective function value of Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha},W)$ is a lower bound on the optimal objective function value of Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha},W)$.

- **Case 2**. $0<\gamma\leq 1-\hat{y}_{i,r}^{\mathcal{D}}$.

Under this case, $\overline{y}_{i,r}^{\mathcal{D}}=\hat{y}_{i,r}^{\mathcal{D}}+\gamma$. From $(A.4)$ and $\overline{y}_{i,r}^{\mathcal{D}}>\hat{y}_{i,r}^{\mathcal{D}}$, we have $T\sum_{i\in W}p_{i,r}\overline{y}_{i,r}^{\mathcal{D}}>T\sum_{i\in W}p_{i,r}\hat{y}_{i,r}^{\mathcal{D}}\geq \xi_r$, which satisfies constraint $(E.1)$ of Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha},W)$. Until now, we have shown that the solution $\overline{y}_{i,r}^{\mathcal{D}}$ under case 2 is feasible solution for Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha},W)$. Now we need to show that the objective function value of this solution in Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha},W)$ is lower than that of $y_{t,i,r}^{\pi(W)}(t,\mathbf{s}_t^{\pi(W)})$ in Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha},W)$. From (A.3), we have

$$T\sum_{r\in\mathcal{R}}p_{i,r}h_{i,r}(\overline{y}_{i,r}^{\mathcal{D}})=C_i.$$

Thus, we have

$$max\Big\{\sum_{t=1}^{T}\sum_{r\in\mathcal{R}}p_{i,r}\mathbb{E}_{\mathbf{s}_t^{\pi(W)}}\Big[h_{i,r}\big(y_{t,i,r}^{\pi(W)}(t,\mathbf{s}_t^{\pi(W)})\big)\Big],C_i\Big\}\geq max\Big\{T\sum_{r\in\mathcal{R}}p_{i,r}h_{i,r}(\overline{y}_{i,r}^{\mathcal{D}}),C_i\Big\}.$$

By summing over $i\in W$, we have

$$\sum_{i\in W}max\Big\{\sum_{t=1}^{T}\sum_{r\in\mathcal{R}}p_{i,r}\mathbb{E}_{\mathbf{s}_t^{\pi(W)}}\Big[h_{i,r}\big(y_{t,i,r}^{\pi(W)}(t,\mathbf{s}_t^{\pi(W)})\big)\Big],C_i\Big\}\geq \sum_{i\in W}max\Big\{T\sum_{r\in\mathcal{R}}p_{i,r}h_{i,r}(\overline{y}_{i,r}^{\mathcal{D}}),C_i\Big\}.$$

Therefore, the optimal objective function value of Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha},W)$ is a lower bound on the optimal objective function value of Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha},W)$ under case 2. Hence, the proof of the first claim is complete.

Now, we show the proof of the second claim. Let us define the policy $\pi^*(W)$ as $\pi^*(W):=y_{t,i,r}^{\pi^*(W)}(t,\mathbf{s}_t^{\pi^*(W)})=y_{i,r}^{*\mathcal{D}}\quad\forall t$. By substituting the policy defined above in (2), it is clear that it is feasible for Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha},W)$. It is easy to see that the objective function value of this policy for Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha},W)$ is $T\sum\limits_{i\in W}\sum\limits_{r\in\mathcal{R}}p_{i,r}h_{i,r}(y_{i,r}^{*\mathcal{D}})$ which completes the proof of the optimality of policy $\pi^*(W)$ for Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha},W)$. The above analysis leads to the following result. □

Proof of Proposition 2.

Let us define solution $\mathbf{y}^*$, with components $y^*_{i,r} = y^*_r,\ \forall r$, in which the expected number of impressions won at any location through the time horizon is equal to the demand at that location. That is, the constraint set $(E.1)$ is tight in solution $\mathbf{y}^*$, or

$$T\omega p_r y^*_r = \alpha_r, \qquad \forall r \in \mathcal{R}. \tag{A.7}$$

Here, our claim is to show that under the identical ad-exchanges case, feasible solution $\mathbf{y}^*$ whose components $y^*_{i,r}$ have the same value y^*_r is optimal solution to Problem $\mathcal{P}^{\mathcal{I}}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$.

Suppose to the contrary, we assume there exists a feasible solution $\breve{\mathbf{y}}$ to the Problem $\mathcal{P}^{\mathcal{I}}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$ which satisfies $T\sum_{i\in W} p_r \breve{y}_{i,r} = \alpha_r,\ \forall r$ and for any given location $r \in \mathcal{R}$, all components $\breve{y}_{i,r}$ have the same value $\breve{y}_r$ except two of them. i.e.,

$$\exists \quad (j,l) \in W \quad s.t. \quad \breve{y}_{j,r} \neq \breve{y}_{l,r} \neq \breve{y}_r \qquad \forall r \in \mathcal{R}.$$

Without loss of generality, we assume that $\breve{y}_{j,r} < \breve{y}_{l,r}$.

Let us define a new feasible solution $\grave{\mathbf{y}}$ whose r^{th} column is defined as

$$\begin{cases} \grave{y}_{i,r} = \breve{y}_{i,r}, \quad \forall i \in \Big\{ W - \{j,l\} \Big\}, \\ \grave{y}_{j,r} = \breve{y}_{j,r} + \varsigma, \\ \grave{y}_{l,r} = \breve{y}_{l,r} - \varsigma, \end{cases}$$

where $\varsigma > 0$. It is easy to verify that the solution $\grave{\mathbf{y}}$ satisfies the constraint set $(E.1)$ as equality. To complete the proof, we need to show that

$$\sum_{\substack{i\in W\\ i\neq j,l}} \max\Big\{\sum_{r\in\mathcal{R}} p_r h_r\big(\grave{y}_{i,r}\big), C\Big\} + \max\Big\{p_r h_r\big(\grave{y}_{j,r}\big), C\Big\} + \max\Big\{p_r h_r\big(\grave{y}_{l,r}\big), C\Big\} <$$
$$\sum_{\substack{i\in W\\ i\neq j,l}} \max\Big\{\sum_{r\in\mathcal{R}} p_r h_r\big(\breve{y}_{i,r}\big), C\Big\} + \max\Big\{p_r h_r\big(\breve{y}_{j,r}\big), C\Big\} + \max\Big\{p_r h_r\big(\breve{y}_{l,r}\big), C\Big\}.$$

It is sufficient to show that

$$\max\Big\{p_r h_r\big(\grave{y}_{j,r}\big), C\Big\} + \max\Big\{p_r h_r\big(\grave{y}_{l,r}\big), C\Big\} < \max\Big\{p_r h_r\big(\breve{y}_{j,r}\big), C\Big\} + \max\Big\{p_r h_r\big(\breve{y}_{l,r}\big), C\Big\}.$$

We substitute the value of $\grave{y}_{j,r}$ and $\grave{y}_{l,r}$ in the above equation, and we get

$$\max\Big\{p_r h_r\big(\breve{y}_{j,r} + \varsigma\big), C\Big\} + \max\Big\{p_r h_r\big(\breve{y}_{l,r} - \varsigma\big), C\Big\} < \max\Big\{p_r h_r\big(\breve{y}_{j,r}\big), C\Big\} + \max\Big\{p_r h_r\big(\breve{y}_{l,r}\big), C\Big\}.$$

and then

$$\max\Big\{p_r h_r\big(\breve{y}_{j,r} + \varsigma\big), C\Big\} - \max\Big\{p_r h_r\big(\breve{y}_{j,r}\big), C\Big\} < \max\Big\{p_r h_r\big(\breve{y}_{l,r}\big), C\Big\} - \max\Big\{p_r h_r\big(\breve{y}_{l,r} - \varsigma\big), C\Big\}.$$

We divide both sides by ς and obtain

$$\frac{\max\Big\{p_r h_r\big(\breve{y}_{j,r}+\varsigma\big),C\Big\}-\max\Big\{p_r h_r\big(\breve{y}_{j,r}\big),C\Big\}}{\varsigma}<\frac{\max\Big\{p_r h_r\big(\breve{y}_{l,r}\big),C\Big\}-\max\Big\{p_r h_r\big(\breve{y}_{l,r}-\varsigma\big),C\Big\}}{\varsigma}.$$

We can rewrite the above as

$$\frac{\max\Big\{p_r h_r\big(\breve{y}_{j,r}+\varsigma\big),C\Big\}-\max\Big\{p_r h_r\big(\breve{y}_{j,r}\big),C\Big\}}{\big(\breve{y}_{j,r}+\varsigma\big)-\breve{y}_{j,r}}<\frac{\max\Big\{p_r h_r\big(\breve{y}_{l,r}\big),C\Big\}-\max\Big\{p_r h_r\big(\breve{y}_{l,r}-\varsigma\big),C\Big\}}{\breve{y}_{l,r}-\big(\breve{y}_{l,r}-\varsigma\big)}. \tag{A.8}$$

Since function $h_r(y)$ is strictly increasing and strictly convex in y, then function $\max\Big\{p_r h_r(y),C\Big\}$ is strictly increasing and strictly convex in y. We also know that $\breve{y}_{j,r}<\grave{y}_{j,r}<\grave{y}_{l,r}<\breve{y}_{l,r}$, then equation $(A.8)$ holds. □

Proof of Theorem 4.

The proof is same as that of Theorem 1 except that when constructing an equivalent problem $\mathcal{P}_E(\Omega)$, we replace the probabilistic constraint $\beta=1/(mn)$, and $\xi_r=1$, $\forall i,r$ by the expectation constraint $\xi_r+z_\beta\sqrt{\xi_r}=1/(mn)$, $\forall i,r$. □

Proof of Lemma 1.

First we need to show that function $\mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big]$ is strictly increasing in ω; i.e.,

$$\frac{\partial\mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega}>0 \tag{A.9}$$

It is clear that

$$\frac{\partial\mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega}=\left(\frac{\partial\mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega\sum_{r\in\mathcal{R}}p_r}\right)\left(\frac{\partial\omega\sum_{r\in\mathcal{R}}p_r}{\partial\omega}\right)$$

We have

$$\frac{\partial\mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega\sum_{r\in\mathcal{R}}p_r}=\sum_{k=0}^{T}g(k)\binom{T}{k}\Big(\omega\sum_{r\in\mathcal{R}}p_r\Big)^{k-1}\Big(1-\omega\sum_{r\in\mathcal{R}}p_r\Big)^{T-k-1}\Big(k-T\omega\sum_{r\in\mathcal{R}}p_r\Big)$$

Since function $g(.)$ is convex, then (A.9) holds.

Now, we just need to prove that function $\mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big]$ is strictly convex in ω; i.e.,

$$\frac{\partial\mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega^2}>0$$

We define

$$\frac{\partial\mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega^2}=\left(\frac{\partial\mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial(\omega\sum_{r\in\mathcal{R}}p_r)^2}\right)\Big(\sum_{r\in\mathcal{R}}p_r\Big)^2 \tag{A.10}$$

Now, we only need to find the sign of the first component in Eq. (A.10). As we know, K_ω follows a binomial distribution with trial number T and success probability $\omega \sum_{r\in\mathcal{R}} p_r$. Hence, the expected value of function $g(K_\omega)$ is calculated as

$$\mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big] = \sum_{k=0}^{T} g(k)\binom{T}{k}\Big(\omega\sum_{r\in\mathcal{R}} p_r\Big)^k\Big(1-\omega\sum_{r\in\mathcal{R}} p_r\Big)^{T-k} \tag{A.11}$$

Now, we need to differentiate twice from Eq. (A.11) with respect to $\omega \sum_{r\in\mathcal{R}} p_r$.

$$\frac{\partial \mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial(\omega\sum_{r\in\mathcal{R}} p_r)^2} =$$

$$\sum_{k=0}^{T} g(k)\binom{T}{k}\Big((T-k)(T-k-1)p^k(1-p)^{T-k-2} + k(k-1)p^{k-2}(1-p)^{T-k} - 2k(T-k)p^{k-1}(1-p)^{T-k-1}\Big)$$

where $p = \sum_{r\in\mathcal{R}} p_r$. We reindex the three components separately with substitutions $k = j$, $k = j+2$, and $k = j+1$, respectively. Then we have

$$\frac{\partial \mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial(\omega\sum_{r\in\mathcal{R}} p_r)^2} = \sum_{j=0}^{T-2}\Big(g(j+2)-2g(j+1)+g(j)\Big)\Big(\frac{T!}{j!(T-j-2)!}\Big)\Big((\omega\sum_{r\in\mathcal{R}} p_r)^j(1-\omega\sum_{r\in\mathcal{R}} p_r)^{T-j-2}\Big)$$

Since function $g(.)$ is convex, then we can say that $\frac{\partial \mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial(\omega\sum_{r\in\mathcal{R}} p_r)^2} > 0$ and hence we can say that function $\mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]$ is convex in ω.

Now, we prove that the function $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ is strictly convex in ω. It is sufficient to show that $\frac{\partial f_1^{\mathcal{I}^2}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)}{\partial\omega^2} > 0$. The second derivative of function $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ with respect to ω is defined as

$$\frac{\partial f_1^{\mathcal{I}^2}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)}{\partial\omega^2} =$$

$$\sum_{r\in\mathcal{R}}\left(\frac{\partial{\psi_r}^2\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}{\partial\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)^2}\left(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega^2}\right)^2 + \frac{\partial\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}{\partial\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}\frac{2(\xi_r + z_\beta\sqrt{\xi_r})^2}{(Tp_r\omega^3)}\right) + \frac{\partial \mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega^2}$$

Since function $h_r(y) = y\psi_r(y)$ is strictly increasing and strictly convex in y and $\mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big]$ is strictly increasing and strictly convex in ω, the above equation is positive and we can say that function $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ is strictly convex in ω.

Now, we prove that the function $f_2^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ is strictly increasing and convex in ω. It is sufficient to show that $\frac{\partial f_2^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)}{\partial\omega} > 0$ and $\frac{\partial f_2^{\mathcal{I}^2}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)}{\partial\omega^2} > 0$, respectively. The first derivative of function $f_2^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ is defined as

$$\frac{\partial f_2^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)}{\partial\omega} = C + \frac{\partial \mathbb{E}_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega}.$$

Now, we need to prove that

$$\frac{\partial f_2^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)}{\partial\omega} > 0.$$

The second derivative of function $f_2^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ with respect to ω is defined as

$$\frac{\partial f_2^{\mathcal{I}^2}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)}{\partial\omega^2} = \frac{\partial\mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega^2}.$$

From Lemma 1, we know that $\frac{\partial\mathbb{E}^2_{K_\omega}\Big[g(K_\omega)\Big]}{\partial\omega^2} > 0$. Therefore, function $f_2^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ is convex in ω. □

Proof of Theorem 6.

Here, we focus on inner solution. Let us first consider a case in which $\min\{\varrho,\varpi\}=\varpi$. Hence, for $\omega\leq\varrho$, function $f_{\mathcal{R}}^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)=f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ and we know that function $f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ changes its direction at point ϖ such that before this point the function is decreasing and after it, it is increasing in ω. Therefore, the function $f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ in two closest integer points $\lfloor\varpi\rfloor$ and $\lceil\varpi\rceil$ has smaller values than any other integer points smaller than $\lfloor\varpi\rfloor$ or greater than $\lceil\varpi\rceil$. We only need to check the value of the function $f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ in two integer points $\lfloor\varpi\rfloor$ and $\lceil\varpi\rceil$ and pick that point in which the function has a smaller value. Now, let us consider case $\min\{\varrho,\varpi\}=\varrho$. Under this case, the function $f_2^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ intersects the decreasing part of function $f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ for the first time at point ϱ which is before point ϖ. Hence, for $\omega\leq\varrho$, function $f_{\mathcal{R}}^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)=f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ and is decreasing in ω. Similarly, we only need to check the value of function $f_{\mathcal{R}}^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ at points $\lfloor\varrho\rfloor$ and $\lceil\varrho\rceil$ and pick that one which gives a smaller value to the function. □

Proof of Proposition 3.

We consider there are n existing incumbent ad-exchanges in the market. Let us define ω^b as the number of the ad-exchanges that the ad-firm works with before the entry of a new ad-exchange, and ω^a is the number of the ad-exchanges that the ad-firm works with after the entry. If $\omega^a=\omega^b$, then after the entry of a new ad-exchange, the expected capacity cost is equal to $\mathbb{E}_{K_{\omega^a}}\Big[g\big(K_{\omega^a}\big)\Big]$ and the expected capacity cost before the entry is $\mathbb{E}_{K_{\omega^b}}\Big[g\big(K_{\omega^b}\big)\Big]$. Since function g(.) is strictly increasing and convex, then we have

$$\mathbb{E}_{K_{\omega^a}}\Big[g\big(K_{\omega^a}\big)\Big]=\mathbb{E}_{K_{\omega^b}}\Big[g\big(K_{\omega^b}\big)\Big].$$

Therefore, the revenue of the cloud provider is not affected by the entry of the ad-change.

After the entry, the winning probability at location r is equal to $y_r^a=\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega^a}$ and winning probability at location r before the entry is equal to $y_r^b=\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega^b}$. Since $\omega^a=\omega^b$, we have $y_r^a=y_r^b$. Since the bid-curve is strictly increasing and convex in y, then the expected bidding cost at y_r^a is smaller than the expected bidding cost at y_r^b. If the expected bidding cost at y_r^a is greater than the participation fee, then the procurement cost after the entry is equal to the bidding cost which is equal to the procurement cost before the entry. If the expected bidding cost at y_r^a is less than participation fee, the procurement cost after the entry is equal to the participation fee, which is again equal to the procurement cost before the entry, i.e., $\max\{C, y^b\psi(y^b)\} = \max\{C, y^a\psi(y^a)\}$. Thus, the total revenue of the ad-exchanges in the market is same as the one before the entry. However, after the entry, the average revenue for each ad-exchange reduces from $\frac{\max\{C, y^b\psi(y^b)\}}{n}$ to $\frac{\max\{C, y^a\psi(y^a)\}}{n+1}$.

If $\omega^a > \omega^b$, since function g(.) is strictly increasing and convex, then we have

$$\mathbb{E}_{K_{\omega^a}}\Big[g\big(K_{\omega^a}\big)\Big] \geq \mathbb{E}_{K_{\omega^b}}\Big[g\big(K_{\omega^b}\big)\Big].$$

This means that the cloud provider gain more revenue after the entry of a new ad-exchange. We also have $y_r^a \leq y_r^b$. Since the bid-curve is strictly increasing and convex, then after the entry, the bidding cost at each incumbent ad-exchange is less than before the ad-exchange. If the bidding cost before the entry is less than the participation fee, then the procurement cost is unaffected by the entry. If it is above the participation fee, then the procurement cost after the entry is less than before the entry. Thus, we have $\max\{C, y^b\psi(y^b)\} \geq \max\{C, y^a\psi(y^a)\}$. For the new ad-exchanges, they clearly earn some money. However, the average revenue for each ad-exchange decreases $\frac{\max\{C, y^b\psi(y^b)\}}{n}$ to $\frac{\max\{C, y^a\psi(y^a)\}}{n+1}$. □

Proof of Lemma 4.

It is clear when ξ_r increases, the value of $\xi_r + z_\beta\sqrt{\xi_r}$ also increases. Hence, for proving that function $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ is increasing in ξ_r for any given ω, we only need to show that the function is increasing in $\xi_r + z_\beta\sqrt{\xi_r}$. For any given ω, we know that function $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ is continuous in $\xi_r + z_\beta\sqrt{\xi_r}$. Now, we need to show that

$$\frac{\partial f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)}{\partial(\xi_r + z_\beta\sqrt{\xi_r})} > 0. \tag{A.12}$$

We have

$$\frac{\partial f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)}{\partial(\xi_r + z_\beta\sqrt{\xi_r})} = \psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big) + \left(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega}\right)\left(\frac{\partial\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}{\partial\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}\right).$$

Since function $h_r(y) = y\psi_r(y)$ is non-negative and increasing in y, then we can say that $(A.12)$ holds and we conclude that function $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ is increasing in ξ_r $\forall r \in \mathcal{R}$.

Now, we need to prove the convexity of function $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\sqrt{\boldsymbol{\xi}}, \omega)$ in $\xi_r + z_\beta\sqrt{\xi_r}$ for any given $\omega \in \{1, 2, ..., N\}$.We only need to show that

$$\frac{\partial f_1^{\mathcal{I}^2}(\boldsymbol{\xi} + z_\beta\sqrt{\boldsymbol{\xi}}, \omega)}{\partial(\xi_r + z_\beta\sqrt{\xi_r})^2} > 0 \tag{A.13}$$

We have

$$\frac{\partial f_1^{\mathcal{I}^2}(\boldsymbol{\xi}+z_\beta\sqrt{\boldsymbol{\xi}},\omega)}{\partial(\xi_r+z_\beta\sqrt{\xi_r})^2}=$$

$$\left(\frac{1}{Tp_r\omega}\right)\left(\frac{\partial\psi_r^2\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}{\partial\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)^2}+\frac{\partial\psi_r\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}{\partial\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}+\frac{\partial\psi_r^2\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}{\partial\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)^2}\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)\right)$$

Since function $h_r(y)=y\psi_r(y)$ is non-negative, strictly increasing and, convex in y, $(A.13)$ holds. □

Proof of Lemma 5.

We know that

$$\frac{\partial f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)}{\partial(\xi_r+z_\beta\sqrt{\xi_r})}=\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)+\left(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\right)\left(\frac{\partial\psi_r\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}{\partial\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)}\right).$$

Now let us consider two values ω_1 and ω_2 such that $\omega_1<\omega_2$. We can see that

$$\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega_2}<\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega_1}. \tag{A.14}$$

Since function $h_r(y)=y\psi_r(y)$ is strictly increasing and convex in y, we can conclude

$$\left.\frac{\partial f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)}{\partial(\xi_r+z_\beta\sqrt{\xi_r})}\right|_{\omega=\omega_2}<\left.\frac{\partial f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)}{\partial(\xi_r+z_\beta\sqrt{\xi_r})}\right|_{\omega=\omega_1}. \tag{A.15}$$

□

Proof of Lemma 6.

Let us assume that the ad-firm considers to work with two following values number of ad-exchanges ω: ω_1 and ω_2 such that $\omega_1<\omega_2$. Suppose to contrary we assume that $\zeta_{\omega_1}>\zeta_{\omega_2}$. Let us now pick demand value, ζ which is in interval $[\zeta_{\omega_2},\zeta_{\omega_1}]$, i.e., $\zeta_{\omega_2}\le\zeta\le\zeta_{\omega_1}$. We can say that for given demand ζ, the expected total cost of working with ω_1 number of ad-exchanges is equal to $f_2^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega_1)$. It means

$$\sum_{r\in\mathcal{R}}\zeta\psi_r\Big(\frac{\zeta}{Tp_r\omega_1}\Big)\le\omega_1C.$$

Similarly, the expected total cost for ω_2 is equal to $f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega_2)$. It means that

$$\sum_{r\in\mathcal{R}}\zeta\psi_r\Big(\frac{\zeta}{Tp_r\omega_2}\Big)\ge\omega_2C. \tag{A.16}$$

Since function $\psi_r(y)$ is increasing in y and $\omega_2<\omega_1$, we have

$$\sum_{r\in\mathcal{R}}\zeta\psi_r\Big(\frac{\zeta}{Tp_r\omega_1}\Big)\le\sum_{r\in\mathcal{R}}\zeta\psi_r\Big(\frac{\zeta}{Tp_r\omega_2}\Big). \tag{A.17}$$

We know that $\omega_2C\le\omega_1C$. Hence, equation $(A.16)$ cannot be held and it is a contradiction. Therefore, when ω increases, the value of ζ_ω increases.

Proof of Proposition 4.

Proof of part (i). When demand at location r, ξ_r, increases, the value of $\xi_r + z_\beta\sqrt{\xi_r}$ also increases. Then we need to study the impact of $\xi_r + z_\beta\sqrt{\xi_r}$ on ω^{up}. From Theorem 6, we know that ω^{up} takes different value under different conditions.

Let us start with case (i) in Theorem 6. This case happens when $\sum_{r\in\mathcal{R}}(\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\dfrac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r}\Big) \leq C$. Therefore, when $\xi_r + z_\beta\sqrt{\xi_r}$ increases, if the condition still holds, i.e., the expected bidding cost incurred by working with minimum available number of ad-exchanges is still less than or equal to participation fee, then the near-optimal number of ad-exchanges is still 1 and does not change. If it does not hold, it means that either the condition of case (ii) holds or that of case (iii). It is clear that under both cases the near-optimal number of ad-exchanges is larger. Then, we can conclude when demand $\xi_r + z_\beta\sqrt{\xi_r}$ increase, ω^{up} calculated in case (i) in Theorem 6 stays the same or increases.

Now, we consider case (ii) in Theorem 6. When demand $\xi_r + z_\beta\sqrt{\xi_r}$ increases, the condition of case (ii) never violates. Since function $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \omega)$ is strictly increasing in $\xi_r + z_\beta\sqrt{\xi_r}$. In fact, when demand at location r increases, the expected bidding cost incurred by working with maximum available number of ad-exchanges N is still above participation fee. Therefore, from case (ii), we don't go to other cases. Therefore, now we need to show how $\omega^{up} = \arg\min_{\omega}\Big\{f_{\mathcal{R}}^{\mathcal{I}}\big(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \lfloor\varpi\rfloor\big), f_{\mathcal{R}}^{\mathcal{I}}\big(\boldsymbol{\xi} + z_\beta\boldsymbol{\xi}_o, \lceil\varpi\rceil\big)\Big\}$ changes when demand $\xi_r + z_\beta\sqrt{\xi_r}$ increases.

Let us first obtain the impact of $\xi_r + z_\beta\sqrt{\xi_r}$ on ϖ. Since both $\lfloor\varpi\rfloor$ and $\lceil\varpi\rceil$ are increasing in ϖ, then we can get the useful insights about how $\xi_r + z_\beta\sqrt{\xi_r}$ affects ω^{up}. Therefore, we need to get the sign of $\dfrac{\mathrm{d}\varpi}{\mathrm{d}(\xi_r + z_\beta\sqrt{\xi_r})}$. From equation (15), let $\Gamma(\varpi, p_r, \xi_r, T) = \sum_{r\in\mathcal{R}}(\xi_r + z_\beta\sqrt{\xi_r})\left(\dfrac{\partial\psi_r\big(y_r^*(\varpi)\big)}{\partial y_r^{*\mathcal{D}}(\varpi)}\;\dfrac{\mathrm{d}y_r^*(\varpi)}{\mathrm{d}\varpi}\right) + \dfrac{\partial\mathbb{E}_{K_\varpi}\Big[g(K_\varpi)\Big]}{\partial\varpi} = 0$. Using the Implicit Function Theorem, see for example, Rudin (1976), we have

$$\frac{\mathrm{d}\varpi}{\mathrm{d}(\xi_r + z_\beta\sqrt{\xi_r})} = -\frac{\dfrac{\partial\Gamma}{\partial(\xi_r + z_\beta\sqrt{\xi_r})}}{\dfrac{\partial\Gamma}{\partial\varpi}}, \tag{A.18}$$

where $\dfrac{\partial\Gamma}{\partial(\xi_r + z_\beta\sqrt{\xi_r})}$ and $\dfrac{\partial\Gamma}{\partial\varpi}$ are calculated as below:

$$\frac{\partial\Gamma}{\partial(\xi_r + z_\beta\sqrt{\xi_r})} = -\left(\frac{\partial\psi_r\big(y_r^*(\varpi)\big)}{\partial y_r^*(\varpi)}\right)\left(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\varpi^2}\right).$$

Since the function $h_r(y) = y\psi_r(y)$ is strictly increasing and convex, then we have $\dfrac{\partial\Gamma}{\partial(\xi_r + z_\beta\sqrt{\xi_r})} < 0$.

$$\frac{\partial\Gamma}{\partial\varpi} = \sum_{r\in\mathcal{R}}\left(\frac{\partial{\psi_r}^2\Big(y_r^*(\varpi)\Big)}{\partial y_r^*(\varpi)^2}\;\frac{(\xi_r + z_\beta\sqrt{\xi_r})^3}{T^2p_r^2\varpi^4} + \frac{\partial\psi_r\Big(y_r^*(\varpi)\Big)}{\partial y_r^*(\varpi)}\;\frac{2(\xi_r + z_\beta\sqrt{\xi_r})^2}{Tp_r\varpi^3}\right) + \frac{\partial\mathbb{E}_{K_\varpi}^2\Big[g(K_\varpi)\Big]}{\partial\varpi^2}.$$

From Lemma 1, we know that function $f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\varpi)$ is strictly convex in ω Therefore, we can say that the second derivative at point ϖ must be positive, i.e., $\frac{\partial {f_1^{\mathcal{I}}}^2(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)}{\partial\omega^2}\Big|_{\omega=\varpi}>0$. Then, we have $\frac{\partial\Gamma}{\partial\varpi}>0$ and $\frac{\mathrm{d}\varpi}{\mathrm{d}(\xi_r+z_\beta\sqrt{\xi_r})}>0$. From the above results and Lemma 5, we can say when $\xi_r+z_\beta\sqrt{\xi_r}$ increases, ω^{up} stays the same or increases.

Now, let us consider case (iii) in Theorem 6. When demand $\xi_r+z_\beta\sqrt{\xi_r}$ increases, the condition of case (iii) might hold or violate. If it does not hold, because of increasing property of function $f_1^{\mathcal{I}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\omega)$ in $\xi_r+z_\beta\sqrt{\xi_r}$ as shown in Lemma 4, condition described for case (ii) in Theorem 6 might occur. If that happens, we can see that the near-optimal number of ad-exchanges cannot decrease. If the condition of case (iii) does not violate, we need to show how $\arg\min_\omega\Big\{f_{\mathcal{R}}^{\mathcal{I}}\big(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\lfloor\min\{\varrho,\varpi\}\rfloor\big),f_{\mathcal{R}}^{\mathcal{I}}\big(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\lceil\min\{\varrho,\varpi\}\rceil\big)\Big\}$ changes when demand $\xi_r+z_\beta\sqrt{\xi_r}$ increases. If $\min\{\varrho,\varpi\}=\varpi$, we have already shown that ϖ is increasing in $\xi_r+z_\beta\sqrt{\xi_r}$. Therefore, ω^{up} stays the same or increases.

But if $\min\{\varrho,\varpi\}=\varrho$, from Theorem 6, we have $\omega^{up}=\arg\min_\omega\Big\{f_{\mathcal{R}}^{\mathcal{I}}\big(\lfloor\varrho\rfloor\big),f_{\mathcal{R}}^{\mathcal{I}}\big(\lceil\varrho\rceil\big)\Big\}$. Now, we need to examine the effect of $\xi_r+z_\beta\sqrt{\xi_r}$ on ω^{up}. Let us first obtain the impact of $\xi_r+z_\beta\sqrt{\xi_r}$ on ϱ. Since both $\lfloor\varrho\rfloor$ and $\lceil\varrho\rceil$ are increasing in ϱ, then we can get the useful insights about how $\xi_r+z_\beta\sqrt{\xi_r}$ affects ω^{up}. Therefore, we need to get the sign of $\frac{\mathrm{d}\varrho}{\mathrm{d}(\xi_r+z_\beta\sqrt{\xi_r})}$. From definition of ϱ, let $\Upsilon(\varrho,p_r,\xi_r,T)=\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\varrho}\Big)-\varrho C=0$. Using the Implicit Function Theorem, we have

$$\frac{\mathrm{d}\varrho}{\mathrm{d}(\xi_r+z_\beta\sqrt{\xi_r})}=-\frac{\frac{\partial\Upsilon}{\partial(\xi_r+z_\beta\sqrt{\xi_r})}}{\frac{\partial\Upsilon}{\partial\varrho}},\tag{A.19}$$

where $\frac{\partial\Upsilon}{\partial(\xi_r+z_\beta\sqrt{\xi_r})}$ and $\frac{\partial\Upsilon}{\partial\varrho}$ are calculated as below:

$$\frac{\partial\Upsilon}{\partial(\xi_r+z_\beta\sqrt{\xi_r})}=\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\varrho}\Big)$$

Since the function $\psi_r(y)$ is nonnegative, then we have $\frac{\partial\Upsilon}{\partial(\xi_r+z_\beta\sqrt{\xi_r})}\geq 0$.

$$\frac{\partial\Upsilon}{\partial\varrho}=-\sum_{r\in\mathcal{R}}\left(\frac{\partial\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\varrho}\Big)}{\partial\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\varrho}\Big)}\;\frac{(\xi_r+z_\beta\sqrt{\xi_r})^2}{Tp_r\varrho^2}\right)-C.$$

Since the function $\psi_r(y)$ is strictly increasing, then we have $\frac{\partial\Upsilon}{\partial\varrho}<0$. Hence, $\frac{\mathrm{d}\varrho}{\mathrm{d}(\xi_r+z_\beta\sqrt{\xi_r})}\geq 0$. From the above results, we can say when $(\xi_r+z_\beta\sqrt{\xi_r})$ increases ω^{up} stays the same or increases.

Now, we need to show how $y_r^*(\omega^{up})=\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}$ changes when demand $\xi_r+z_\beta\sqrt{\xi_r}$ increases. From above analyses, we can see when $\xi_r+z_\beta\sqrt{\xi_r}$ increases, the near-optimal number of ad-exchanges ω^{up} stays the same or increases. If ω^{up} stays the same, it is clear that $y_r^*(\omega^{up}) = \dfrac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}$ increases when demand at location r increases.

Proof of part (ii). We have $y_r^{*\mathcal{D}}(\omega^{up}) = \dfrac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}$. When the demand at location r increases and the near-optimal number of ad-exchanges stays the same, it is trivial that the value of $y_r^{*\mathcal{D}}(\omega^{up})$ increases. When the near-optimal number of ad-exchanges increases from ω^{up} to $\omega^{up}+1$, we say there exists a positive number ϵ such that $\lim_{\epsilon\to 0}\dfrac{\xi_r + \epsilon + z_\beta\sqrt{\xi_r+\epsilon}}{Tp_r(\omega^{up}+1)} = \dfrac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r(\omega^{up}+1)}$. Therefore, it is clear that $\dfrac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r(\omega^{up}+1)} < \dfrac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}$. Then we can say when the near-optimal number of ad-exchanges increases then, the winning probability decreases.

Proof of part (iii). Suppose to the contrary we assume that there exist demand level D_r for location r and a small number ϵ such that if the demand at location r increases from D_r to $D_r+\epsilon$, then the expected total cost decreases. Let us assume when demand at location r is D_r, the ad-firm works with ω_1 number of ad-exchanges and when the demand at location r is $D_r+\epsilon$, the ad-firm works with ω_2 number of ad-exchanges. Therefore, under demand D_r, we have

$$f_{\mathcal{R}}^{\mathcal{I}}(D_r,\omega_1) < f_{\mathcal{R}}^{\mathcal{I}}(D_r,\omega_2)$$

and under demand $D_r+\epsilon$, we have

$$f_{\mathcal{R}}^{\mathcal{I}}(D_r+\epsilon,\omega_1) > f_{\mathcal{R}}^{\mathcal{I}}(D_r+\epsilon,\omega_2)$$

From above results, we have that ω_1 is not optimal solution for problem with demand level D_r, since ω_2 can satisfy the demand constraint under a larger demand at location r (i.e., $D_r+\epsilon$) and causes a lower cost which is a contradiction.

Now, we need to show when demand at location r increases from D_r to $D_r+\epsilon$ (we assume that ϵ is a very small number), the number of ad-exchanges stays the same if the participation fee is relatively large, otherwise it increases.

Let us first assume when demand at location r is equal to D_r, the ad-firm works with ω^{up} number of ad-exchanges. If demand increases to $D_r+\epsilon$ and expected bidding cost of working with the same number of ad-exchanges (i.e., ω^{up}) is smaller than participation fee, i.e., $\sum_{r\in\mathcal{R}}(D_r+\epsilon)\psi_r(\dfrac{D_r+\epsilon}{Tp_r\omega^{up}}) < \omega^{up}C$, then we have

$$\sum_{r\in\mathcal{R}} D_r\psi_r(\frac{D_r}{Tp_r\omega^{up}}) < \omega^{up}C.$$

Since bidding cost is increasing in demand at location r. Therefore, the ad-firm keeps working with the same number of ad-exchanges ω^{up}.

Now let us assume when demand at location r is D_r, the ad-firm works with $\omega^{up} = \lceil\varpi\rceil$, then the expected bidding cost is

$$\sum_{r\in\mathcal{R}} D_r\psi_r(\frac{D_r}{Tp_r\omega^{up}}) > \omega^{up}C.$$

Since the expected bidding cost is increasing in demand at location r, then we have

$$\sum_{r\in\mathcal{R}}(D_r+\epsilon)\psi_r\big(\frac{(D_r+\epsilon)}{Tp_r\omega^{up}}\big) > \omega^{up}C.$$

From the above result and the fact that the expected total cost is strictly increasing in ω for any $\omega > \lceil\varpi\rceil$, then we can say that the ad-firm works with the same number of ad-exchanges when $\frac{\sum_{r\in\mathcal{R}}(D_r+\epsilon)\psi_r\big(\frac{(D_r+\epsilon)}{Tp_r\omega^{up}}\big)}{\omega^{up}} > C$.

Now let us assume when demand at location r is D_r, the ad-firm works with $\omega^{up} < \lceil\omega\rceil$ number of ad-exchanges and $\frac{\sum_{r\in\mathcal{R}}D_r\psi_r\big(\frac{D_r}{Tp_r\omega^{up}}\big)}{\omega^{up}} > C$. Since the expected bidding cost is decreasing in ω for any given $\omega < \omega^{up}$, we have

$$\sum_{r\in\mathcal{R}}(D_r+\epsilon)\psi_r\big(\frac{(D_r+\epsilon)}{Tp_r\omega}\big) > \omega^{up}C.$$

Under the current case, the ad-firm keeps working with the same number of ad-exchanges if the expected total cost of working with $\omega^{up}+1$ number of ad-exchanges is larger. Therefore, under the following condition the ad-firm works with the same number of ad-exchange when the demand at location r increases.

$$C \le \frac{\mathbb{E}_{K_{\omega^{up}}}\Big[g\big(K_{\omega^{up}}\big)\Big] - \mathbb{E}_{K_{(\omega^{up}+1)}}\Big[g\big(K_{(\omega^{up}+1)}\big)\Big] + \sum_{r\in\mathcal{R}}(D_r+\epsilon)\psi_r\big(\frac{(D_r+\epsilon)}{Tp_r\omega}\big)}{\omega^{up}+1}.$$

Therefore, the ad-firm works with more ad-exchanges if

$$C \le \sum_{r\in\mathcal{R}}(D_r+\epsilon)\psi_r\big(\frac{(D_r+\epsilon)}{Tp_r\omega}\big),$$

and

$$\frac{\mathbb{E}_{K_{\omega^{up}}}\Big[g\big(K_{\omega^{up}}\big)\Big] - \mathbb{E}_{K_{(\omega^{up}+1)}}\Big[g\big(K_{(\omega^{up}+1)}\big)\Big] + \sum_{r\in\mathcal{R}}(D_r+\epsilon)\psi_r\big(\frac{(D_r+\epsilon)}{Tp_r\omega}\big)}{\omega^{up}+1} \le C.$$

□

Proof of Proposition 5.

Here we first want to show that the number of ad-exchanges which the ad-firm works with under the case where a fraction of supply of ad-space is considered for bid computation cannot be less than that under the existing strategy. On other hand, we are showing that

(i) $\omega^{up-F} = \omega^{up}$, if $\rho^*(\omega^{up-F}) = 1$

(ii) $\omega^{up-F} \ge \omega^{up}$, if $\rho^*(\omega^{up-F}) < 1$

The proof of part (i) is trivial. Let us first assume that ω^{up} is calculated from case (ii) in Theorem 6. Now, we want to prove that $\omega^{up-F} \ge \omega^{up}$. To proceed the proof, we first need to show that the absolute difference between the expected total cost at point $(y_r^{F^*},\rho^*(\omega^{up-F}),\omega)$ and expected total cost at point $(y_r^*,1,\omega)$ is increasing in ω. We need to show that

$$\Bigg| \sum_{r\in\mathcal{R}} (\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega_2}\Big) + \mathbb{E}_{K_{\omega_2}}\Big[g\big(K_{\omega_2}\big)\Big] - \sum_{r\in\mathcal{R}} (\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega_2}\Big) - \mathbb{E}_{K_{\omega_2}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega_2}\big)\Big]\Bigg|$$
$$-\Bigg| \sum_{r\in\mathcal{R}} (\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega_1}\Big) + \mathbb{E}_{K_{\omega_1}}\Big[g\big(K_{\omega_1}\big)\Big] - \sum_{r\in\mathcal{R}} (\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega_1}\Big) - \mathbb{E}_{K_{\omega_1}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega_1}\big)\Big]\Bigg| \geq 0 \tag{A.20}$$

where $|.|$ denotes the absolute value.

To show that $(A.20)$ holds, we prove that for any given ω_1 and ω_2 where $\omega_1 \leq \omega_2$,

$$\Bigg(\mathbb{E}_{K_{\omega_2}}\Big[g\big(K_{\omega_2}\big)\Big] - \mathbb{E}_{K_{\omega_2}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega_2}\big)\Big]\Bigg) - \Bigg(\mathbb{E}_{K_{\omega_1}}\Big[g\big(K_{\omega_1}\big)\Big] - \mathbb{E}_{K_{\omega_1}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega_1}\big)\Big]\Bigg) \geq 0, \tag{A.21}$$

and

$$\Bigg(\sum_{r\in\mathcal{R}} (\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega_2}\Big) - \sum_{r\in\mathcal{R}} (\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega_2}\Big)\Bigg) - \Bigg(\sum_{r\in\mathcal{R}} (\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega_1}\Big) - \sum_{r\in\mathcal{R}} (\xi_r + z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega_1}\Big)\Bigg) \leq 0, \tag{A.22}$$

First, we prove that $(A.21)$ holds. The left side of (A.21) is equal to

$$\sum_{k=0}^{T}\binom{T}{k}\Big(\sum_{r\in\mathcal{R}} p_r\Big)^k\Big(g(k) - g(\rho^*(\omega^{up-F})k)\Big)\Bigg(\omega_2^k\Big(1-\omega_2\sum_{r\in\mathcal{R}} p_r\Big)^{T-k} - \omega_1^k\Big(1-\omega_1\sum_{r\in\mathcal{R}} p_r\Big)^{T-k}\Bigg)$$

Since function $g(.)$ is strictly increasing and convex, then $(A.21)$ holds.

Now, we need to prove that $(A.22)$ holds. Since $\omega_1 \leq \omega_2$, it is clear that

$$\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega_2} \leq \frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega_1}$$

and

$$\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega_2} \leq \frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega_1}$$

Therefore, we have

$$\frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega_2} - \frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega_2} \leq \frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega_1} - \frac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega_1}$$

Since function $h_r(.)$ is strictly increasing and convex, then $(A.22)$ holds. Until now, we have shown that the absolute difference between the expected total cost at point $(y_r^{F^*}$,$\rho^*(\omega^{up-F})$,$\omega^{up})$ and expected total cost at point $(y_r^*,1,\omega^{up})$ is greater than the absolute difference between the expected total cost at point $(y_r^{F^*},\rho^*(\omega^{up-F}),\omega^{up}-1)$ and expected total cost at point $(y_r^*,1,\omega^{up}-1)$. Since the expected total cost at point $(y_r^*,1,\omega^{up})$ is smaller than the expected total cost at point $(y_r^*,1,\omega^{up}-1)$, we can conclude that $\omega^{up-F}\geq\omega^{up}$.

Let us assume that ω^{up} is calculated from case (iii) in Theorem 6. If $\min\{\varrho,\varpi\}=\varpi$, the proof is similar to the previous proof. If $\min\{\varrho,\varpi\}=\varrho$ and $\omega^{up}=\lfloor\varrho\rfloor$, the proof is again similar to the previous proof. But if $\min\{\varrho,\varpi\}=\varrho$ and $\omega^{up}=\lceil\varrho\rceil$, we need to show that

$$\left(\max\left\{\omega^{up}C,\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega^{up}}\Big)\right\}-\omega^{up}C\right)$$
$$-\left(\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})(\omega^{up}-1)}\Big)-\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r(\omega^{up}-1)}\Big)\right)\leq 0, \tag{A.23}$$

If $\max\left\{\omega^{up}C,\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega^{up}}\Big)\right\}=\omega^{up}C$, then $(A.23)$ holds. If $\max\left\{\omega^{up}C,\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega^{up}}\Big)\right\}=\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega^{up}}\Big)$, from $(A.22)$ and the fact that $\omega^{up}C\leq\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r(\omega^{up}-1)}\Big)$, we can say that $(A.23)$ holds.

Now, we show when $\rho^*(\omega^{up-F})<1$, the number of ad-exchanges which the ad-firm works with under selective bidding is

(i) $\omega^{up-F}=\omega^{up}, \quad if \quad C\geq\chi_2$,

(ii) $\omega^{up-F}=\omega^{up}, \quad if \quad C\leq\chi_2$ and $\omega^{up}=\lceil\varpi\rceil$,

(iii) $\omega^{up-F}=\omega^{up}, \quad if \quad \chi_1\leq C\leq\chi_2$ and $\omega^{up}<\lceil\varpi\rceil$,

(iv) $\omega^{up-F}>\omega^{up}, \quad if \quad C\leq\chi_1$ and $\omega^{up}<\lceil\varpi\rceil$,

where $\chi_1\leq\chi_2$ and $\chi_1=\frac{\mathbb{E}_{K_{\omega^{up}}}\left[g\left(\rho^*(\omega^{up-F})K_{\omega^{up}}\right)\right]-\mathbb{E}_{K_{(\omega^{up}+1)}}\left[g\left(\rho^*(\omega^{up-F})K_{(\omega^{up}+1)}\right)\right]+\chi_2}{\omega^{up}+1}$ and $\chi_2=\frac{\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega^{up}}\Big)}{\omega^{up}}$.

(i) When $\rho^*(\omega^{up-F})<1$, since function $\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho\omega^{up}}\Big)$ is decreasing in ρ, if $\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega^{up}}\Big)<\omega^{up}C$, we have

$$\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}\Big)<\omega^{up}C$$

Therefore, from $(A.23)$, we can see that ω^{up-F} is same as ω^{up}.

(ii) When $\omega^{up}=\lceil\varpi\rceil$, it means that

$$\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}\Big)>\omega^{up}C$$

Since function $\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho\omega^{up}}\Big)$ is decreasing in ρ, if $\rho^*(\omega^{up-F})<1$, then we have

$$\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega^{up}}\Big)>\omega^{up}C$$

From the above results and the fact that expected total cost is strictly increasing in ω for any $\omega>\lceil\varpi\rceil$, we have $\omega^{up-F}=\omega^{up}$.

(iii) and (iv). Since the expected bidding cost is decreasing in ω, when $\omega^{up}<\lceil\varpi\rceil$ and $\omega^{up}C\le\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}\Big)$, for $\rho^*(\omega^{up-F})<1$ and $\omega\le\omega^{up}$, we have

$$\omega^{up}C\le\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega}\Big)$$

Therefore, the expected total cost at point ω^{up} is

$$\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\frac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega}\Big)+\mathbb{E}_{K_{\omega^{up}}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega^{up}}\big)\Big]$$

The ad-firm keeps working with the same number of ad-exchange ω^{up}, if the expected total cost of working with $\omega^{up}+1$ is greater. Therefore, if

$$C\le\frac{\mathbb{E}_{K_{\omega^{up}}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega^{up}}\big)\Big]-\mathbb{E}_{K_{(\omega^{up}+1)}}\Big[g\big(\rho^*(\omega^{up-F})K_{(\omega^{up}+1)}\big)\Big]+\chi_2}{\omega^{up}+1}$$

Then, the ad-firm works with larger number of ad-exchanges. □

Proof of Proposition 6.

Part (i) is trivial. Let us focus on the proof of parts (ii) and (iii). If $\rho^*(\omega^{up-F})=1$, then it is clear that $\omega^{up-F}=\omega^{up}$. Therefore, under the selective bidding, the expected capacity cost is equal to $\mathbb{E}_{K_{\omega^{up-F}}}\Big[g\big(K_{\omega^{up-F}}\big)\Big]$ which is equal to the expected capacity cost under the existing bidding strategy $\mathbb{E}_{K_{\omega^{up}}}\Big[g\big(K_{\omega^{up}}\big)\Big]$. The winning probability at location r is equal to $y_r^{F^*}=\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up-F}}$ which is equal to winning probability at location r under the existing bidding strategy $y_r^*=\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}$. Hence, the procurement cost is the same under both bidding strategies.

If $\rho^*(\omega^{up-F})<1$ and $\omega^{up-F}=\omega^{up}$, then under the case where a fraction of supply of ad-exchange is considered for bid computation, the expected capacity cost is equal to $\mathbb{E}_{K_{\omega^{up-F}}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega^{up-F}}\big)\Big]$ and the expected capacity cost under the existing bidding strategy is $\mathbb{E}_{K_{\omega^{up}}}\Big[g\big(K_{\omega^{up}}\big)\Big]$. Since $\omega^{up-F}=\omega^{up}$ and $\rho^*(\omega^{up-F})<1$ and function g(.) is strictly increasing and convex, then we have

$$\mathbb{E}_{K_{\omega^{up-F}}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega^{up-F}}\big)\Big]\le\mathbb{E}_{K_{\omega^{up}}}\Big[g\big(K_{\omega^{up}}\big)\Big].$$

Under the selective bidding, the winning probability at location r is equal to $y_r^{F^*} = \dfrac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\rho^*(\omega^{up-F})\omega^{up-F}}$ and winning probability at location r under the existing bidding strategy is equal to $y_r^* = \dfrac{\xi_r + z_\beta\sqrt{\xi_r}}{Tp_r\omega^{up}}$.Since $\omega^{up-F} = \omega^{up}$ and $\rho^*(\omega^{up-F}) < 1$. Then we have $\rho^*(\omega^{up-F})\omega^{up-F} \leq \omega^{up}$ which results in

$$y_r^* \leq y_r^{F^*}.$$

Since the bid-curve is strictly increasing and convex in y, then the expected bidding cost at y_r^* is smaller than the expected bidding cost at $y_r^{F^*}$. If the expected bidding cost at y_r^* is greater than the participation fee, then the procurement cost under the case where a fraction of supply of ad-exchange is considered for bid computation is larger than that under the existing bidding strategy. If the expected bidding cost at y_r^* is less than participation fee, the procurement cost under the selective bidding is either the same or larger.

If $\rho^*(\omega^{up-F}) < 1$ and $\omega^{up-F} > \omega^{up}$, since $\rho^*(\omega^{up-F})\omega^{up-F} \geq \omega^{up}$ and function g(.) is strictly increasing and convex, then we have

$$\mathbb{E}_{K_{\omega^{up-F}}}\Big[g\big(\rho^*(\omega^{up-F})K_{\omega^{up-F}}\big)\Big] \geq \mathbb{E}_{K_{\omega^{up}}}\Big[g\big(K_{\omega^{up}}\big)\Big].$$

We also have $y_r^{F^*} \leq y_r^*$. Since the bid-curve is strictly increasing and convex, then the expected bidding cost under the selective bidding is less than that under the existing bidding strategy for the existing ad-exchanges. If the expected bidding cost under the existing bidding strategy is less than the participation fee, then the procurement cost is the same under both bidding strategies. If it is above the participation fee, then the procurement cost is less under the selective bidding. For the new ad-exchanges, they clearly earn some money. $\square$

Appendix B. The Proposed Deterministic Problem in Section 3.1.1

For any given subset $W \subseteq \Omega$, we introduce a following *deterministic* Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha}, W)$ and then establish its equivalence to Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$; that is, we show that an optimal policy for Problem $\mathcal{P}_{R}(\boldsymbol{\alpha}, W)$ can be obtained from an optimal solution to Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha}, W)$. For proving that claim, we first show that the optimal objective function value of Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha}, W)$ is a lower bound for the optimal objective function value of the Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$. Then we need to show that there is a feasible solution for Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$ whose objective function value is same as the optimal objective function value of Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha}, W)$.

Problem $\mathcal{P}_{\mathcal{D}}(\boldsymbol{\alpha}, W)$

$$f_{\mathcal{D}}(\boldsymbol{\alpha}, W) := \min_{\mathbf{y}^{\mathcal{D}}} \quad \sum_{i \in W} max\left\{T \sum_{r \in \mathcal{R}} p_{i,r} h_{i,r}\big(y_{i,r}^{\mathcal{D}}\big), C_i\right\} + \mathbb{E}_{K_W}\Big[g\big(K_W\big)\Big]$$

subject to

$$T \sum_{i \in W} p_{i,r} y_{i,r}^{\mathcal{D}} \geq \alpha_r, \qquad \forall r, \tag{B.1}$$

$$y_{i,r}^{\mathcal{D}} \in [0,1], \qquad \forall i, r, \tag{B.2}$$

where superscript $\mathcal{D}$ denotes the problem with state-independent policy, $\mathbf{y}^{\mathcal{D}}$.

Appendix C. Optimization Under Selective Bidding

Let parameter ρ denote the supply fraction for which the ad-firm responds with a non-zero bid when it obtains its supply from the set of ad-exchanges W. The ad-firm's bidding policy is $\pi_F(W) := y_{t,i,r}^{\pi_F(W)}(t, \mathbf{s}_t^{\pi_F(W)})$, where the subscript F denotes selective bidding case.

Formally, the ad-firm's problem under selective bidding can be stated as below. Note the presence of ρ in the objective function and as a new decision variable.

Problem $\mathcal{P}_F(\Omega)$

$$\min_{W \subseteq \Omega, \rho} \quad f_F(\beta, \rho, W), \tag{C.1}$$

where function $f_F(\beta, \rho, W)$ is given by solving Problem $\mathcal{P}_F(\beta, \rho, W)$.

Problem $\mathcal{P}_F(\beta, \rho, W)$

$$f_F(\beta, \rho, W) := \min_{\pi_F(W)} \sum_{i \in W} max \left\{ \sum_{t=1}^{T} \sum_{r \in \mathcal{R}} \rho p_{i,r} \mathbb{E}_{\mathbf{S}_t^{\pi_F(W)}} \left[h_{i,r}\big(y_{t,i,r}^{\pi_F(W)}(t, \mathbf{s}_t^{\pi_F(W)})\big) \right], C_i \right\} + \mathbb{E}_{K_W} \left[g\big(\rho K_W\big) \right]$$

subject to:

$$\mathbb{P}\left[\mathbf{S}_{r,T+1}^{\pi_F(W)} \geq \xi_r \right] \geq \beta, \qquad \forall r, \tag{C.2}$$

$$y_{t,i,r}^{\pi_F(W)}(t, \mathbf{s}_t^{\pi_F(W)}) \in [0,1], \qquad \forall t, i, r, s_t^{\pi_F(W)}. \tag{C.3}$$

The solution to the above problem is similar to the case without selective bidding. The parameter ρ affects the probability of a bid-request arriving from an ad-exchange location combination: $p_{i,r}^F = \rho p_{i,r}$, $p_{i,r}^F$ is the probability with fractional supply. To find the optimal value of the fraction ρ, we need to solve the problem for various values of ρ between 0 and 1 and pick the best value. However, for obtaining structural results, we revert to the problem with identical ad-exchanges for further analysis.

Selective Bidding Under Identical Ad-exchanges In this section, we analyze selective bidding under identical ad-exchanges. Our analyses reveal that selective bidding always makes the ad-firm better off (never worse off), but the ad-exchanges or the cloud provider may be better off or worse off. For any given ω and ρ, the optimal solution of the following problem is a feasible solution for Problem $\mathcal{P}_F^{\mathcal{I}}(\beta, \rho, \omega)$.

Problem $\mathcal{P}_{F\mathcal{R}}^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, \rho, \omega)$

$$f_{F\mathcal{R}}^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, \rho, \omega) := \min_{\pi_F^{\mathcal{I}}(\omega)} \sum_{i=1}^{\omega} max \left\{ \sum_{t=1}^{T} \sum_{r \in \mathcal{R}} \rho p_r \mathbb{E}_{\mathbf{s}_t^{\pi_F^{\mathcal{I}}(\omega)}} \left[h_{i,r}\big(y_{t,i,r}^{\pi_F^{\mathcal{I}}(\omega)}(t, \mathbf{s}_t^{\pi_F^{\mathcal{I}}(\omega)})\big) \right], C \right\} + \mathbb{E}_{K_\omega} \left[g\big(\rho K_\omega\big) \right]$$

subject to:

$$\sum_{i=1}^{\omega} \sum_{t=1}^{T} \rho p_r \mathbb{E}_{\mathbf{s}_t^{\pi_F^{\mathcal{I}}(\omega)}} \left[y_{t,i,r}^{\pi_F^{\mathcal{I}}(\omega)}(t, \mathbf{s}_t^{\pi_F^{\mathcal{I}}(\omega)}) \right] \geq \xi_r + z_\beta \sqrt{\xi_r}, \qquad \forall r, \tag{C.4}$$

$$y_{t,i,r}^{\pi_F^{\mathcal{I}}(\omega)}(t, \mathbf{s}_t^{\pi_F^{\mathcal{I}}(\omega)}) \in [0,1], \qquad \forall t, i, r, s_t^{\pi_F^{\mathcal{I}}(\omega)}. \tag{C.5}$$

Similar to the proof of Proposition 1 and Proposition 2, for any given ω and ρ, it is shown that the optimal winning probabilities for Problem $\mathcal{P}^{\mathcal{I}}_{F\mathcal{R}}(\boldsymbol{\alpha},\rho,\omega)$, where $\boldsymbol{\alpha}=\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o$, is state-independent and all impressions arriving from location r are procured with the same probability which is equal to $y_r^{F^*}=\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho\omega}$. Then, the optimal expected total cost is $f^{\mathcal{I}}_{F\mathcal{R}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho,\omega)=\max\Big\{f^{\mathcal{I}}_{F1}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho,\omega),f^{\mathcal{I}}_{F2}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho,\omega)\Big\}$, where $f^{\mathcal{I}}_{F1}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho,\omega)=\sum\limits_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho\omega}\Big)+\mathbb{E}_{K_\omega}\Big[g\big(\rho K_\omega\big)\Big]$ and $f^{\mathcal{I}}_{F2}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho,\omega)=\omega C+\mathbb{E}_{K_\omega}\Big[g\big(\rho K_\omega\big)\Big]$.

Now, for any $\omega\in\{1,2,...,N\}$, we need to calculate the optimal value of ρ, i.e., ρ^*. Since function $f^{\mathcal{I}}_{F1}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho,\omega)$ is strictly convex in ρ, then there exists an unique point $\overline{\rho}$ in which $\dfrac{\partial f^{\mathcal{I}}_{F1}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho,\omega)}{\partial\rho}\Big|_{\overline{\rho}}=0$. We can also see that function $f^{\mathcal{I}}_{F2}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho,\omega)$ is strictly increasing in ρ. Hence, there might exist a point, Λ, in which $f^{\mathcal{I}}_{F1}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\Lambda,\omega)=f^{\mathcal{I}}_{F2}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\Lambda,\omega)$. The following lemma characterizes ρ^*. For any given $\omega\in\{1,2,...,N\}$, the optimal fraction, ρ^*, is

(i) $\min\{1,\overline{\rho}\}$, if $\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\omega}\Big)>C$,

(ii) $\min\{\Lambda,\overline{\rho}\}$, otherwise.

The Effect of Selective Bidding on Ad-exchange Selection By knowing the value of $y_r^{F^*}$ and ρ^*, we propose the following solution as a near-optimal solution to Problem $\mathcal{P}^{\mathcal{I}}_F(\Omega)$. A feasible solution to Problem $\mathcal{P}^{\mathcal{I}}_F(\Omega)$, ω^{up-F}, is

(i) 1, if $\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*}\Big)\le C$,

(ii) $\arg\min\limits_{\omega}\Big\{f^{\mathcal{I}}_{F\mathcal{R}}\big(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho^*,\lfloor\varpi^F\rfloor\big),f^{\mathcal{I}}_{F\mathcal{R}}\big(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho^*,\lceil\varpi^F\rceil\big)\Big\}$, if $\sum_{r\in\mathcal{R}}(\xi_r+z_\beta\sqrt{\xi_r})\psi_r\Big(\dfrac{\xi_r+z_\beta\sqrt{\xi_r}}{Tp_r\rho^*N}\Big)>C$,

(iii) $\arg\min\limits_{\omega}\Big\{f^{\mathcal{I}}_{F\mathcal{R}}\big(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho^*,\big\lfloor\min\{\varrho^F,\varpi^F\}\big\rfloor\big),f^{\mathcal{I}}_{F\mathcal{R}}\big(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho^*,\big\lceil\min\{\varrho^F,\varpi^F\}\big\rceil\big)\Big\}$, otherwise,

where ϱ^F is a point in which $f^{\mathcal{I}}_{F1}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho^*,\varrho^F)=f^{\mathcal{I}}_{F2}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho^*,\varrho^F)$ and ϖ^F is a point in which $\dfrac{\partial f^{\mathcal{I}}_{F1}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o,\rho^*,\varrho^F)}{\partial\omega}\Big|_{\omega=\varpi^F}=0$ when the integrality constraint on ω is removed.

The proof is similar to the proof of Theorem 6.

Appendix D: An Enhanced Hill Climbing Policy for Ad-Exchange Selection Problem

Here, we propose an approximation algorithm to solve the ad-exchanges selection problem for the *non-identical ad-exchanges* case. We first introduce Problem $\mathcal{P}_E(\Omega)$, although $\mathcal{P}_E(\Omega)$ is different from the original problem $\mathcal{P}(\Omega)$, we show that an approximate solution to $\mathcal{P}_E(\Omega)$ is also an approximate solution to $\mathcal{P}(\Omega)$. In the rest of this section, we focus on solving $\mathcal{P}_E(\Omega)$.

Problem $\mathcal{P}_E(\Omega)$

$$\min_{W \subseteq \Omega} \quad f_{\mathcal{R}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, W),$$

In the rest of this section, we focus on providing an approximate solution to Problem $\mathcal{P}_E(\Omega)$. Based on Proposition 1, we know that for the choice $\boldsymbol{\alpha} = \boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o$, Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$ can be solved optimally for any given W. Given an optimal solution $y^{*\mathcal{D}}_{i,r}$ to $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$, if there exists any $i \in W$ such that $T\sum_{r\in\mathcal{R}} p_{i,r} y^{*\mathcal{D}}_{i,r} \psi_{i,r}\big(y^{*\mathcal{D}}_{i,r}\big) < C_i$, we increase the value of $y^{*\mathcal{D}}_{i,r}$ for an arbitrary r such that $T\sum_{r\in\mathcal{R}} p_{i,r} y^{*\mathcal{D}}_{i,r} \psi_{i,r}\big(y^{*\mathcal{D}}_{i,r}\big) = C_i$. It is easy to verify that the new solution is still optimal since (1) all demands can still be satisfied due to the original solution is feasible and the new solution increases some bidding amounts of the original solution, and (2) for every $i \in \Omega$, the cost $\max\{T\sum_{r\in\mathcal{R}} p_{i,r} y^{*\mathcal{D}}_{i,r} \psi_{i,r}\big(y^{*\mathcal{D}}_{i,r}\big), C_i\}$ does not increase. In the rest of this paper, we assume that $\forall i \in \Omega, T\sum_{r\in\mathcal{R}} p_{i,r} y^{*\mathcal{D}}_{i,r} \psi_{i,r}\big(y^{*\mathcal{D}}_{i,r}\big) \geq C_i$. Consider a fixed W, we can partition its cost $f_{\mathcal{R}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, W)$ into two parts: hard cost $f_{\text{hard}}(W) = \sum_{i\in W} C_i + \mathbb{E}_{K_W}\Big[g\big(K_W\big)\Big]$ and soft cost $f_{\text{soft}}(W) = T\sum_{i\in W}\sum_{r\in\mathcal{R}} p_{i,r} y^{*\mathcal{D}}_{i,r} \psi_{i,r}\big(y^{*\mathcal{D}}_{i,r}\big) - \sum_{i\in W} C_i$. Notice that for any fixed W, the hard cost $f_{\text{hard}}(W)$, which is composed of participation fees and capacity cost, is fixed, regardless of the bidding strategy. It is easy to observe that $f_{\text{hard}}(W)$ is a generalization of $f_2^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, \omega)$ as defined in the identical case. Similarly, the soft cost $f_{\text{soft}}(W)$ can be considered as a modification of $f_1^{\mathcal{I}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, \omega)$, as defined in the identical case, by excluding the capacity cost. Now let us assume that the optimal solution to $\mathcal{P}_E(\Omega)$ contracts with W^*, let $B^* = f_{\text{hard}}(W^*)$ denote the hard cost of W^*. Problem $\mathcal{P}_E(\Omega)$ can be rewritten as:

$$\begin{aligned} &\min_{W\subseteq\Omega} \quad f_{\text{soft}}(W) \\ &\text{subject to:} \quad f_{\text{hard}}(W) \leq B^*, \end{aligned} \tag{D.1}$$

Let $\theta^* = \arg\min_{i\in W^*} f_{\text{soft}}(\{i\})$ denote the mobile ad-exchange in W^* that has the minimum soft cost. Given θ^*, the above problem can be further rewritten as:

Problem $\mathcal{P}_{\mathcal{K}}(\theta^*, B^*)$

$$\begin{aligned} &\max_{W\subseteq\Omega} \quad f_{\text{soft}}(\{\theta^*\}) - f_{\text{soft}}(\{\theta^*\} \cup W) \\ &\text{subject to:} \quad f_{\text{hard}}(\{\theta^*\} \cup W) \leq B^* \end{aligned} \tag{D.2}$$

It is easy to verify that the objective function $f_{\text{soft}}(\{\theta^*\}) - f_{\text{soft}}(\{\theta^*\} \cup W)$ is an non-decreasing set function of W and $f_{\text{hard}}(\{\theta^*\} \cup W)$ is also an non-decreasing set function of W. Thus, Problem $\mathcal{P}_{\mathcal{K}}(\theta^*, B^*)$ is to maximize an non-decreasing objective function subject to a budget constraint. Although θ^* and the value of B^* are not known in advance, we can guess their value through partial enumeration. Given the guess θ and B, we then try to find a feasible solution with θ and the budget constraint equal to B using a hill climbing algorithm.

Let $B_{\max} = f_{\text{hard}}(\Omega)$ be the largest possible budget constraint. For a given guess θ, let $\Delta_\theta = \{f_{\text{hard}}(\{\theta\})(1+\varepsilon)^l, l = 0, \cdots, L_\theta\}$ where $L_\theta = \lceil \log \frac{B_{\max}}{f_{\text{hard}}(\{\theta\})\varepsilon} \rceil$ and $\varepsilon \geq 0$ is a turning parameter that controls the tradeoff between the time complexity and the performance of our policy. To simplify notation, define $\phi_\theta(W) = f_{\text{soft}}(\theta) - f_{\text{soft}}(\{\theta\} \cup W)$, i.e., $\phi_{\theta^*}(W)$ is the objective function of Problem $\mathcal{P}_{\mathcal{K}}(\theta^*, B^*)$. Notice that for any given W and θ, the value of $f_{\text{soft}}(\{\theta\} \cup W)$, thus the value of $\phi_\theta(W)$, can be obtained using the solution proposed in 3.1.1. We next present the design of our policy φ^{EHC}.

1. (Line 1-2 in Algorithm 1) Estimating θ^* and B^* through enumeration;
2. (Line 3-9 in Algorithm 1) Given an estimation of θ^* and B^*, say θ and B, we solve $\mathcal{P}_{\mathcal{K}}(\theta, B)$ using a simple hill climbing algorithm. Our algorithm works round by round. At each round j, we add the mobile ad-exchange with the largest benefit-to-cost ratio to existing solution, i.e., assume W_{j-1} is the solution returned from round $j-1$, in the next round j, we add $\arg\max_{i \in \Omega} \frac{\phi_\theta(W_{j-1} \cup \{i\}) - \phi_\theta(W_{j-1})}{f_{\text{hard}}(\{\theta\} \cup W_{j-1} \cup \{i\}) - f_{\text{hard}}(\{\theta\} \cup W_{j-1})}$ to W_{j-1} (the value of $\phi_\theta(W_{j-1} \cup \{i\})$ and $\phi_\theta(W_{j-1})$ can be obtained using the solution proposed in Section 3.1.1);
3. (Line 10 in Algorithm 1) After enumerating all θ and B, return the solution with the minimum cost.

Algorithm 1 $\mathsf{EnhancedHillClimbing}$ Policy φ^{EHC}

```
1:  for θ ∈ Ω do
2:      for B ∈ Δ_θ do
3:          W_{θ,B} = {θ}; S = Ω_{≥θ}; { Ω_{≥θ} = {i | i ∈ Ω, f_soft(i) ≥ f_soft(θ)}}
4:          while S ≠ ∅ do
5:              i* ← arg max_{i∈S} (φ_θ(W∪{i}) − φ_θ(W)) / (f_hard(W∪{i}) − f_hard(W)); {call the solution proposed in Section 3.1.1 to calculate
                the value of φ_θ(W_{j−1} ∪ {i}) and φ_θ(W_{j−1})}
6:              if f_hard(W ∪ {i*}) > B then
7:                  S = S \ {i*};
8:              else
9:                  W_{θ,B} ← W_{θ,B} ∪ {i*}; S ← S \ {i*};
10: return W_{θ,B} that minimizes f_R(ξ + z_β ξ_o, W_{θ,B}) over θ ∈ Ω and B ∈ Δ_θ
```

We next briefly explain the intuition behind our algorithm. Notice that given a fixed θ and B, in each round of step 2, we select the ad-exchange that maximizes $\frac{\phi_\theta(W_{j-1}\cup\{i\})-\phi_\theta(W_{j-1})}{f_{\text{hard}}(\{\theta\}\cup W_{j-1}\cup\{i\})-f_{\text{hard}}(\{\theta\}\cup W_{j-1})}$ where $\phi_\theta(W_{j-1}\cup\{i\})-\phi_\theta(W_{j-1})$ captures the amount of saved bidding cost after adding i and $f_{\text{hard}}(\{\theta\}\cup W_{j-1}\cup\{i\})-f_{\text{hard}}(\{\theta\}\cup W_{j-1})$ captures the increased participation fee and capacity cost after adding i. In other words, during each round of selection, we prefer the ad-exchange that maximizes the ratio of the saved bidding cost and the increased participation fee plus capacity cost. A detailed description of our policy φ^{EHC} is presented in Algorithm 1. Because of Theorem 2 and the fact that the solution returned from φ^{EHC} satisfies constraint explained in Appendix B, we have the following theorem.

Performance Guarantee of $\mathsf{EnhancedHillClimbing}$ **Policy** φ^{EHC}**:** Next we provide a performance bound of the policy φ^{EHC} for $\mathcal{P}(\Omega)$. Assume W^{EHC} is the final solution returned from φ^{EHC}. Let $Cost(\varphi^{EHC}) = f_{\mathcal{R}}(\boldsymbol{\xi}+z_\beta\boldsymbol{\xi}_o, W^{EHC})$ represent the cost incurred by φ^{EHC} and $Cost(\varphi^{opt})$ be the cost of the optimal solution φ^{opt} to $\mathcal{P}(\Omega)$.

To give a lower bound on $Cost(\varphi^{opt})$, we introduce a new problem $\mathcal{P}'_E(\Omega)$. It is clear that the optimal solution to $\mathcal{P}'_E(\Omega)$ is a lower bound of $Cost(\varphi^{opt})$.

Problem $\mathcal{P}'_E(\Omega)$

$$\min_{W\subseteq\Omega} \quad f_{\mathcal{R}}(\beta\boldsymbol{\xi}, W),$$

Given any W, we define its soft cost under $\mathcal{P}'_E(\Omega)$ as $f'_{\text{soft}}(W) = T\sum_{i\in W}\sum_{r\in\mathcal{R}} p_{i,r} y^{*'\mathcal{D}}_{i,r}\psi_{i,r}\big(y^{*'\mathcal{D}}_{i,r}\big) - \sum_{i\in W} C_i$ where $y^{*'\mathcal{D}}_{i,r}$ is the optimal solution to $\mathcal{P}_{\mathcal{R}}(\beta\boldsymbol{\xi}, W)$. Let $W^{*'}$ denote the optimal solution to $\mathcal{P}'_E(\Omega)$ and $\theta^{*'} = \arg\min_{i\in W^{*'}} f'_{\text{soft}}(\{i\})$. We apply φ^{EHC} to solve $\mathcal{P}'_E(\Omega)$ and assume $W^{EHC'}$ is the final solution. We use W^G to denote an intermediate solution found by φ^{EHC} when $\theta = \theta^{*'}$ and $B \in [f_{\text{hard}}(W^{*'}), (1+\epsilon)f_{\text{hard}}(W^{*'})]$. We next give the performance bound of φ^{EHC}.

The policy φ^{EHC} achieves the following performance guarantee for Problem $\mathcal{P}(\Omega)$:

$$\frac{Cost(\varphi^{EHC})}{Cost(\varphi^{opt})} \leq \frac{\eta f_{\mathcal{R}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, W^{EHC})}{f_{\mathcal{R}}(\beta \boldsymbol{\xi}, W^{EHC'})}, \tag{D.3}$$

where $\eta = (\eta_4 - (1 - e^{-\frac{\eta_3}{\eta_1 \eta_2}})(1 - \eta_4) + (1 + \epsilon)\eta_5)\frac{1}{1+\eta_5}$, $\eta_1 = \min_{i \in \Omega, Z \subseteq \Omega} \frac{f'_{\text{soft}}(\emptyset) - f'_{\text{soft}}(\{i\})}{f'_{\text{soft}}(Z) - f'_{\text{soft}}(\{i\} \cup Z)}$, $\eta_2 = \min_{i \in \Omega, Z \subseteq \Omega} \frac{f_{\text{hard}}(\{i\}) - f_{\text{hard}}(\emptyset)}{f_{\text{hard}}(\{i\} \cup Z) - f_{\text{hard}}(Z)}$, $\eta_3 = \frac{f_{\text{hard}}(W^G)}{(1+\epsilon) f_{\text{hard}}(W^{*'})}$, $\eta_4 = \frac{f'_{\text{soft}}(\{\theta^{*'}\})}{f'_{\text{soft}}(W^{*'})}$, and $\eta_5 = \frac{f_{\text{hard}}(W^{*'})}{f'_{\text{soft}}(W^{*'})}$. Assuming we use stochastic gradient descent algorithm (Bottou 2010) to solve Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$, the time complexity of φ^{EHC} is bounded by $O(L_{\max} N^4 R/\epsilon)$ where $L_{\max} = \max_{\theta \in \Omega} L_\theta$ and ϵ is the accuracy of the stochastic gradient descent algorithm.

To give a lower bound on the optimal solution, we solve $\mathcal{P}'_E(\Omega)$ approximately using φ^{EHC}. For a given θ, we redefine $\Omega_{\geq \theta}$ as the set of all ad-exchanges whose soft cost is no smaller than $f'_{\text{soft}}(\theta)$. Redefine $\phi_\theta(W) = f'_{\text{soft}}(\{\theta\}) - f'_{\text{soft}}(\{\theta\} \cup W)$. Define the curvature of function $\phi_\theta(W)$ (resp. $f_{\text{hard}}(W)$) to be $\eta_1 = \min_{i \in \Omega, Z \subseteq \Omega} \frac{\phi_\theta(\{i\}) - \phi_\theta(\emptyset)}{\phi_\theta(\{i\} \cup Z) - \phi_\theta(Z)}$ (resp. $\eta_2 = \min_{i \in \Omega, Z \subseteq \Omega} \frac{f_{\text{hard}}(\{i\}) - f_{\text{hard}}(\emptyset)}{f_{\text{hard}}(\{i\} \cup Z) - f_{\text{hard}}(Z)}$). We first show that for any guess θ and B, our policy (Line 3-9 in Algorithm 1) is able to find an approximation solution to the following problem. Our proof is based on some standard techniques studied in the context of submodular optimization (Nemhauser et al. 1978).

Problem $\mathcal{P}_{\mathcal{K}}(\theta, B)$

$$\max_{W \subseteq \Omega} \quad \phi_\theta(W)$$
$$\text{subject to:} \quad f_{\text{hard}}(\{\theta\} \cup W) \leq B \tag{D.4}$$

Since θ is fixed, we use $\phi(W)$ to denote $\phi_\theta(W)$ for short. Let $O = \{o_i, i \in [1, m]\}$ denote the optimal solution to $\mathcal{P}_{\mathcal{K}}(\theta, B)$ and W_j denote the solution returned from φ^{EHC} at round j. For a given function f, define $f(A; X) = f(X \cup A) - f(X)$ as the marginal increase of f with respect to A and X. We have

$$\begin{aligned}
\phi(O) &\leq \phi(W_j \cup O) \\
&\leq \phi(W_j) + \phi(o_1; W_j) + \phi(o_1 \cup o_2; W_j \cup o_1) + \cdots + \phi(W_j \cup o_1 \cup \cdots \cup o_m; W_j \cup o_1 \cup \cdots \cup o_{m-1}) \\
&\leq \phi(W_j) + \eta_1 \phi(o_1; W_j) + \eta_1 \phi(o_2; W_j) + \cdots + \eta_1 \phi(o_m; W_j)
\end{aligned}$$

Thus,

$$\frac{1}{\eta_1}(\phi(O) - \phi(W_j)) \leq \sum_{i \in [1,m]} \phi(o_i; W_j) \tag{D.5}$$

Because $\frac{\phi(o_i; W_j)}{f_{\text{hard}}(o_i; W_j)} \leq \frac{\phi(w_{j+1}; W_j)}{f_{\text{hard}}(w_{j+1}; W_j)}$, we have

$$\phi(o_i; W_j) \leq f_{\text{hard}}(o_i; W_j) \frac{\phi(w_{j+1}; W_j)}{f_{\text{hard}}(w_{j+1}; W_j)}$$

Thus,

$$\sum_{i\in[1,m]}\phi(o_i;W_j)\le\left(\sum_{i\in[1,m]}f_{\text{hard}}(o_i;W_j)\right)\frac{\phi(w_{j+1};W_j)}{f_{\text{hard}}(w_{j+1};W_j)}\le\eta_2B\frac{\phi(w_{j+1};W_j)}{f_{\text{hard}}(w_{j+1};W_j)} \tag{D.6}$$

(D.5) and (D.6) together imply that

$$\frac{1}{\eta_1\eta_2}\frac{\phi(O)-\phi(W_j)}{B}\le\frac{\phi(w_{j+1};W_j)}{f_{\text{hard}}(w_{j+1};W_j)}$$

Recall that $W_{\theta,B}$ is the solution found by φ^{EHC} for any guess θ and B, we have $\phi(W_{\theta,B})\ge(1-e^{-\frac{f_{\text{hard}}(W_{\theta,B})}{\eta_1\eta_2B}})\phi(O)$.

Let W^G denote the solution found by φ^{EHC} when $\theta=\theta^{*'}$ and $B\in[f_{\text{hard}}(W^{*'}),(1+\epsilon)f_{\text{hard}}(W^{*'})]$ (this is guaranteed to happen through partial enumeration), assume that $\eta_3=\frac{f_{\text{hard}}(W^G)}{(1+\epsilon)f_{\text{hard}}(W^{*'})}$, we have

$$\begin{aligned}f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^G)&=f'_{\text{soft}}(\{\theta^{*'}\})-\phi(\theta^{*'},W^G)+f_{\text{hard}}(\{\theta^{*'}\}\cup W^G)\\&\le f'_{\text{soft}}(\{\theta^{*'}\})-(1-e^{-\frac{\eta_3}{\eta_1\eta_2}})\phi(\theta^{*'},W^{*'})+(1+\epsilon)f_{\text{hard}}(W^{*'})\\&=f'_{\text{soft}}(\{\theta^{*'}\})-(1-e^{-\frac{\eta_3}{\eta_1\eta_2}})\left(f'_{\text{soft}}(\{\theta^{*'}\})-f'_{\text{soft}}(\{\theta^{*'}\}\cup W^{*'})\right)+(1+\epsilon)f_{\text{hard}}(W^{*'})\end{aligned}$$

Assume that $f'_{\text{soft}}(\{\theta^{*'}\})=\eta_4f'_{\text{soft}}(W^{*'})$, $f_{\text{hard}}(W^{*'})=\eta_5f'_{\text{soft}}(W^{*'})$, and the optimal solution to the original problem $\mathcal{P}(\Omega)$ contracts with W^{opt}, it follows that

$$\begin{aligned}f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^G)&\le(\eta_4-(1-e^{-\frac{\eta_3}{\eta_1\eta_2}})(1-\eta_4)+(1+\epsilon)\eta_5)f'_{\text{soft}}(W^{*'})\\&\le(\eta_4-(1-e^{-\frac{\eta_3}{\eta_1\eta_2}})(1-\eta_4)+(1+\epsilon)\eta_5)\frac{1}{1+\eta_5}f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^{*'})\\&\le(\eta_4-(1-e^{-\frac{\eta_3}{\eta_1\eta_2}})(1-\eta_4)+(1+\epsilon)\eta_5)\frac{1}{1+\eta_5}f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^{opt})\end{aligned} \tag{D.7}$$

Assume that $W^{EHC'}$ is the final solution obtained from φ^{EHC} after solving $\mathcal{P}'_E(\Omega)$. Recall that W^G denotes the intermediate solution found by φ^{EHC}, it is clear that

$$f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^{EHC'})\le f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^G) \tag{D.8}$$

Let

$$\eta=(\eta_4-(1-e^{-\frac{\eta_3}{\eta_1\eta_2}})(1-\eta_4)+(1+\epsilon)\eta_5)\frac{1}{1+\eta_5} \tag{D.9}$$

According to (D.7), we have

$$f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^G)\le\eta f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^{opt}) \tag{D.10}$$

(D.8) and (D.10) together imply that

$$f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^{EHC'})\le\eta f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^{opt}) \tag{D.11}$$

It is clear that

$$Cost(\varphi^{opt})\ge f_{\mathcal{R}}(\beta\boldsymbol{\xi},W^{opt}) \tag{D.12}$$

(D.11) and (D.12) imply that $Cost(\varphi^{opt}) \geq \frac{1}{\eta} f_{\mathcal{R}}(\beta\boldsymbol{\xi}, W^{EHC'})$. Since $Cost\big(\varphi^{EHC}\big) = f_{\mathcal{R}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, W^{EHC})$, we have

$$\frac{Cost\big(\varphi^{EHC}\big)}{Cost(\varphi^{opt})} \leq \frac{\eta f_{\mathcal{R}}(\boldsymbol{\xi} + z_\beta \boldsymbol{\xi}_o, W^{EHC})}{f_{\mathcal{R}}(\beta\boldsymbol{\xi}, W^{EHC'})} \tag{D.13}$$

.

We next analyze the time complexity of φ^{EHC}. Notice that we need to enumerate at most $NL_{\max}$ possible pairs of (θ, B). Given a pair of (θ, B), our hill climbing algorithm takes $O(N^2)$ rounds to find a group of ad-exchanges, this is because the size of the group is at most N and it takes $O(N)$ steps to find an ad-exchange that maximizes the benefit-to-cost ratio. Each round requires solving an instance of Problem $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$. If we use the stochastic gradient descent algorithm (Bottou 2010) to solve $\mathcal{P}_{\mathcal{R}}(\boldsymbol{\alpha}, W)$, the time complexity is $O(NR/\epsilon)$ where ϵ is the accuracy of the stochastic gradient descent algorithm. Thus the time complexity of φ^{EHC} is $O(L_{\max} N^4 R/\epsilon)$.

According to the definition of η, both η_1 and η_2 have negative effect on the performance bound. Based on the definition of submodular function (Lovász 1983), we know that if $-f'_{\text{soft}}(W)$ and $f_{\text{hard}}(W)$ are submodular functions then $\eta_1 = 1$ and $\eta_2 = 1$. It follows that as $-f'_{\text{soft}}(W)$ and $f_{\text{hard}}(W)$ approach submodular functions, our solution is closer to the optimal solution. We conjecture that $-f'_{\text{soft}}(W)$ is monotone and submodular, e.g., the marginal bidding cost saved by adding one more ad-exchange is non-increasing.

Appendix E. The Proposed Deterministic Problem in Section 5.2

For any given subset $W \subseteq \Omega$, we introduce a following *deterministic* Problem $\mathcal{P}_{\Delta_{\mathcal{D}}}(\boldsymbol{\alpha}, W)$ and then establish its equivalence to Problem $\mathcal{P}_{\Delta_{\mathcal{R}}}(\boldsymbol{\alpha}, W)$; that is, we show that an optimal policy for Problem $\mathcal{P}_{\Delta_{\mathcal{R}}}(\boldsymbol{\alpha}, W)$ can be obtained from an optimal solution to Problem $\mathcal{P}_{\Delta_{\mathcal{D}}}(\boldsymbol{\alpha}, W)$. For proving that claim, we first show that the optimal objective function value of Problem $\mathcal{P}_{\Delta_{\mathcal{D}}}(\boldsymbol{\alpha}, W)$ is a lower bound for the optimal objective function value of the Problem $\mathcal{P}_{\Delta_{\mathcal{R}}}(\boldsymbol{\alpha}, W)$. Then we need to show that there is a feasible solution for Problem $\mathcal{P}_{\Delta_{\mathcal{R}}}(\boldsymbol{\alpha}, W)$ whose objective function value is same as the optimal objective function value of Problem $\mathcal{P}_{\Delta_{\mathcal{D}}}(\boldsymbol{\alpha}, W)$.

Problem $\mathcal{P}_{\Delta_{\mathcal{D}}}(\boldsymbol{\alpha}, W)$

$$\min_{\mathbf{y}^{\mathcal{D}}} \quad \sum_{i \in W} max \left\{ \sum_{j=1}^{B} |\Delta_j| \sum_{r \in \mathcal{R}} p_{i,j,r} h_{i,j,r}\big(y^{\mathcal{D}}_{i,j,r}\big), C_i \right\} + \mathbb{E}_{K_W}\Big[g\big(K_W\big)\Big]$$

subject to

$$\sum_{i \in W} \sum_{j=1}^{B} |\Delta_j| p_{i,j,r} y^{\mathcal{D}}_{i,j,r} \geq \alpha_r, \qquad \forall r \tag{E.1}$$

$$y^{\mathcal{D}}_{i,j,r} \in [0,1], \qquad \forall i,j,r \tag{E.2}$$

Appendix F: Figures

Figure 1 Operational Details of Real-Time Bidding

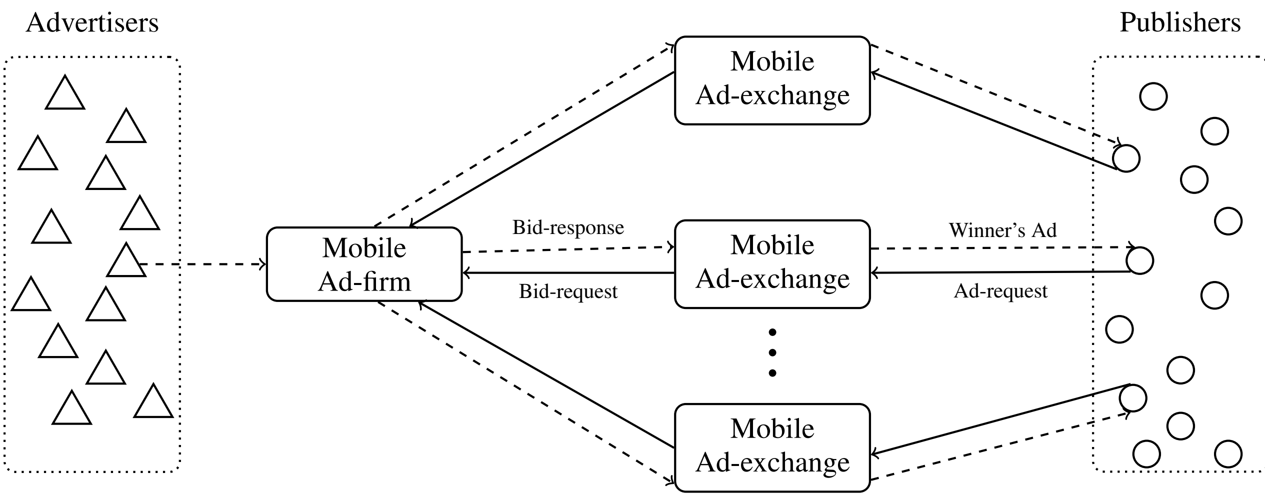
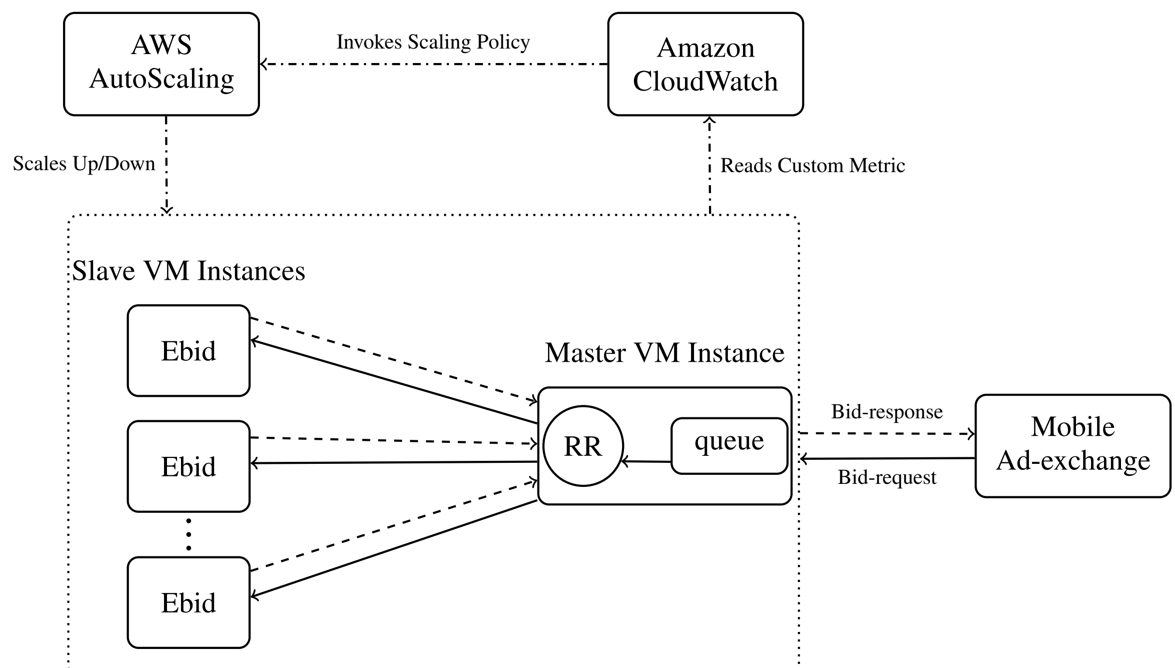
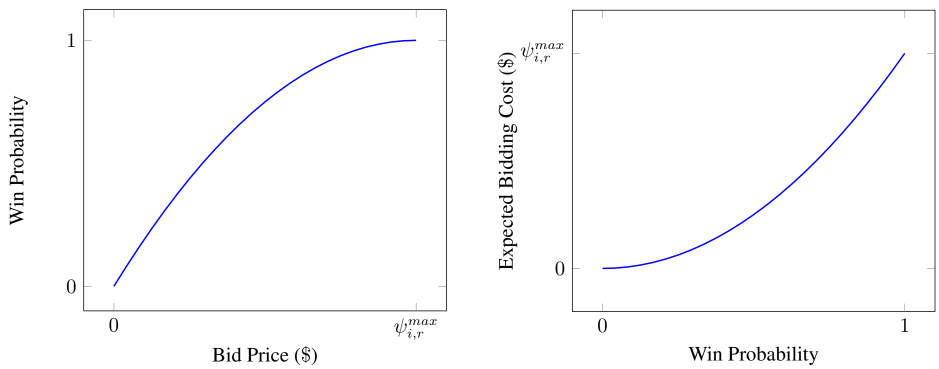


Figure 2 Illustrative Real-time Bidding Architecture



Notes: *Ebid is the bidding algorithm running on each server (VM instance). The autoscaling feature adjusts the number of servers based upon a metric (e.g., queue length) continually observed by a savant algorithm called CloudWatch. RR stands for Round-Robin: each new bid-request is allocated to a server using a Round-Robin scheme.*

Figure 3 Win-Curve and Expected Bidding Cost



(a) Win-Curve (b) Expected Bidding Cost

Figure 4 Overview of Solution

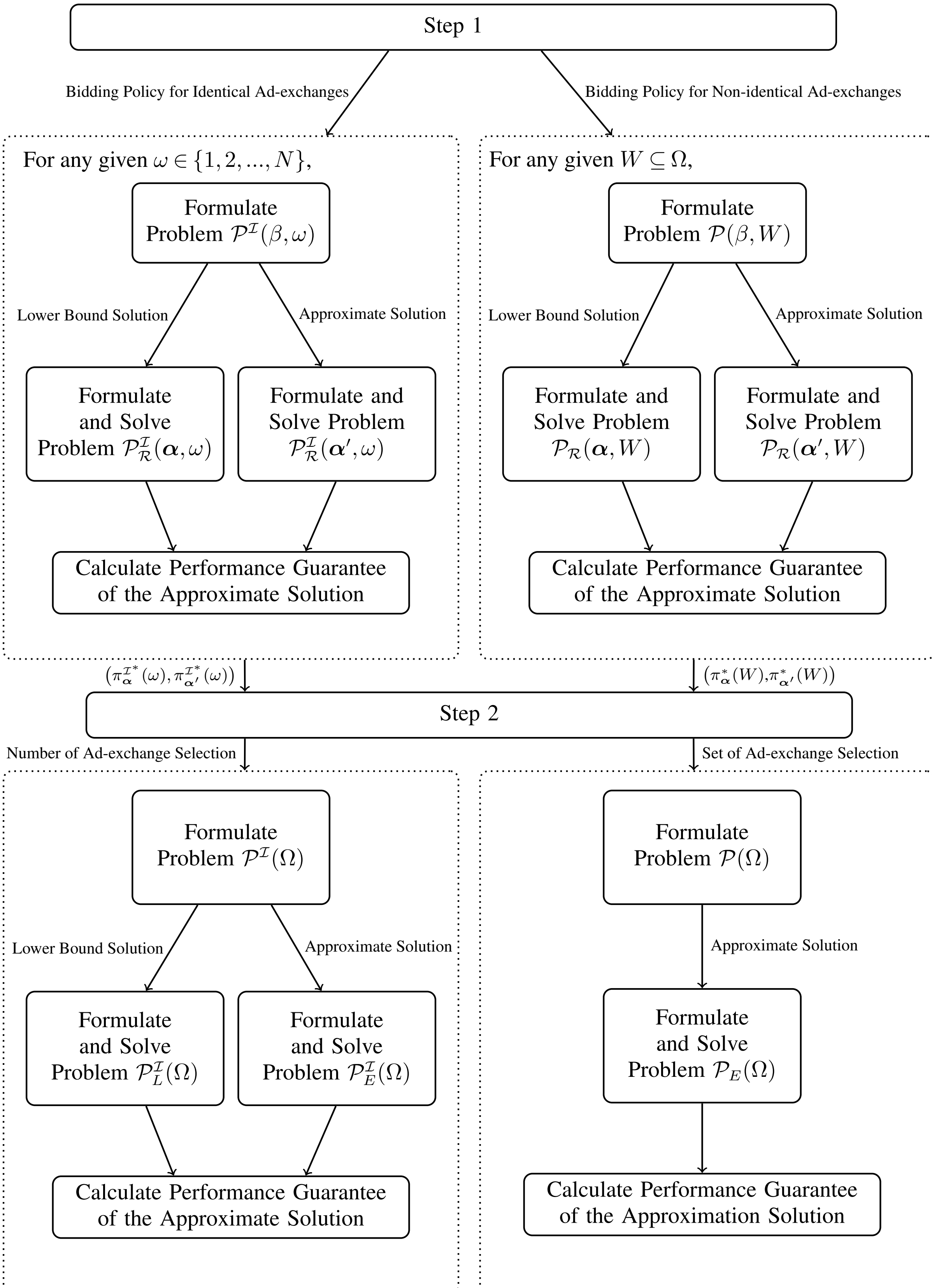
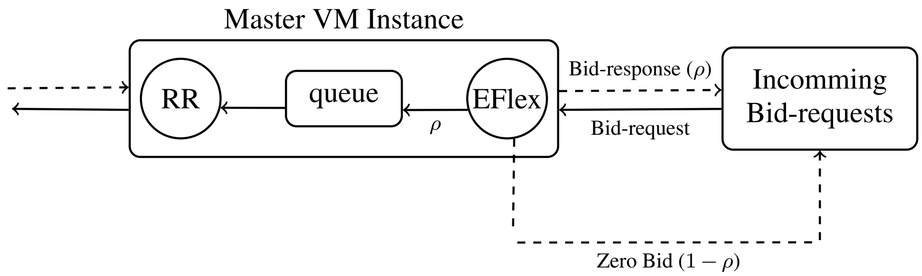


Figure 6 The Proposed Selective Bidding Strategy



Appendix G: Details of Cloud-Based Bidding Architecture

As demonstrated in Figure 2, an ad-firm rents servers (i.e., Virtual Machine (VM) instances) from a cloud provider such as Amazon Web Services (AWS) and creates a cluster of these instances; a cluster consists of one master node (master VM instance) and multiple slave nodes (slave VM instances). When a bid-request arrives, an appropriate bid amount is computed by a bidding algorithm called *Ebid* that is installed on each of the slave VM instances. First, all bid-requests are stored in a queue in master VM instance. Then, each bid-request is allocated to one of the slave VM instances by a load balancer installed on the master VM instance using a Round Robin (RR) algorithm.

The response time of a bid-request is the sum of the service time and the waiting time. The service time – the time needed to calculate the bid amount – does not vary much across bid-requests, and can be assumed to be deterministic. However, the waiting time of a bid-request is random and depends on the bid-request traffic (which is stochastic) and the number and size of the VM instances in the cluster. If the traffic is too high for the current capacity, the number of timeouts can be expected to increase. To ensure that timeouts do not exceed an acceptable limit, the ad-firm can use Auto Scaling feature that automatically scales the capacity (higher or lower) to adjust for the traffic using a predefined scaling policy. By specifying the scaling policy, the Auto Scaling monitors the value of a custom metric and launches or terminates slave VM instances whenever the custom metric's value goes above or below a set target value.

The ad-firm can use backlog per VM instance as a metric with the target value being the acceptable backlog per VM instance to maintain. The backlog per VM instance is calculated using the number of bid-requests waiting in queue at the master VM instance and the number of available slave VM instances. To determine the target value, the ad-firm takes the acceptable time limit to respond to a bid-request (it is typically 100 milliseconds) and divides it by service time. To illustrate with an example, let us assume that the current number of bid-requests waiting in queue is 20 and the current number of slave VM instances in the cluster is 10. If the service time is 20 milliseconds for each bid-request and the acceptable time limit is 100 milliseconds, then the acceptable backlog per VM instance is $\frac{100}{20}$, which equals 5. This means that 5 is target value for the scaling policy. Since the backlog per VM instance is currently about 7 ($\frac{20}{\sqrt{10}}$), we get the required number of VM instances as $\frac{20}{\sqrt{m}} = 5$, or $m = 16$. Hence, 6 more slave instances need to be added to achieve the target value.

It is notable that the Auto Scaling feature is offered at no additional fee; the ad-firm pays only for the VM instances that are rented. At the end of a specific time period, the cloud provider sends a bill that charges the firm based on its usage of VM instances during the time period (and hence the *capacity cost*, or computing cost, for this period) is proportional to total number of bid-requests processed in this period because the usage of VM instances is also proportional to the arriving traffic of bid-requests.

Appendix H: Details of the Logistic Regression for Estimating the Expected Bidding Cost for Case Study

We first ran twelve different regressions on data obtained from Cidewalk, corresponding to each of the exchange-location combination, to estimate the win-curves in the respective combinations. In each regression, the dependent variable is winning status and the bid amount is considered as independent variable. We estimated the coefficients of the regressions and evaluated the goodness-of-fit of the regression models using the McFadden pseudo R squared, McFadden adjusted pseudo R squared, Veall-Zimmermann pseudo R squared, and Mckelvey-Zavonia pseudo R squared. For exchange-location combination (i,r), we estimate the win-probability for a bid amount ψ as $y_{i,r} = \frac{e^{\kappa_0^{i,r}+\kappa_1^{i,r}\psi}}{1+e^{\kappa_0^{i,r}+\kappa_1^{i,r}\psi}}$. Table 5 and 6 present the descriptive statistics for variables, bid amount and win-probability. The results of the regression and the goodness-of-fit test for each exchange-location are reported in Table 7 and 8.

Table 5 Descriptive Statistics for Six Locations and Ad-Exchange $\mathcal{A}$

		Bid Amount (ψ)				Win-Probability (y)			
Location	No. Obs	Min	Max	Mean	St. Dev	Min	Max	Mean	St. Dev
1	12651	0.001	2.587	0.250	0.303	0	1	0.050	0.218
2	4046	0.002	1.850	0.379	0.217	0	1	0.079	0.270
3	14809	0.002	3.392	0.294	0.381	0	1	0.049	0.218
4	8882	0.003	2.594	0.428	0.415	0	1	0.118	0.322
5	10764	0.002	2.60	0.415	0.388	0	1	0.102	0.302
6	16101	0.012	3.40	1.063	0.689	0	1	0.366	0.482

Note. No. Obs, Number of Observations; St. Dev, Standard Deviation.

Table 6 Descriptive Statistics for Six Locations and Ad-Exchange $\mathcal{B}$

		Bid Amount (ψ)				Win-Probability (y)			
Location	No. Obs	Min	Max	Mean	St. Dev	Min	Max	Mean	St. Dev
1	1455	0.013	2.489	0.399	0.331	0	1	0.435	0.496
2	2392	0.011	2.60	0.300	0.264	0	1	0.288	0.453
3	1409	0.035	3.371	0.728	0.582	0	1	0.687	0.463
4	2179	0.049	2.579	0.564	0.379	0	1	0.540	0.498
5	1364	0.088	2.568	0.641	0.426	0	1	0.659	0.474
6	7360	0.120	3.394	0.857	0.635	0	1	0.608	0.488

Note. No. Obs, Number of Observations; St. Dev, Standard Deviation.

Table 7 Estimation Results for Six Locations and Ad-exchange $\mathcal{A}$

	Ad-Exchange $\mathcal{A}$					
Location	Intercept ($\kappa_0^{\mathcal{A},r}$)	Coefficient for Bid ($\kappa_1^{\mathcal{A},r}$)	McFadden Pseudo R Squared	McFadden Adjusted Pseudo R Squared	Veall-Zimmermann Pseudo R Squared	Mckelvey-Zavoina Pseudo R Squared
1	-4.04058^{***} (0.06697)	2.7858^{***} (0.09866)	0.198	0.195	0.354	0.280
2	-3.7393^{***} (0.1502)	2.9393^{***} (0.2826)	0.216	0.202	0.341	0.268
3	-1.99688^{***} (0.02302)	2.15227^{***} (0.03477)	0.226	0.209	0.358	0.326
4	-1.35044^{***} (0.02357)	1.48896^{***} (0.03215)	0.201	0.196	0.337	0.276
5	-1.61715^{***} (0.02406)	1.76308^{***} (0.03386)	0.238	0.212	0.355	0.299
6	-0.75765^{***} (0.02052)	0.64176^{***} (0.01894)	0.215	0.203	0.343	0.279

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$.

Table 8 Estimation Results for Six Locations and Ad-exchange $\mathcal{B}$

	Ad-Exchange $\mathcal{B}$					
Location	Intercept ($\kappa_0^{\mathcal{B},r}$)	Coefficient for Bid ($\kappa_1^{\mathcal{B},r}$)	McFadden Pseudo R Squared	McFadden Adjusted Pseudo R Squared	Veall-Zimmermann Pseudo R Squared	Mckelvey-Zavoina Pseudo R Squared
1	-0.89441^{***} (0.09034)	1.59160^{***} (0.18561)	0.208	0.198	0.342	0.276
2	-2.32853^{***} (0.08811)	4.3459^{***} (0.21266)	0.253	0.214	0.348	0.294
3	0.66132^{***} (0.09254)	$0.17424^{\cdot}$ (0.10286)	0.226	0.198	0.358	0.294
4	$-0.13546^{\cdot}$ (0.07879)	0.53064^{***} (0.11934)	0.227	0.206	0.368	0.305
5	0.2201^{*} (0.1084)	0.7104^{***} (0.1546)	0.239	0.208	0.336	0.284
6	-0.48800^{***} (0.04479)	1.17435^{***} (0.05181)	0.241	0.217	0.359	0.301

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$.