Online Appendices

For convenience, we summarize the notation used in our baseline model.

Symbol	Description
$\theta \in [\theta_{\mathrm{L}}, \theta_{\mathrm{H}}]$	User type (private information)
$F(\theta)$, $f(\theta)$	CDF and PDF of user types
$\psi(\theta)$	Virtual-value function $\left(\theta - \frac{1-F(\theta)}{f(\theta)}\right)$
α	Ad-nuisance parameter
r	Ad-revenue rate
k	Baseline usage without ads
T	Subscription period (days)
μ	Mechanism class; $\mu \in \{\mathrm{ADS}, \mathrm{SUB}, \mathrm{BAS}, \mathrm{GAS}\}$
$n_\mu(\theta)$, $p_\mu(\theta)$	Ad-intensity and subscription-fee functions under μ
$d_\mu(\theta) \in \{0,1\}$	Participation decision under μ
$t^*_\mu(\theta)$	Optimal usage under mechanism μ
$V_\mu(\hat{\theta};\theta)$	Net utility of user with type θ who reveals type $\hat{\theta}$ under mechanism μ
π_μ	Revenue under mechanism μ

Table 3 Our notation for the baseline model.

Appendix A: Proof of Theorem 1 (Optimal GAS Mechanism)

Recall that $V_{\mathrm{GAS}}(\hat{\theta};\theta) = \frac{1}{2}T\theta k\left((1-\alpha n_{\mathrm{GAS}}(\hat{\theta}))_+\right)^2 - p_{\mathrm{GAS}}(\hat{\theta})$. Let $V_{\mathrm{GAS}}(\theta) := V_{\mathrm{GAS}}(\theta;\theta)$ denote the net utility to a user of type θ under an incentive-compatible GAS mechanism. The incentive compatibility constraints imply that:

$$V_{\mathrm{GAS}}(\theta) = \max_{\hat{\theta}\in[\theta_{\mathrm{L}},\theta_{\mathrm{H}}]} \frac{1}{2}T\theta k\left((1-\alpha n_{\mathrm{GAS}}(\hat{\theta}))_+\right)^2 - p_{\mathrm{GAS}}(\hat{\theta}).$$

Note that $V_{\mathrm{GAS}}(\cdot)$ is a maximum of a family of affine functions and is therefore a convex function. Thus, $V_{\mathrm{GAS}}(\cdot)$ is absolutely continuous and is differentiable almost everywhere in the interior of its domain. Using Envelope Theorem, we get

$$\frac{dV_{\mathrm{GAS}}(\theta)}{d\theta} = \frac{Tk}{2}\left((1-\alpha\cdot n_{\mathrm{GAS}}(\theta))_+\right)^2. \tag{32}$$

Integrating, we get

$$V_{\mathrm{GAS}}(\theta) = V_{\mathrm{GAS}}(\theta_{\mathrm{L}}) + \int_{\theta_{\mathrm{L}}}^{\theta} \frac{Tk}{2}\left((1-\alpha\cdot n_{\mathrm{GAS}}(x))_+\right)^2 dx. \tag{33}$$

Thus, the subscription fee function can be now be written as:

$$p_{\mathrm{GAS}}(\theta) = \frac{\theta Tk}{2}\left((1-\alpha n_{\mathrm{GAS}}(\theta))_+\right)^2 - \int_{\theta_{\mathrm{L}}}^{\theta} \frac{Tk}{2}\left((1-\alpha\cdot n_{\mathrm{GAS}}(x))_+\right)^2 dx - V_{\mathrm{GAS}}(\theta_{\mathrm{L}}). \tag{34}$$

The expected revenue to the DCP is:

$$\int_{\theta_L}^{\theta_H}\left(\underbrace{p_{\text{GAS}}(\theta)}_{\text{subscription revenue}} + \underbrace{T\,r\,n_{\text{GAS}}(\theta)\,t^*_{\text{GAS}}(\theta)}_{\text{ad revenue}}\right) d_{\text{GAS}}(\theta) f(\theta) d\theta$$
$$= \int_{\theta_L}^{\theta_H}\left(\frac{\theta T k}{2}\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)^2 - \int_{\theta_L}^{\theta}\frac{Tk}{2}\left((1-\alpha\cdot n_{\text{GAS}}(x))_+\right)^2 dx - V_{\text{GAS}}(\theta_L)\right) d_{\text{GAS}}(\theta) f(\theta) d\theta \,+$$
$$\int_{\theta_L}^{\theta_H} T\,r\,n_{\text{GAS}}(\theta)\,k\left((1-\alpha n_{\text{GAS}}(\theta))_+\right) d_{\text{GAS}}(\theta) f(\theta) d\theta.$$

Note that

$$\int_{\theta_L}^{\theta_H}\int_{\theta_L}^{\theta}\frac{Tk}{2}\left((1-\alpha n_{\text{GAS}}(x))_+\right)^2 dx d_{\text{GAS}}(\theta) f(\theta) d\theta = \int_{\theta_L}^{\theta_H}\frac{Tk}{2}\left((1-\alpha n_{\text{GAS}}(x))_+\right)^2\left(\int_x^{\theta_H} d_{\text{GAS}}(\theta) f(\theta) d\theta\right) dx. \tag{35}$$

Let us define

$$h_{\text{GAS}}(\theta) := d_{\text{GAS}}(\theta) f(\theta),$$
$$\Psi_{\text{GAS}}(\theta) := \begin{cases}\theta - \frac{\int_\theta^{\theta_H} h_{\text{GAS}}(x)dx}{h_{\text{GAS}}(\theta)} & \text{if } h_{\text{GAS}}(\theta) > 0\\ 0 & \text{otherwise.}\end{cases}$$

With this definition, we have $\int_\theta^{\theta_H} h_{\text{GAS}}(\theta) d\theta = (\theta - \Psi_{\text{GAS}}(\theta))\, h_{\text{GAS}}(\theta)$. Consequently,

$$\int_{\theta_L}^{\theta_H}\frac{Tk}{2}\left((1-\alpha n_{\text{GAS}}(x))_+\right)^2\left(\int_x^{\theta_H} d_{\text{GAS}}(\theta) f(\theta) d\theta\right) dx$$
$$= \int_{\theta_L}^{\theta_H}\frac{Tk}{2}\left((1-\alpha n_{\text{GAS}}(x))_+\right)^2\left(\int_x^{\theta_H} h_{\text{GAS}}(\theta) d\theta\right) dx$$
$$= \int_{\theta_L}^{\theta_H}\frac{Tk}{2}\left((1-\alpha n_{\text{GAS}}(x))_+\right)^2 (x - \Psi_{\text{GAS}}(x))\, h_{\text{GAS}}(x) dx$$
$$= \int_{\theta_L}^{\theta_H}\frac{Tk}{2}\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)^2 (\theta - \Psi_{\text{GAS}}(\theta))\, h_{\text{GAS}}(\theta) d\theta.$$

The expected revenue to the DCP now simplifies to:

$$\int_{\theta_L}^{\theta_H}\frac{Tk\theta}{2}\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)^2 h_{\text{GAS}}(\theta) d\theta - \int_{\theta_L}^{\theta_H}\frac{Tk}{2}\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)^2 (\theta - \Psi_{\text{GAS}}(\theta))\, h_{\text{GAS}}(\theta) d\theta \,+$$
$$\int_{\theta_L}^{\theta_H} Trk\, n_{\text{GAS}}(\theta)\left((1-\alpha n_{\text{GAS}}(\theta))_+\right) h_{\text{GAS}}(\theta) d\theta - V_{\text{GAS}}(\theta_L)\int_{\theta_L}^{\theta_H} h_{\text{GAS}}(\theta) d\theta$$
$$= \int_{\theta_L}^{\theta_H}\left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2}\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)^2 + Trk\, n_{\text{GAS}}(\theta)\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)\right) h_{\text{GAS}}(\theta) d\theta$$
$$- V_{\text{GAS}}(\theta_L)\int_{\theta_L}^{\theta_H} h_{\text{GAS}}(\theta) d\theta.$$

Notice that the expected revenue to the DCP is decreasing in $V_{\text{GAS}}(\theta_L)$. Since the individual-rationality constraint for the type θ_L requires $V_{\text{GAS}}(\theta_L) \geq 0$, a revenue-maximizing mechanism rule will have $V_{\text{GAS}}(\theta_L) = 0$. The DCP's revenue-maximization problem $\mathcal{P}$-GAS can now be written as

$$\max_{\substack{n_{\text{GAS}}(\cdot)\geq 0\\ d_{\text{GAS}}(\cdot)\in\{0,1\}}} \int_{\theta_L}^{\theta_H}\left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2}\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)^2 + Trk\, n_{\text{GAS}}(\theta)\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)\right) h_{\text{GAS}}(\theta) d\theta \tag{36}$$

subject to:

$$V_{\text{GAS}}(\theta;\theta) \geq V_{\text{GAS}}(\hat{\theta};\theta)\ \forall\theta\in[\theta_{\text{L}},\theta_{\text{H}}] \qquad \text{(incentive compatibility)}$$

$$d_{\text{GAS}}(\theta) = \begin{cases} 1 & \text{if } V_{\text{GAS}}(\theta;\theta)\geq 0, \\ 0 & \text{otherwise} \end{cases} \quad \forall\theta\in[\theta_{\text{L}},\theta_{\text{H}}], \qquad \text{(individual rationality)}$$

where

$$h_{\text{GAS}}(\theta) = d_{\text{GAS}}(\theta)f(\theta) \quad \text{and} \quad \Psi_{\text{GAS}}(\theta) = \theta - \frac{\int_{\theta}^{\theta_{\text{H}}} h_{\text{GAS}}(\theta)d\theta}{h_{\text{GAS}}(\theta)}.$$

Consider a relaxation of problem $\mathcal{P}$-GAS obtained by neglecting the incentive compatibility and individual rationality constraints:

$$\max_{\substack{n_{\text{GAS}}(\cdot)\geq 0 \\ d_{\text{GAS}}(\cdot)\in\{0,1\}}} \int_{\theta_{\text{L}}}^{\theta_{\text{H}}} \left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2}\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)^2 + Trk\, n_{\text{GAS}}(\theta)\left((1-\alpha n_{\text{GAS}}(\theta))_+\right)\right) h_{\text{GAS}}(\theta)d\theta \quad (\mathcal{P}^{\text{RELAX}})$$

Since the objective function is equal to 0 for $n_{\text{GAS}}(\theta) > \frac{1}{\alpha}$, it is sufficient to restrict the search space of $n_{\text{GAS}}(\theta)$ to $\left[0,\frac{1}{\alpha}\right]$. Thus, we have:

$$\max_{\substack{0\leq n_{\text{GAS}}(\cdot)\leq\frac{1}{\alpha} \\ d_{\text{GAS}}(\cdot)\in\{0,1\}}} \int_{\theta_{\text{L}}}^{\theta_{\text{H}}} \left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2}\left(1-\alpha n_{\text{GAS}}(\theta)\right)^2 + Trk\, n_{\text{GAS}}(\theta)\left(1-\alpha n_{\text{GAS}}(\theta)\right)\right) h_{\text{GAS}}(\theta)d\theta$$

Let us fix $d_{\text{GAS}}(\theta)$ for all $\theta\in[\theta_{\text{L}},\theta_{\text{H}}]$. With $d_{\text{GAS}}(\cdot)$ fixed, we can optimize the relaxed objective point-wise for each value of θ by solving

$$\max_{n_{\text{GAS}}(\theta)\geq 0} \left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2}\left(1-\alpha n_{\text{GAS}}(\theta)\right)^2 + Trk\, n_{\text{GAS}}(\theta)\left(1-\alpha n_{\text{GAS}}(\theta)\right)\right) h_{\text{GAS}}(\theta)$$

Let

$$g(n_{\text{GAS}}(\theta)) := \left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2}\left(1-\alpha n_{\text{GAS}}(\theta)\right)^2 + Trk\, n_{\text{GAS}}(\theta)\left(1-\alpha n_{\text{GAS}}(\theta)\right)\right) h_{\text{GAS}}(\theta)$$

denote the objective function and let $n^*_{\text{GAS}}(\theta)$ denote an optimal solution to the above point-wise optimization problem.

Next, $\forall n_{\text{GAS}}(\theta)\in\left[0,\frac{1}{\alpha}\right]$, we have

$$g'(n_{\text{GAS}}(\theta)) = \big(-\alpha Tk\Psi_{\text{GAS}}(\theta)\left(1-\alpha n_{\text{GAS}}(\theta)\right) + Trk\left(1-2\alpha n_{\text{GAS}}(\theta)\right)\big)h_{\text{GAS}}(\theta),$$
$$g''(n_{\text{GAS}}(\theta)) = \left(\alpha^2 Tk\Psi_{\text{GAS}}(\theta) - 2rTk\alpha\right)f(\theta) = \big(\alpha Tk\left(\alpha\Psi_{\text{GAS}}(\theta)-2r\right)\big)h_{\text{GAS}}(\theta).$$

Consider two cases based on the value of $(\alpha\Psi_{\text{GAS}}(\theta)-2r)$:

Case 1: $\alpha\Psi_{\text{GAS}}(\theta) < 2r$.

$$\alpha\Psi_{\text{GAS}}(\theta) < 2r \implies g''(n_{\text{GAS}}(\theta)) < 0 \quad \forall n_{\text{GAS}}(\theta).$$

Therefore, $g(\cdot)$ is a concave function and its critical value is obtained solving $g'(n_{\text{GAS}}(\theta)) = 0$ in the domain $[0, 1/\alpha]$. This yields

$$n^*_{\text{GAS}}(\theta) = \frac{(r - \alpha\Psi_{\text{GAS}}(\theta))_+}{\alpha\,(2r - \alpha\Psi_{\text{GAS}}(\theta))},$$

and gives us:

$$g(n^*_{\text{GAS}}(\theta)) = \begin{cases} \frac{T\,k\,r^2\,h_{\text{GAS}}(\theta)}{2\alpha(2r-\alpha\Psi_{\text{GAS}}(\theta))} & \text{if } \Psi_{\text{GAS}}(\theta) \le \frac{r}{\alpha}, \\ \frac{T\,k\,\Psi_{\text{GAS}}(\theta)h_{\text{GAS}}(\theta)}{2} & \text{if } \Psi_{\text{GAS}}(\theta) > \frac{r}{\alpha}. \end{cases}$$

Case 2: $\alpha\Psi_{\text{GAS}}(\theta) \ge 2r$.

In this case,

$$\alpha\Psi_{\text{GAS}}(\theta) \ge 2r \implies g''(n_{\text{GAS}}(\theta)) \ge 0.$$

Therefore, $g(\cdot)$ is a convex function and the optimal solution lies on the boundary of the feasible region. Observe that

$$\begin{aligned} g(0) &= \frac{Tk\Psi_{\text{GAS}}(\theta)h_{\text{GAS}}(\theta)}{2} \\ g(1/\alpha) &= 0 \\ \text{Therefore, } g(0) - g(1/\alpha) &= \frac{Tk\Psi_{\text{GAS}}(\theta)h_{\text{GAS}}(\theta)}{2} \\ &\ge \frac{T\,k\,h_{\text{GAS}}(\theta)\cdot\frac{2r}{\alpha}}{2} \quad (\text{since } \alpha\Psi_{\text{GAS}}(\theta) \ge 2r) \\ &\ge 0 \end{aligned}$$

Thus, in this case, the optimal solution to the point-wise optimization problem is:

$$n^*_{\text{GAS}}(\theta) = 0,$$

and consequently,

$$g(n^*_{\text{GAS}}(\theta)) = \frac{Tk\Psi_{\text{GAS}}(\theta)h_{\text{GAS}}(\theta)}{2}.$$

Combining the solution from the above two cases and simplifying, we get:

$$n^*_{\text{GAS}}(\theta) = \begin{cases} \frac{r-\alpha\Psi_{\text{GAS}}(\theta)}{\alpha(2r-\alpha\Psi_{\text{GAS}}(\theta))} & \text{if } \Psi_{\text{GAS}}(\theta) < \frac{r}{\alpha} \\ 0 & \text{if } \Psi_{\text{GAS}}(\theta) \ge \frac{r}{\alpha}, \end{cases} \tag{37}$$

$$g(n^*_{\text{GAS}}(\theta)) = \begin{cases} \frac{T\,k\,r^2\,h_{\text{GAS}}(\theta)}{2\alpha(2r-\alpha\Psi_{\text{GAS}}(\theta))} & \text{if } \Psi_{\text{GAS}}(\theta) < \frac{r}{\alpha}, \\ \frac{T\,k\,\Psi_{\text{GAS}}(\theta)h_{\text{GAS}}(\theta)}{2} & \text{if } \Psi_{\text{GAS}}(\theta) \ge \frac{r}{\alpha} \end{cases} \tag{38}$$

Next, let us substitute this optimal solution and consider the optimization problem with respect to $d_{\text{GAS}}(\cdot)$:

$$\max_{d_{\text{GAS}}(\cdot)\in\{0,1\}} \int_{\theta_{\text{L}}}^{\theta_{\text{H}}} g(n^*_{\text{GAS}}(\theta)) \cdot h_{\text{GAS}}(\theta) d\theta$$

We will now show that $d^*_{\text{GAS}}(\theta) = 1\ \forall\theta \in [\theta_{\text{L}}, \theta_{\text{H}}]$ is an optimal solution to the above problem. We prove this via contradiction.

Suppose that the optimal solution $d^*_{\mathrm{GAS}}(\theta) \neq 1\ \forall \theta \in [\theta_{\mathrm{L}}, \theta_{\mathrm{H}}]$ (except for finitely many points in the domain). Then, there exists a set $\mathcal{S}$ of users, $\mathcal{S} := \{\theta \mid d^*_{\mathrm{GAS}}(\theta) = 0\}$, such that $|\mathcal{S}| > 0$. Since $\mathcal{S}$ has a positive measure, there exists $\theta_0 \in (\theta_{\mathrm{L}}, \theta_{\mathrm{H}})$ and an infinitesimally small $\epsilon > 0$, such that $[\theta_0, \theta_0 + \epsilon] \subset \mathcal{S}$. Consider a solution obtained by flipping the participation decision of the users $[\theta_0, \theta_0 + \epsilon]$:

$$\hat{d}^*_{\mathrm{GAS}}(\theta) = \begin{cases} 1 & \text{if } \theta \in [\theta_0, \theta_0 + \epsilon] \\ d^*_{\mathrm{GAS}}(\theta) & \text{otherwise.} \end{cases}$$

To obtain the revenue under $\hat{d}^*_{\mathrm{GAS}}(\theta)$, it is convenient to define:

$$\begin{aligned}
\hat{h}_{\mathrm{GAS}}(\theta) &:= \hat{d}_{\mathrm{GAS}}(\theta) f(\theta), \\
\hat{\Psi}_{\mathrm{GAS}}(\theta) &:= \begin{cases} \theta - \frac{\int_\theta^{\theta_{\mathrm{H}}} \hat{h}_{\mathrm{GAS}}(x)dx}{\hat{h}_{\mathrm{GAS}}(\theta)} & \text{if } h_{\mathrm{GAS}}(\theta) > 0 \\ 0 & \text{otherwise,} \end{cases} \\
\hat{g}(n^*_{\mathrm{GAS}}(\theta)) &= \begin{cases} \frac{T\, k\, r^2\, \hat{h}_{\mathrm{GAS}}(\theta)}{2\alpha(2r - \alpha \hat{\Psi}_{\mathrm{GAS}}(\theta))} & \text{if } \hat{\Psi}_{\mathrm{GAS}}(\theta) < \frac{r}{\alpha}, \\ \frac{T\, k\, \hat{\Psi}_{\mathrm{GAS}}(\theta) \hat{h}_{\mathrm{GAS}}(\theta)}{2} & \text{if } \hat{\Psi}_{\mathrm{GAS}}(\theta) \geq \frac{r}{\alpha}. \end{cases}
\end{aligned}$$

To stress that the function $g(\cdot)$ depends on $d^*_{\mathrm{GAS}}(\theta)$ via $\Psi_{\mathrm{GAS}}(\theta)$, we define $G(\Psi_{\mathrm{GAS}}(\theta)) := g(n^*_{\mathrm{GAS}}(\theta))$. Thus, $\hat{g}(n^*_{\mathrm{GAS}}(\theta)) = G(\hat{\Psi}_{\mathrm{GAS}}(\theta))$.

Let $\pi, \hat{\pi}$ denote the revenue to the DCP under the solution $d^*_{\mathrm{GAS}}(\theta)$ and $\hat{d}^*_{\mathrm{GAS}}(\theta)$. Then,

$$\pi = \int_{\theta_{\mathrm{L}}}^{\theta_0} G(\Psi_{\mathrm{GAS}}(\theta)) \cdot h_{\mathrm{GAS}}(\theta) d\theta + \underbrace{\int_{\theta_0}^{\theta_0+\epsilon} G(\Psi_{\mathrm{GAS}}(\theta)) \cdot h_{\mathrm{GAS}}(\theta) d\theta}_{=0} + \int_{\theta_0+\epsilon}^{\theta_{\mathrm{H}}} G(\Psi_{\mathrm{GAS}}(\theta)) \cdot h_{\mathrm{GAS}}(\theta) d\theta,$$

$$\hat{\pi} = \int_{\theta_{\mathrm{L}}}^{\theta_0} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) d\theta + \int_{\theta_0}^{\theta_0+\epsilon} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) d\theta + \int_{\theta_0+\epsilon}^{\theta_{\mathrm{H}}} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) d\theta.$$

Note that over the domain $[\theta_0 + \epsilon, \theta_{\mathrm{H}}]$,

$$\begin{aligned}
\hat{d}^*_{\mathrm{GAS}}(\theta) &= d^*_{\mathrm{GAS}}(\theta), \\
\hat{h}_{\mathrm{GAS}}(\theta) &= h_{\mathrm{GAS}}(\theta), \\
\hat{\Psi}_{\mathrm{GAS}}(\theta) &= \Psi_{\mathrm{GAS}}(\theta), \\
G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) &= G(\Psi_{\mathrm{GAS}}(\theta)).
\end{aligned}$$

Therefore,

$$\int_{\theta_0+\epsilon}^{\theta_{\mathrm{H}}} G(\Psi_{\mathrm{GAS}}(\theta)) \cdot h_{\mathrm{GAS}}(\theta) d\theta = \int_{\theta_0+\epsilon}^{\theta_{\mathrm{H}}} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) d\theta$$

Thus,

$$\begin{aligned}
\hat{\pi} - \pi &= \int_{\theta_{\mathrm{L}}}^{\theta_0} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) d\theta - \int_{\theta_{\mathrm{L}}}^{\theta_0} G(\Psi_{\mathrm{GAS}}(\theta)) \cdot h_{\mathrm{GAS}}(\theta) d\theta + \int_{\theta_0}^{\theta_0+\epsilon} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) d\theta \\
&= \int_{\theta_{\mathrm{L}}}^{\theta_0} \big(G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) - G(\Psi_{\mathrm{GAS}}(\theta)) \cdot h_{\mathrm{GAS}}(\theta) \big) d\theta + \int_{\theta_0}^{\theta_0+\epsilon} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) d\theta
\end{aligned}$$

Over the interval $[\theta_0, \theta_0 + \epsilon]$, $\hat{h}_{\mathrm{GAS}}(\theta) = f(\theta) > 0$. Further, $G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) = G(\Psi_{\mathrm{GAS}}(\theta)) > 0$. Therefore,

$$\begin{aligned}
\int_{\theta_0}^{\theta_0+\epsilon} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot \hat{h}_{\mathrm{GAS}}(\theta) d\theta &= \int_{\theta_0}^{\theta_0+\epsilon} G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) \cdot f(\theta) d\theta \\
&\approx\ G(\hat{\Psi}_{\mathrm{GAS}}(\theta_0)) \cdot f(\theta_0) \cdot \epsilon + O(\epsilon^2) > 0.
\end{aligned}$$

For $\theta \in [\theta_{\mathrm{L}}, \theta_0]$, we have:

$$\begin{aligned}
\int_{\theta}^{\theta_{\mathrm{H}}} \hat{h}_{\mathrm{GAS}}(x)dx &= \int_{\theta}^{\theta_0} \hat{h}_{\mathrm{GAS}}(x)dx + \int_{\theta_0}^{\theta_0+\epsilon} \hat{h}_{\mathrm{GAS}}(x)dx + \int_{\theta_0+\epsilon}^{\theta_{\mathrm{H}}} \hat{h}_{\mathrm{GAS}}(x)dx \\
&= \int_{\theta}^{\theta_0} h_{\mathrm{GAS}}(x)dx + \int_{\theta_0}^{\theta_0+\epsilon} \hat{h}_{\mathrm{GAS}}(x)dx + \int_{\theta_0+\epsilon}^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x)dx \\
&= \int_{\theta}^{\theta_0} h_{\mathrm{GAS}}(x)dx + \underbrace{\int_{\theta_0}^{\theta_0+\epsilon} h_{\mathrm{GAS}}(x)dx}_{=0} + \int_{\theta_0+\epsilon}^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x)dx + \int_{\theta_0}^{\theta_0+\epsilon} \hat{h}_{\mathrm{GAS}}(x)dx \\
&= \int_{\theta}^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x)dx + \int_{\theta_0}^{\theta_0+\epsilon} \hat{h}_{\mathrm{GAS}}(x)dx \\
&= \int_{\theta}^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x)dx + \int_{\theta_0}^{\theta_0+\epsilon} f(x)dx.
\end{aligned}$$

Therefore, over the interval $[\theta_{\mathrm{L}}, \theta_0]$, we get:

$$\begin{aligned}
\hat{\Psi}_{\mathrm{GAS}}(\theta) &= \begin{cases} \theta - \frac{\int_{\theta}^{\theta_{\mathrm{H}}} \hat{h}_{\mathrm{GAS}}(x)dx}{\hat{h}_{\mathrm{GAS}}(\theta)} & \text{if } \hat{h}_{\mathrm{GAS}}(\theta) > 0 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \theta - \frac{\int_{\theta}^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x)dx + \int_{\theta_0}^{\theta_0+\epsilon} f(x)dx}{h_{\mathrm{GAS}}(\theta)} & \text{if } h_{\mathrm{GAS}}(\theta) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (\text{since } \hat{h}_{\mathrm{GAS}}(\theta) = h_{\mathrm{GAS}}(\theta) \text{ for } \theta \in [\theta_{\mathrm{L}}, \theta_0]) \\
&\leq \Psi_{\mathrm{GAS}}(\theta).
\end{aligned} \tag{39}$$

In addition, we have the following over the interval $[\theta_{\mathrm{L}}, \theta_0]$:

$$\hat{\Psi}_{\mathrm{GAS}}(\theta) - \Psi_{\mathrm{GAS}}(\theta) = -\frac{\int_{\theta_0}^{\theta_0+\epsilon} f(x)dx}{h_{\mathrm{GAS}}(\theta)} \approx \frac{-f(\theta_0)\cdot\epsilon}{h_{\mathrm{GAS}}(\theta)} + O(\epsilon^2)$$

Notice that function $G(\cdot)$ is strictly increasing; that is, $G'(\Psi_{\mathrm{GAS}}(\theta)) > 0$. Thus, over the interval $[\theta_{\mathrm{L}}, \theta_0]$, we get

$$G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) - G(\Psi_{\mathrm{GAS}}(\theta)) < 0.$$

Since $\hat{h}_{\mathrm{GAS}}(\theta) = h_{\mathrm{GAS}}(\theta)$ for $\theta \in [\theta_{\mathrm{L}}, \theta_0]$), we get

$$\begin{aligned}
\int_{\theta_{\mathrm{L}}}^{\theta_0} \big(G(\hat{\Psi}_{\mathrm{GAS}}(\theta))\cdot\hat{h}_{\mathrm{GAS}}(\theta) - G(\Psi_{\mathrm{GAS}}(\theta))\cdot h_{\mathrm{GAS}}(\theta)\big)d\theta &= \int_{\theta_{\mathrm{L}}}^{\theta_0} \big(G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) - G(\Psi_{\mathrm{GAS}}(\theta))\big)\cdot h_{\mathrm{GAS}}(\theta)d\theta \\
&< 0.
\end{aligned}$$

Let us now obtain a lower bound on $\int_{\theta_{\mathrm{L}}}^{\theta_0} \big(G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) - G(\Psi_{\mathrm{GAS}}(\theta))\big)\cdot h_{\mathrm{GAS}}(\theta)d\theta$. Notice that $G(\cdot)$ is continuous and differentiable in its domain. Thus, using Mean-Value Theorem, there exists $\xi(\theta) \in [\hat{\Psi}_{\mathrm{GAS}}(\theta), \Psi_{\mathrm{GAS}}(\theta)]$, such that

$$\begin{aligned}
G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) - G(\Psi_{\mathrm{GAS}}(\theta)) &= G'(\xi(\theta))\cdot\Big(\hat{\Psi}_{\mathrm{GAS}}(\theta) - \Psi_{\mathrm{GAS}}(\theta)\Big) \\
&\approx G'(\xi(\theta))\cdot\left(\frac{-f(\theta_0)\cdot\epsilon}{h_{\mathrm{GAS}}(\theta)}\right) + O(\epsilon^2) \\
\implies \big(G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) - G(\Psi_{\mathrm{GAS}}(\theta))\big)\cdot h_{\mathrm{GAS}}(\theta) &\approx -G'(\xi(\theta))\cdot f(\theta_0)\cdot\epsilon + O(\epsilon^2).
\end{aligned}$$

Therefore,

$$\int_{\theta_{\mathrm{L}}}^{\theta_0} \big(G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) - G(\Psi_{\mathrm{GAS}}(\theta))\big) \cdot h_{\mathrm{GAS}}(\theta) d\theta$$
$$\approx - f(\theta_0) \cdot \epsilon \cdot \int_{\theta_{\mathrm{L}}}^{\theta_0} G'(\xi(\theta)) d\theta + O(\epsilon^2).$$

Note that

$$G(\Psi_{\mathrm{GAS}}(\theta)) = \begin{cases} \frac{T\,k\,r^2\,h_{\mathrm{GAS}}(\theta)}{2\alpha(2r-\alpha\Psi_{\mathrm{GAS}}(\theta))} & \text{if } \Psi_{\mathrm{GAS}}(\theta) < \frac{r}{\alpha}, \\ \frac{T\,k\,\Psi_{\mathrm{GAS}}(\theta) h_{\mathrm{GAS}}(\theta)}{2} & \text{if } \Psi_{\mathrm{GAS}}(\theta) \geq \frac{r}{\alpha}. \end{cases}$$

$$G'(\Psi_{\mathrm{GAS}}(\theta)) = \frac{dG(\Psi_{\mathrm{GAS}}(\theta))}{d\Psi_{\mathrm{GAS}}(\theta)} = \begin{cases} \dfrac{T\,k\,r^2\,h_{\mathrm{GAS}}(\theta)}{2\,(2r-\alpha\,\Psi_{\mathrm{GAS}}(\theta))^2}, & \Psi_{\mathrm{GAS}}(\theta) < \dfrac{r}{\alpha}, \\ \dfrac{T\,k\,h_{\mathrm{GAS}}(\theta)}{2}, & \Psi_{\mathrm{GAS}}(\theta) \geq \dfrac{r}{\alpha}, \end{cases}$$

$$G''(\Psi_{\mathrm{GAS}}(\theta)) = \frac{d^2G(\Psi_{\mathrm{GAS}}(\theta))}{d\Psi_{\mathrm{GAS}}(\theta)^2} = \begin{cases} \dfrac{T\,k\,r^2\,\alpha\,h_{\mathrm{GAS}}(\theta)}{(2r-\alpha\,\Psi_{\mathrm{GAS}}(\theta))^3}, & \Psi_{\mathrm{GAS}}(\theta) < \dfrac{r}{\alpha}, \\ 0, & \Psi_{\mathrm{GAS}}(\theta) \geq \dfrac{r}{\alpha}. \end{cases}$$

Since $G'(\cdot) \geq 0$, $G''(\cdot) \geq 0$, $G(\cdot)$ is a convex non-decreasing function of its argument. Further, since $\xi(\theta) \leq \Psi_{\mathrm{GAS}}(\theta))$ and $G'(\cdot) > 0$, we get:

$$\int_{\theta_{\mathrm{L}}}^{\theta_0} \big(G(\hat{\Psi}_{\mathrm{GAS}}(\theta)) - G(\Psi_{\mathrm{GAS}}(\theta))\big) \cdot h_{\mathrm{GAS}}(\theta) d\theta \geq -f(\theta_0) \cdot \epsilon \cdot \int_{\theta_{\mathrm{L}}}^{\theta_0} G'(\Psi_{\mathrm{GAS}}(\theta)) d\theta + O(\epsilon^2).$$

Next, note that when $d_{\mathrm{GAS}}(\theta) = 1$, we have $h_{\mathrm{GAS}}(\theta) = f(\theta)$ and

$$\Psi_{\mathrm{GAS}}(\theta) = \theta - \frac{\int_\theta^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{f(\theta)},$$

$$\begin{aligned} \Psi'_{\mathrm{GAS}}(\theta) = \frac{d\Psi_{\mathrm{GAS}}(\theta)}{d\theta} &= 1 - \frac{f(\theta) \cdot (-h_{\mathrm{GAS}}(\theta)) - f'(\theta) \cdot \int_\theta^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{f(\theta)^2} \\ &= 1 - \frac{f(\theta) \cdot (-h_{\mathrm{GAS}}(\theta)) - f'(\theta) \cdot \int_\theta^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{f(\theta)^2} \\ &= 1 - \frac{f(\theta) \cdot (-f(\theta)) - f'(\theta) \cdot \int_\theta^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{f(\theta)^2} \\ &= 2 + \frac{f'(\theta) \cdot \int_\theta^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{f(\theta)^2} \end{aligned}$$

Further, the monotone hazard rate assumption implies

$$\frac{d\left(\frac{f(\theta)}{1-F(\theta)}\right)}{d\theta} \geq 0 \implies (1-F(\theta)) f'(\theta) + f(\theta)^2 \geq 0 \implies \frac{(1-F(\theta)) f'(\theta)}{f(\theta)^2} \geq -1.$$

Therefore,

$$\Psi'_{\mathrm{GAS}}(\theta) = 2 + \frac{f'(\theta) \cdot \int_\theta^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{f(\theta)^2}.$$

If $\frac{f'(\theta)}{f(\theta)^2} \geq 0$, then $\frac{f'(\theta) \cdot \int_\theta^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{f(\theta)^2} \geq 0 \geq -1$. If $\frac{f'(\theta)}{f(\theta)^2} < 0$, then $\frac{f'(\theta) \cdot \int_\theta^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{f(\theta)^2} \geq \frac{f'(\theta) \cdot \int_\theta^{\theta_{\mathrm{H}}} f(x) dx}{f(\theta)^2} = \frac{(1-F(\theta)) f'(\theta)}{f(\theta)^2} \geq -1$. In either case, we get:

$$\Psi'_{\mathrm{GAS}}(\theta) \geq 2 + (-1) = 1.$$

Using the above bound on $\Psi'_{\text{GAS}}(\theta)$ and the substitution $t = \Psi_{\text{GAS}}(\theta)$, we get $dt = \Psi'_{\text{GAS}}(\theta)d\theta$, which implies $d\theta = \frac{dt}{\Psi'_{\text{GAS}}(\theta)} \leq 1$. This implies:

$$\begin{aligned}-f(\theta_0)\cdot\epsilon\cdot\int_{\theta_{\text{L}}}^{\theta_0} G'(\Psi_{\text{GAS}}(\theta))d\theta + O(\epsilon^2) &\geq -f(\theta_0)\cdot\epsilon\cdot\int_{\Psi(\theta_{\text{L}})}^{\Psi(\theta_0)} G'(t)dt + O(\epsilon^2)\\ &= -f(\theta_0)\cdot\epsilon\cdot(G(\Psi(\theta_0)) - G(\Psi(\theta_{\text{L}}))) + O(\epsilon^2).\end{aligned}$$

We can now write the revenue difference as:

$$\begin{aligned}\hat{\pi} - \pi &= \int_{\theta_{\text{L}}}^{\theta_0}\big(G(\hat{\Psi}_{\text{GAS}}(\theta))\cdot\hat{h}_{\text{GAS}}(\theta) - G(\Psi_{\text{GAS}}(\theta))\cdot h_{\text{GAS}}(\theta)\big)d\theta + \int_{\theta_0}^{\theta_0+\epsilon} G(\hat{\Psi}_{\text{GAS}}(\theta))\cdot\hat{h}_{\text{GAS}}(\theta)d\theta\\ &\geq -f(\theta_0)\cdot\epsilon\cdot(G(\Psi(\theta_0)) - G(\Psi(\theta_{\text{L}}))) + G(\hat{\Psi}_{\text{GAS}}(\theta_0))\cdot f(\theta_0)\cdot\epsilon + O(\epsilon^2)\\ &= f(\theta_0)\cdot\epsilon\cdot G(\Psi(\theta_{\text{L}})) + O(\epsilon^2)\end{aligned}$$

Since $G(\Psi(\theta_{\text{L}})) > 0$, we get that $\hat{\pi} - \pi > 0$. This proves that any mechanism which leaves a positive-measure set of types out of the subscription cannot be revenue-maximizing. In other words, the optimal participation rule must satisfy $d^*_{\text{GAS}}(\theta) = 1$ for all $\theta \in [\theta_{\text{L}}, \theta_{\text{H}}]$.

Since $d^*_{\text{GAS}}(\theta) = 1$ for all $\theta \in [\theta_{\text{L}}, \theta_{\text{H}}]$, we get $\Psi_{\text{GAS}}(\theta) = \psi(\theta)$. Thus, the optimal ad-intensity can be simplified to:

$$n^*_{\text{GAS}}(\theta) = \begin{cases}\frac{(r-\alpha\psi(\theta))_+}{\alpha(2r-\alpha\psi(\theta))} & \text{if } \psi(\theta) < \frac{r}{\alpha}\\ 0 & \text{if } \psi(\theta) \geq \frac{r}{\alpha}.\end{cases}$$

Next, we show that $n^*_{\text{GAS}}(\theta)$ is a weakly non-increasing function of θ. For $\theta \geq \psi^{-1}(r/\alpha)$, $n^*_{\text{GAS}}(\theta) = 0$ is constant. For $\theta < \psi^{-1}(r/\alpha)$, we have $n^*_{\text{GAS}}(\theta) = \frac{r-\alpha\psi(\theta)}{\alpha(2r-\alpha\psi(\theta))}$. Differentiating with respect to θ, we get:

$$\begin{aligned}\frac{dn^*_{\text{GAS}}(\theta)}{d\theta} &= \frac{(-\alpha\psi'(\theta))\cdot\alpha(2r-\alpha\psi(\theta)) - (r-\alpha\psi(\theta))\cdot(-\alpha^2\psi'(\theta))}{\alpha^2\left(2r-\alpha\psi(\theta)\right)^2}\\ &= \frac{-r\,\psi'(\theta)}{\left(2r-\alpha\psi(\theta)\right)^2}.\end{aligned} \tag{40}$$

Since $r > 0$ and $\psi'(\theta) \geq 0$ under the MHR assumption, (40) implies $\frac{dn^*_{\text{GAS}}(\theta)}{d\theta} \leq 0$. Moreover, $n^*_{\text{GAS}}(\psi^{-1}(r/\alpha)) = 0$, confirming continuity at the threshold. Therefore, $n^*_{\text{GAS}}(\theta)$ is weakly non-increasing in θ on $[\theta_{\text{L}}, \theta_{\text{H}}]$.

For $n^*_{\text{GAS}}(\theta)$ to be an optimal solution to problem $\mathcal{P}$-GAS, we need to show that it satisfies incentive compatibility and individual rationality constraints. We now present sufficient conditions under which a mechanism μ is incentive compatible and individually rational for the users.

Lemma 1. *A GAS mechanism is incentive compatible and individually rational if:*

- $n_{GAS}(\theta) \leq n_{GAS}(\hat{\theta})\ \forall\theta \geq \hat{\theta}$, *where* $\theta, \hat{\theta} \in [\theta_L, \theta_H]$ *and* $0 \leq n_{GAS}(\cdot) \leq 1/\alpha$,
- $p_{GAS}(\theta) = \frac{\theta Tk(1-\alpha n_{GAS}(\theta))^2}{2} - \int\limits_{\theta_L}^{\theta}\frac{Tk(1-\alpha n_{GAS}(x))^2}{2}dx\ \forall\theta \in [\theta_L, \theta_H]$.

Proof: To show that a GAS mechanism is incentive-compatible, we need to show that

$$V_{\text{GAS}}(\theta;\theta) - V_{\text{GAS}}(\hat{\theta};\theta) \geq 0\ \forall, \theta, \hat{\theta} \in [\theta_{\text{L}}, \theta_{\text{H}}].$$

Consider any $\theta, \hat{\theta} \in [\theta_{\mathrm{L}}, \theta_{\mathrm{H}}]$. Then,

$$\begin{aligned}
&V_{\mathrm{GAS}}(\theta;\theta) - V_{\mathrm{GAS}}(\hat{\theta};\theta) \\
=&\frac{\theta Tk}{2}\left((1-\alpha n_{\mathrm{GAS}}(\theta))_+\right)^2 - p_{\mathrm{GAS}}(\theta) - \frac{\theta Tk}{2}\left((1-\alpha n_{\mathrm{GAS}}(\hat{\theta}))_+\right)^2 + p_{\mathrm{GAS}}(\hat{\theta}).
\end{aligned}$$

Using $0 \le n_{\mathrm{GAS}}(\cdot) \le 1/\alpha$ and $p_{\mathrm{GAS}}(\theta) = \frac{\theta Tk(1-\alpha n_{\mathrm{GAS}}(\theta))^2}{2} - \int_{\theta_{\mathrm{L}}}^{\theta} \frac{Tk(1-\alpha n_{\mathrm{GAS}}(x))^2}{2}dx$, we get

$$\begin{aligned}
&V_{\mathrm{GAS}}(\theta;\theta) - V_{\mathrm{GAS}}(\hat{\theta};\theta) \\
&= \int_{\theta_{\mathrm{L}}}^{\theta} \frac{Tk\left(1-\alpha n_{\mathrm{GAS}}(x)\right)^2}{2}dx - \frac{\theta Tk}{2}\left(1-\alpha n_{\mathrm{GAS}}(\hat{\theta})\right)^2 + \frac{\hat{\theta} Tk}{2}\left(1-\alpha n_{\mathrm{GAS}}(\hat{\theta})\right)^2 - \int_{\theta_{\mathrm{L}}}^{\hat{\theta}} \frac{Tk\left(1-\alpha n_{\mathrm{GAS}}(x)\right)^2}{2}dx \\
&= \int_{\hat{\theta}}^{\theta} \frac{Tk\left(1-\alpha n_{\mathrm{GAS}}(x)\right)^2}{2}dx - \frac{\theta Tk}{2}\left(1-\alpha n_{\mathrm{GAS}}(\hat{\theta})\right)^2 + \frac{\hat{\theta} Tk}{2}\left(1-\alpha n_{\mathrm{GAS}}(\hat{\theta})\right)^2 \\
&= \int_{\hat{\theta}}^{\theta} \frac{Tk\left(1-\alpha n_{\mathrm{GAS}}(x)\right)^2}{2}dx - \int_{\hat{\theta}}^{\theta} \frac{Tk}{2}\left(1-\alpha n_{\mathrm{GAS}}(\hat{\theta})\right)^2 dx \\
&= \int_{\hat{\theta}}^{\theta} \left(\frac{Tk\left(1-\alpha n_{\mathrm{GAS}}(x)\right)^2}{2} - \frac{Tk\left(1-\alpha n_{\mathrm{GAS}}(\hat{\theta})\right)^2}{2}\right)dx \\
&\ge 0 \ \left(\text{since } n_{\mathrm{GAS}}(\theta) \le n_{\mathrm{GAS}}(\hat{\theta}) \ \forall \theta \ge \hat{\theta}\right).
\end{aligned}$$

Further, notice that for all user types $\theta \in [\theta_{\mathrm{L}}, \theta_{\mathrm{H}}]$, we have

$$\begin{aligned}
V_{\mathrm{GAS}}(\theta;\theta) &= \frac{\theta Tk}{2}\left(1-\alpha n_{\mathrm{GAS}}(\theta)\right)^2 - p_{\mathrm{GAS}}(\theta) \\
&= \int_{\theta_{\mathrm{L}}}^{\theta} \frac{Tk\left(1-\alpha n_{\mathrm{GAS}}(x)\right)^2}{2}dx \\
&\ge 0.
\end{aligned}$$

Thus, all the users obtain non-negative utility upon participation and hence the mechanism is individually rational. This concludes the proof of Lemma 1. ∎

Consider the following GAS* mechanism:

$$n^*_{\mathrm{GAS}}(\theta) = \begin{cases} \frac{r-\alpha\psi(\theta)}{\alpha(2r-\alpha\psi(\theta))} & \text{if } \theta < \psi^{-1}\left(\frac{r}{\alpha}\right) \\ 0 & \text{if } \theta \ge \psi^{-1}\left(\frac{r}{\alpha}\right), \end{cases} \tag{41}$$

$$p^*_{\mathrm{GAS}}(\theta) = \frac{\theta Tk\left(1-\alpha n^*_{\mathrm{GAS}}(\theta)\right)^2}{2} - \int_{\theta_{\mathrm{L}}}^{\theta} \frac{Tk\left(1-\alpha n^*_{\mathrm{GAS}}(x)\right)^2}{2}dx, \tag{42}$$

$$d^*_{\mathrm{GAS}}(\theta) = 1. \tag{43}$$

Note that $(n^*_{\mathrm{GAS}}(\theta), d^*_{\mathrm{GAS}}(\theta))$ is an optimal solution to problem $\mathcal{P}^{\text{RELAX}}$. Using Lemma 1, we conclude that the GAS* mechanism satisfies the incentive compatibility and individual rationality constraints. Therefore, the GAS* mechanism is an optimal solution to problem $\mathcal{P}$-GAS. ∎

Finally, we show that $p^*_{\mathrm{GAS}}(\theta)$ is weakly non-decreasing in θ. Differentiating $p^*_{\mathrm{GAS}}(\theta)$ with respect to θ, we get:

$$\begin{aligned}\frac{dp^*_{\mathrm{GAS}}(\theta)}{d\theta} &= \frac{Tk\left(1-\alpha n^*_{\mathrm{GAS}}(\theta)\right)^2}{2} + \theta Tk\left(1-\alpha n^*_{\mathrm{GAS}}(\theta)\right)(-\alpha)\frac{dn^*_{\mathrm{GAS}}(\theta)}{d\theta} - \frac{Tk\left(1-\alpha n^*_{\mathrm{GAS}}(\theta)\right)^2}{2} \\ &= -\theta\,\alpha\,Tk\left(1-\alpha n^*_{\mathrm{GAS}}(\theta)\right)\frac{dn^*_{\mathrm{GAS}}(\theta)}{d\theta}. \end{aligned} \tag{44}$$

Since $\theta \geq 0$, $\alpha, T, k > 0$, $(1-\alpha n^*_{\mathrm{GAS}}(\theta)) \geq 0$, and $\frac{dn^*_{\mathrm{GAS}}(\theta)}{d\theta} \leq 0$ from (40), equation (44) implies $\frac{dp^*_{\mathrm{GAS}}(\theta)}{d\theta} \geq 0$. Therefore, $p^*_{\mathrm{GAS}}(\theta)$ is weakly non-decreasing in θ on $[\theta_{\mathrm{L}}, \theta_{\mathrm{H}}]$. ■

Appendix B: Proof of Proposition 1 (GAS* as a Posted-Price Mechanism)

Recall from Theorem 1 that the GAS* mechanism is given by:

$$\begin{aligned} n^*_{\mathrm{GAS}}(\theta) &= \begin{cases} \frac{r-\alpha\psi(\theta)}{\alpha(2r-\alpha\psi(\theta))} & \text{if } \theta < \psi^{-1}\left(\frac{r}{\alpha}\right) \\ 0 & \text{if } \theta \geq \psi^{-1}\left(\frac{r}{\alpha}\right), \end{cases} \\ p^*_{\mathrm{GAS}}(\theta) &= \frac{\theta Tk\left(1-\alpha n^*_{\mathrm{GAS}}(\theta)\right)^2}{2} - \int_{\theta_{\mathrm{L}}}^{\theta} \frac{Tk\left(1-\alpha n^*_{\mathrm{GAS}}(x)\right)^2}{2}dx, \\ d^*_{\mathrm{GAS}}(\theta) &= 1. \end{aligned}$$

We want to obtain a posted-price mechanism – that is, subscription price as a function of ad-intensity, $\mathcal{P}(n)$ – that implements the GAS* mechanism. Recall that under the GAS* mechanism, under any mechanism μ, if the ad intensity $n_{\mathrm{GAS}}(\hat{\theta}) > \frac{1}{\alpha}$, then $t^*_{\mathrm{GAS}}(\hat{\theta}) = 0$. That is, if the ad-intensity is above the threshold $\frac{1}{\alpha}$ then users do not consume any content and consequently, the platform does not obtain any revenue. Therefore, it suffices to focus on ad intensities in the range $\left[0, \frac{1}{\alpha}\right]$.

Note that, for $\theta < \psi^{-1}\left(\frac{r}{\alpha}\right)$, $n^*_{\mathrm{GAS}}(\theta) = \frac{r-\alpha\psi(\theta)}{\alpha(2r-\alpha\psi(\theta))}$. Since $n^*_{\mathrm{GAS}}(\theta)$ is strictly decreasing in θ and becomes 0 for $\theta \geq \psi^{-1}\left(\frac{r}{\alpha}\right)$, we can invert this relationship on $\left[0, \frac{1}{\alpha}\right]$ and obtain:

$$\theta(n) = \psi^{-1}\left(\frac{r(1-2\alpha n)}{\alpha(1-\alpha n)}\right) \;\; \forall n \in \left[0, \frac{1}{\alpha}\right]. \tag{45}$$

Here, $\theta(n)$ denotes the user type for which ad intensity n is intended in the GAS* mechanism. Next, for each $n \in \left[0, \frac{1}{\alpha}\right]$, define

$$\mathcal{P}(n) := \frac{Tk\theta(n)(1-\alpha n)^2}{2} - \int_{\theta_{\mathrm{L}}}^{\theta(n)} \frac{Tk\left(1-\alpha\, n^*_{\mathrm{GAS}}(x)\right)^2}{2}dx. \tag{46}$$

Observe that this definition precisely matches the payment function of the GAS* mechanism, namely $p^*_{\mathrm{GAS}}(\theta(n))$. The function $\mathcal{P}(n)$ is set in this manner so that the user of type $\theta(n)$ (the user for whom ad-intensity n is intended) gets the exact same subscription price that he would pay under the GAS* mechanism, if she reported her type truthfully.

Now consider the posted-price mechanism $\mathcal{M} := \{(n, \mathcal{P}(n)) : n \in [0,\, 1/\alpha]\}$. Here, $\mathcal{M}$ can be viewed as a menu of pairs of ad intensities and subscription prices. We will now show that the posted-price mechanism, $\mathcal{M}$, replicates the outcomes of the direct GAS* mechanism.

Incentive Compatibility. We need to show that for every type θ, the best choice from the posted-price menu $\mathcal{M}$ is precisely the pair $\big(n^*_{\mathrm{GAS}}(\theta),\, \mathcal{P}\big(n^*_{\mathrm{GAS}}(\theta)\big)\big)$. In other words, if a type-θ user deviates to any other pair $\big(n,\, \mathcal{P}(n)\big)$ with $n \neq n^*_{\mathrm{GAS}}(\theta)$, then she ends up with strictly lower utility.

For any type θ, choosing $n' \neq n^*_{\mathrm{GAS}}(\theta)$ yields:

$$\begin{aligned} V(\theta, n', \mathcal{P}(n')) &= \frac{\theta T k(1-\alpha n')^2}{2} - \mathcal{P}(n') \\ &= \frac{Tk(1-\alpha n')^2}{2}(\theta - \theta(n')) + \int_{\theta_{\mathrm{L}}}^{\theta(n')} \frac{Tk(1-\alpha n^*_{\mathrm{GAS}}(x))^2}{2} dx. \end{aligned} \tag{47}$$

Further, we have

$$V(\theta, n^*_{\mathrm{GAS}}(\theta), \mathcal{P}(n^*_{\mathrm{GAS}}(\theta))) = \int_{\theta_{\mathrm{L}}}^{\theta} \frac{Tk(1-\alpha n^*_{\mathrm{GAS}}(x))^2}{2} dx. \tag{48}$$

Thus, the utility difference between deviation and truth-telling is:

$$\begin{aligned} V(\theta, n', \mathcal{P}(n')) - V(\theta, n^*_{\mathrm{GAS}}(\theta), \mathcal{P}(n^*_{\mathrm{GAS}}(\theta))) = &\frac{Tk(1-\alpha n')^2}{2}(\theta - \theta(n')) \\ &+ \int_{\theta(n')}^{\theta} \frac{Tk(1-\alpha n^*_{\mathrm{GAS}}(x))^2}{2} dx. \end{aligned} \tag{49}$$

To evaluate the utility difference, we consider two cases:

- **Case 1**: $n' < n^*_{\mathrm{GAS}}(\theta)$ (choosing a lower ad intensity than intended).

Since $n^*_{\mathrm{GAS}}(\theta)$ is strictly decreasing in θ, we have $\theta(n') > \theta$. This gives us:

$$\frac{Tk(1-\alpha n')^2}{2}(\theta - \theta(n')) < 0.$$

Furthermore, since $\theta(n') > \theta$, we also get

$$\int_{\theta(n')}^{\theta} \frac{Tk(1-\alpha n^*_{\mathrm{GAS}}(x))^2}{2} dx < 0.$$

Thus, for $n' < n^*_{\mathrm{GAS}}(\theta)$, we have $U(\theta, n', p(n')) - U(\theta, n^*_{\mathrm{GAS}}(\theta), \mathcal{P}(n^*_{\mathrm{GAS}}(\theta)) < 0$.

- **Case 2:** $n' > n^*_{\mathrm{GAS}}(\theta)$ (choosing a higher ad intensity than intended).

Since $n^*_{\mathrm{GAS}}(x) \geq n^*_{\mathrm{GAS}}(\theta)$ for all $x \leq \theta$, we have

$$(1-\alpha n^*_{\mathrm{GAS}}(x))^2 \leq (1-\alpha n^*_{\mathrm{GAS}}(\theta))^2. \tag{50}$$

Thus,

$$\int_{\theta(n')}^{\theta} \frac{Tk(1-\alpha n^*_{\mathrm{GAS}}(x))^2}{2} dx \leq (\theta - \theta(n'))\frac{Tk(1-\alpha n^*_{\mathrm{GAS}}(\theta))^2}{2}. \tag{51}$$

Plugging this bound into the utility difference, we obtain

$$\begin{aligned} V(\theta, n', \mathcal{P}(n')) - V(\theta, n^*_{\mathrm{GAS}}(\theta), \mathcal{P}(n^*_{\mathrm{GAS}}(\theta))) &\leq \frac{Tk(1-\alpha n')^2}{2}(\theta - \theta(n')) \\ &\quad - (\theta - \theta(n'))\frac{Tk(1-\alpha n^*_{\mathrm{GAS}}(\theta))^2}{2} \\ &= \frac{Tk(\theta - \theta(n'))}{2}\Big[(1-\alpha n')^2 - (1-\alpha n^*_{\mathrm{GAS}}(\theta))^2\Big]. \end{aligned} \tag{52}$$

Since $n' > n^*_{\mathrm{GAS}}(\theta)$, it follows that $(1-\alpha n')^2 < (1-\alpha n^*_{\mathrm{GAS}}(\theta))^2$. Further, since the mapping $n \mapsto \theta(n)$ is strictly decreasing, we have $\theta - \theta(n') > 0$. Therefore,

$$V(\theta, n', \mathcal{P}(n')) - V(\theta, n^*_{\mathrm{GAS}}(\theta), \mathcal{P}(n^*_{\mathrm{GAS}}(\theta))) < 0. \tag{53}$$

Thus, from both cases, we infer that a user of type θ will select a pair of $(n, \mathcal{P}(n))$ from menu $\mathcal{M}$ such that $n = n^*_{\text{GAS}}(\theta)$ and the corresponding subscription price $\mathcal{P}(n) = p^*_{\text{GAS}}(\theta(n))$.

Individual Rationality. For a user of type θ, note that:

$$V(\theta, n^*_{\text{GAS}}(\theta), \mathcal{P}(n^*_{\text{GAS}}(\theta))) = \int_{\theta_{\text{L}}}^{\theta} \frac{Tk(1-\alpha n^*_{\text{GAS}}(x))^2}{2} dx \geq 0. \tag{54}$$

In summary, by constructing the inverse mapping $\theta(n)$ and defining the posted-price menu as

$$\{(n, \mathcal{P}(n)) : n \in [0, 1/\alpha]\} \quad \text{with} \quad \mathcal{P}(n) = p^*_{\text{GAS}}(\theta(n)),$$

we ensure that every user of type θ self-selects the option $\left(n^*_{\text{GAS}}(\theta), p^*_{\text{GAS}}(\theta)\right)$. Hence, the outcomes, regarding both ad intensity and subscription fee, are identical to those under the direct mechanism GAS*. Thus, the posted-price mechanism $\mathcal{M}$ implements GAS* mechanism. ■

Appendix C: Sensitivity Analysis of the Mechanisms

In this section, we present a detailed analysis of revenue comparison under the four optimal mechanisms – ADS*, SUB*, BAS*, and GAS* – by analyzing how it varies with key parameters: the ad-nuisance rate (α), the ad-revenue rate (r), the variance in the distribution of user types (θ), and the baseline usage (k). We set the nominal values of the parameters as described in Table 1.

C.1. Sensitivity to Ad-Nuisance and Ad-Revenue Rates

Figure 11 (left panel) illustrates how the expected revenue to the DCP varies across the four mechanisms with respect to the ad-nuisance (α). Since there are no ads under the SUB* mechanism, the expected revenue remains invariant to changes in ad-nuisance. In contrast, the revenue under the ADS* mechanism declines sharply as ad-nuisance increases. This decline stems from reduced engagement of users: higher ad-nuisance leads to lower viewing times, which in turn reduces the total number of ad impressions and, consequently, advertising revenue. On the flip side, the BAS* and GAS* mechanisms – which leverage the ad revenue as well as the subscription revenue – also experience a reduction in revenue as ad-nuisance increases, albeit at a much slower rate compared to the ADS* mechanism. The rationale is that the BAS* and GAS* mechanisms mitigate the losses more effectively by offering high-type users to opt for reduced ad exposure in exchange for a subscription fee.

The right panel of Figure 11 illustrates how the expected revenue varies with the ad-revenue rate (r). As expected, the SUB* mechanism remains unaffected, whereas the ADS* mechanism experiences a linear increase in revenue since it solely depends on ads. For the BAS* and GAS* mechanisms, the increase in the revenue is more nuanced. As the GAS* mechanism offers a more flexible menu of ad intensity and subscription fee compared to the BAS* mechanism, it yields a higher expected revenue to the DCP.

Collectively, Figure 11 shows that all ad-supported mechanisms (ADS*, BAS*, and GAS*) benefit from higher ad revenue rates but are sensitive to ad-nuisance, while BAS* and GAS* mitigate the impact of parameter fluctuations by leveraging subscription fees.

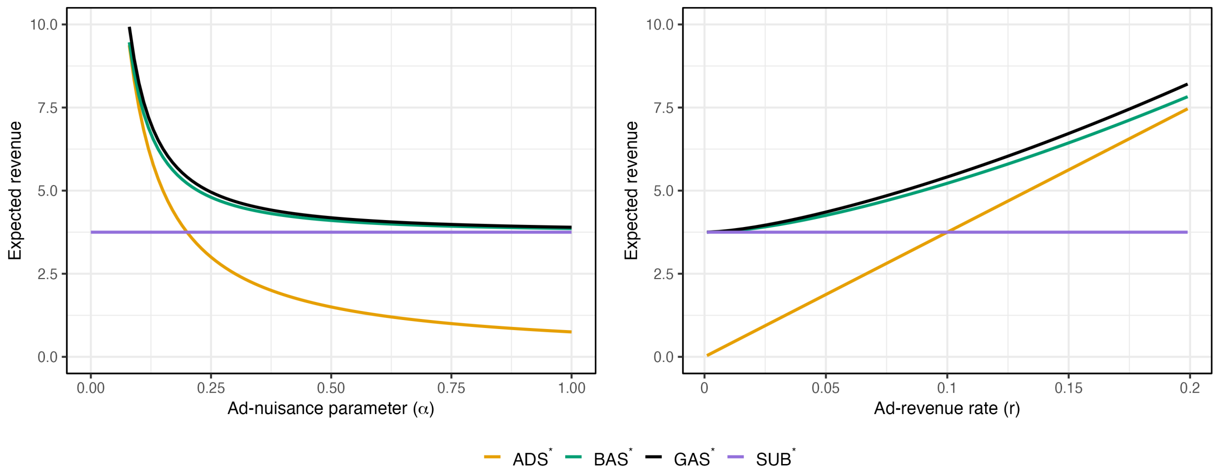


Figure 11 **Expected revenue as a function of ad-nuisance parameter α (left) and ad-revenue rate r (right) across the four mechanisms. Both panels use nominal and range values as described in Table 1.**

C.2. Sensitivity to the Distribution of User Types

The performance of each revenue mechanism also depends on the distribution of user types, which captures the heterogeneity in the value that users derive from the net usage benefit. To assess this dependence, we model user types with a flexible Beta distribution, which can accommodate a wide range of shapes, ranging from uniform to highly concentrated or skewed. Further, we restrict the parameters (a,b) of the distribution to values that satisfy the monotone hazard rate (MHR) condition. We now consider two contrasting shifts in the distribution of user types.

Figure 12 illustrates how the expected revenue under each of the four mechanisms responds to variations in the distribution of user types. In the left panel of the figure, the Beta distribution transitions from a uniform distribution, Beta(m,m) with $m=1$, to a distribution that closely approximates a bell curve by setting $m=6$.[12] Under the ADS* mechanism, every user faces the same ad intensity, irrespective of their type. Consequently, all users watch the same amount of content, generating a fixed number of ad impressions per user. As a result, the total expected revenue remains constant across different distributions of user types. On the other hand, the other three mechanisms – SUB*, BAS*, and GAS* – do respond to how the distribution of user types is spread out. In particular, as m increases, the distribution shifts towards moderate-to-high values, which implies that more users are willing to subscribe and pay for the service. Further, with an increase in m, the distribution also gets concentrated (around 0.5), which reduces uncertainty about how many users will ultimately subscribe to the service. Collectively, the combined effect of more potential subscribers and less uncertainty drives up the expected revenue under the three mechanisms.

[12] A Beta distribution Beta(m,m) with $m=1$ is uniform on the interval [0,1]. As the parameter m increases, the distribution concentrates around 0.5.

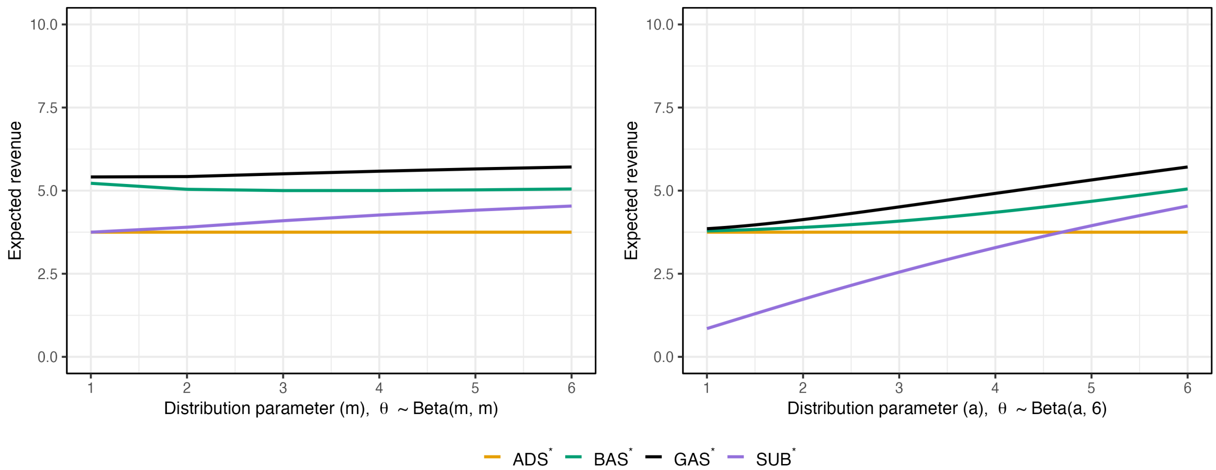


Figure 12 Expected revenue per user depending on distribution shape Beta(a, b) by mechanism. Both panels are produced using nominal and range values from Table 1.

In the right panel of Figure 12, the distribution of user types shifts from right-skewed, namely $\text{Beta}(1,6)$, to near-normal $\text{Beta}(6,6)$ distribution. Under the highly skewed distribution, most users have low WTP and do not find subscription fees worthwhile, making SUB* the least effective mechanism. In contrast, mechanisms that incorporate advertising – ADS*, BAS*, and GAS* – perform better by monetizing a larger user base. As the distribution shifts to $\text{Beta}(6,6)$, more users shift into the mid-to-high-WTP range, which increases the viability of subscription fees. This transition improves the performance of the SUB*, BAS*, and GAS* mechanisms.

C.3. Sensitivity to Baseline Usage

In this section, we study how the performance of each mechanism varies with baseline usage (i.e., in the absence of advertising). Recall from (3) that usage equals k when ad-intensity is zero, so k captures baseline usage. Thus, a higher value of k corresponds to users spending more time consuming content in the absence of ads, which in turn reflects stronger demand for the DCP (e.g., better content and engagement, strong habits).

Figure 13 plots the expected revenue under each mechanism as a function of k for two levels of ad-nuisance: $\alpha = 0.25$ (left panel) and $\alpha = 0.15$ (right panel). Across both panels, revenues increase with k under all four mechanisms, since higher baseline usage raises total consumption and enables the DCP to extract higher surplus through subscriptions and/or ads. The GAS* and BAS* mechanisms yield the highest revenues throughout, and GAS* mechanism maintains a slight advantage over the BAS* mechanism.

The relative ranking of the SUB* and ADS* mechanisms depends on the ad-nuisance environment. When ad-nuisance is high (left panel, $\alpha = 0.25$), SUB* dominates ADS* because the high disutility from ads reduces the usage and ad impressions. In contrast, when ad-nuisance is low (right panel, $\alpha = 0.15$), ADS* dominates SUB* as the DCP can more effectively leverage ad revenue.

In summary, while the level of k amplifies revenue uniformly across mechanisms, it does not alter the qualitative ranking.

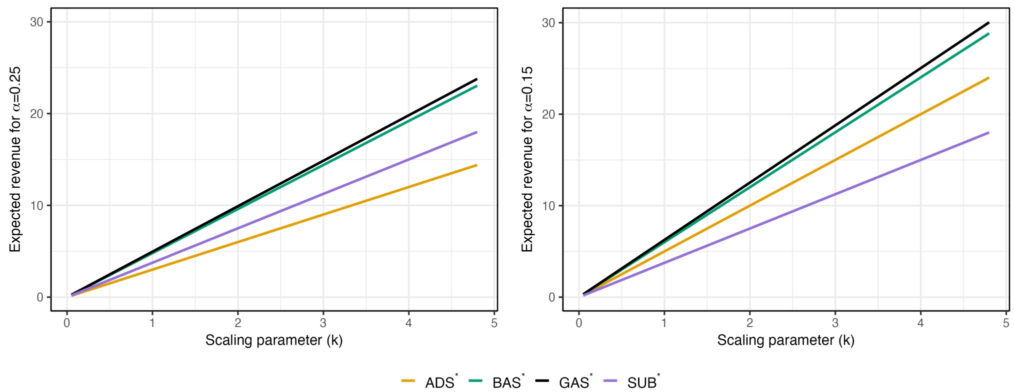


Figure 13 **Expected revenue as a function of the baseline usage (k) across the four mechanisms. Both panels are produced using nominal and range values from Table 1.**

C.4. Sensitivity of Optimal Access Options under GAS* Mechanism

In the preceding sections, we examined how expected revenue varies with key model parameters. A key observation is that the revenue under the GAS* mechanism is relatively stable across parameters values. A natural follow-up question is whether this stability reflects the stability in the underlying access options (that is, the subscription fee $p^*_{\text{GAS}}(\theta)$ and ad-intensity $n^*_{\text{GAS}}(\theta)$ offered to each user type θ). To better understand this question, we plot the optimal access options under the GAS* mechanism as a function of the distribution of the user types and ad-nuisance.

Variation in the distribution of user types: Figure 14 plots the optimal subscription-fee $p^*_{\text{GAS}}(\theta)$ (left panel) and ad-intensity $n^*_{\text{GAS}}(\theta)$ (right panel) under the GAS* mechanism for several symmetric Beta distributions $\theta \sim \text{Beta}(m,m)$, with $m \in \{1,2,3,4,5,6\}$. As m increases, the distribution of user types becomes more concentrated around the center of the support.

First, we note that the subscription-fee curve shifts across distribution specifications: under the uniform distribution $(m=1)$, the fee rises gradually over the entire range of types, whereas under more concentrated distributions $(m \geq 3)$ the fee increases more steeply for intermediate values of θ and levels off at a lower maximum. Second, the ad-intensity curve also shifts in a pronounced manner. Under $m=1$, advertising is sustained at high levels over a wide range of types and declines gradually. In contrast, as m increases, the maximum ad-intensity decreases, the ad-intensity curve becomes steeper, and the cutoff type – the value of θ at which advertising is reduced to zero – shifts to a lower value θ. In other words, more concentrated distributions lead the platform to reduce advertising earlier and rely more heavily on subscription revenue from moderate- and high-type users. We note that these adjustments in the subscription-fee and ad-intensity

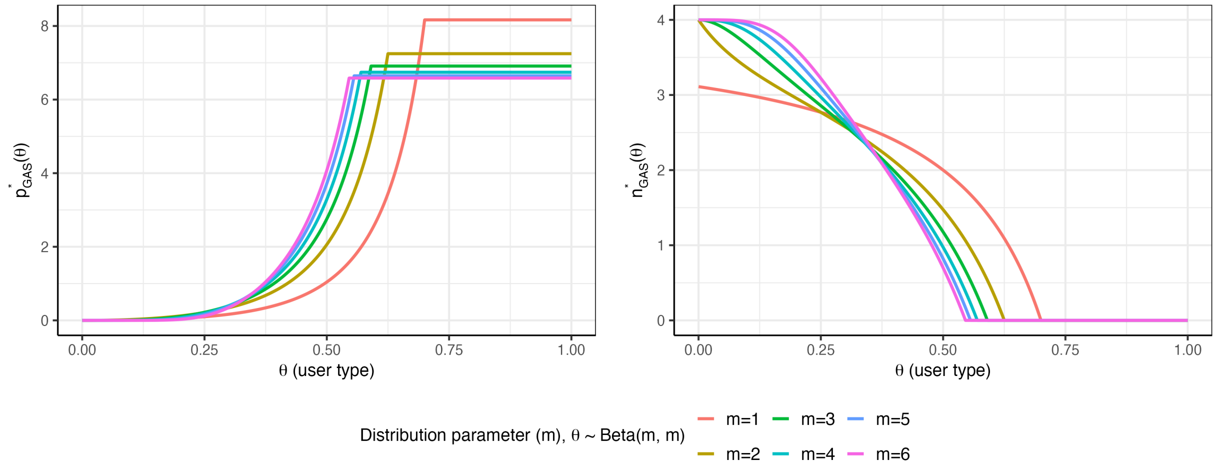


Figure 14 **Subscription-fee and ad-intensity under GAS* mechanism as a function of concentration of user-type distribution. Both panels use nominal values as described in Table 1.**

curves are substantial even though the total expected revenue remains relatively stable (as shown in Figure 12) across the same range of distribution parameters.

Variation in the ad-nuisance parameter: Figure 15 plots the optimal subscription-fee $p^*_{\text{GAS}}(\theta)$ (left panel) and ad-intensity $n^*_{\text{GAS}}(\theta)$ (right panel) under the GAS* mechanism for several values of the ad-nuisance parameter $\alpha \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$.

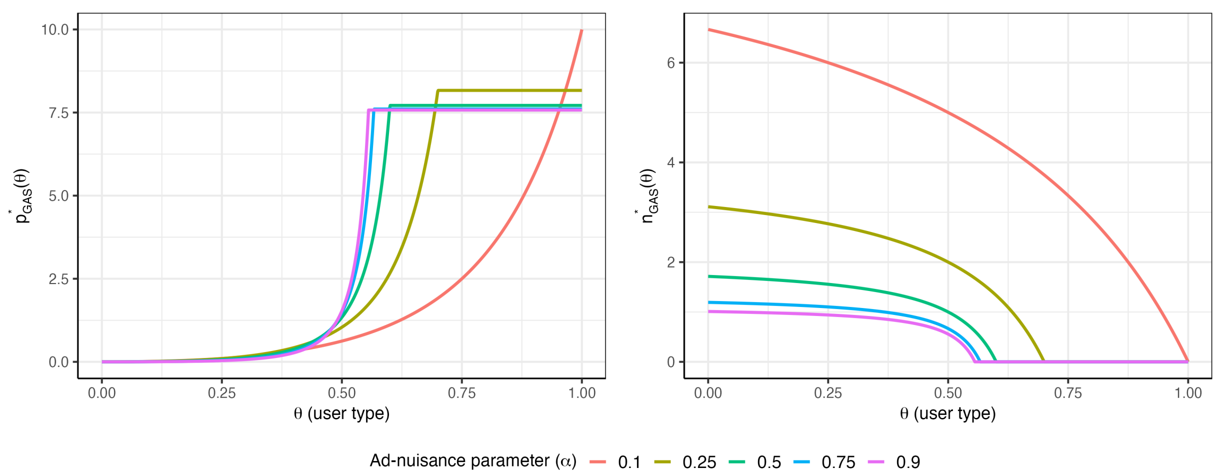


Figure 15 **Subscription-fee and ad-intensity under GAS* mechanism as a function of ad-nuisance parameter α. Both panels use nominal values as described in Table 1.**

When α is low, ads impose relatively little disutility on users, and the optimal solution features high ad-intensity over a wide range of types: the DCP monetizes most users primarily through ads, with only the highest-type users paying a subscription fee to avoid ads. As α increases, the disutility from ads grows, and the optimal ad-intensity sharply declines: both the maximum level of advertising and the range of types exposed to ads shrink. At high values of α, the optimal ad-intensity is low across all types, and the no-ad region expands toward lower types. On the other hand, the subscription-fee curve is relatively stable across values of α for high-type users. This reflects that these users consistently prefer the ad-free option. However, for low- and moderate-type users, the subscription-fee curve adjusts according to the changes in the ad-intensities across users. Thus, the access options shift substantially with variation in the ad-nuisance.

In sum, Figures 14 and 15 demonstrate that the stability of the revenue under the GAS* mechanism is not driven by stability in the optimal access options. Rather, the DCP adjusts the composition of the optimal menu across the users by shifting the balance between the two revenue levers: advertising and subscription fees. This adjustment is a distinctive feature of the GAS* mechanism. The continuous menu gives it the flexibility to continuously trade off ad-intensity against subscription fees across user types.

Appendix D: Welfare Analysis

In this section, we examine how total welfare – defined as the sum of the DCP's expected revenue and expected consumer surplus – varies with the key parameters of the model. We report results for all four mechanisms in Figures 16 and 17.

Figure 16 (left panel) shows total welfare as a function of the ad-nuisance parameter α. As expected, the SUB* mechanism is invariant to α since it does not involve advertising. The ADS* mechanism experiences the sharpest decline: higher ad-nuisance reduces both viewing time and, consequently, the surplus that users derive from the platform. In contrast, the welfare under GAS* and BAS* mechanisms decrease at a slower rate, as users who find ads particularly costly can self-select into lower ad-intensity tiers. The welfare gap between GAS* and BAS* is small throughout. This shows that the continuous menu of GAS* extracts only marginally more surplus than the two-tier menu of BAS*.

The right panel of Figure 16 plots total welfare against the ad-revenue rate r. Higher ad-revenue rates raise total welfare under all ad-supported mechanisms because the revenue of the DCP increases whereas the disutility incurred by users due to ad-nuisance does not (since α is fixed). The welfare under ADS* mechanism increases steeply and is higher than that under SUB* after moderate values of r. The GAS* and BAS* mechanisms dominate the other two mechanisms. The welfare under GAS* is slightly higher than that under BAS*.

Figure 17 examines sensitivity to the distribution of user types. In the left panel, where $\theta \sim \mathrm{Beta}(m, m)$, increasing m concentrates the distribution around 0.5. Total welfare under GAS* and BAS* remains high and nearly flat, indicating that these two mechanisms are robust to changes in the distribution of user types. The SUB* mechanism improves noticeably as m increases, similar to that in revenue analysis: a more concentrated distribution places more users near the subscription threshold which raises both participation and the welfare. The ADS* mechanism, by contrast, is nearly invariant to the shape of the distribution, since all users face the same ad intensity regardless of type.

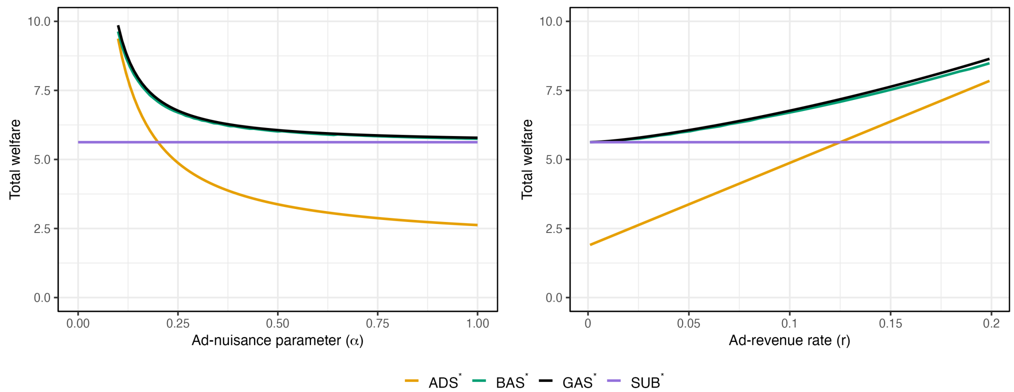


Figure 16 **Expected total welfare as a function of ad-nuisance parameter α (left) and ad-revenue rate r (right) across the four mechanisms. Both panels use nominal values as described in Table 1.**

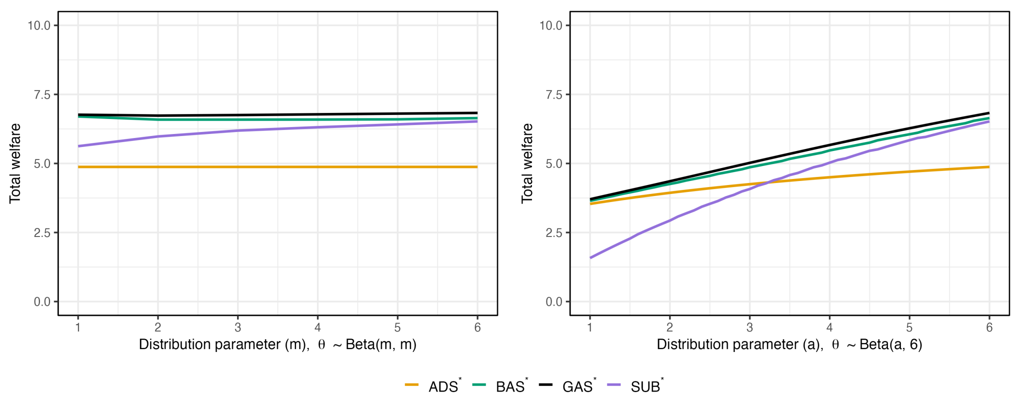


Figure 17 **Expected total welfare depending on distribution shape Beta(a, b) by mechanism. Both panels are produced using nominal and range values from Table 1.**

The right panel of Figure 17 shifts the distribution from a right-skewed Beta$(1,6)$ to a near-symmetric Beta$(6,6)$. Total welfare increases under all mechanisms as the mass of high-type users grows. The welfare ordering is consistent with the revenue ranking: GAS* dominates, followed closely by BAS*, then SUB*, and finally ADS*. Thus, the flexibility of the GAS* mechanism benefits not only the DCP but also the users, as a richer menu enables better matching between user preferences and ad-intensity.

Appendix E: Numerical Study of the Revenue Gap Between GAS* and BAS*

Recall that an analytical characterization of the revenue gap between GAS* and BAS* is not feasible because the BAS* solution does not admit a closed-form expression. Consequently, we study the gap between GAS* and BAS* numerically. In our study, we vary all the parameters of our model according to the nominal values reported in Table 1: ad-nuisance (α), ad-revenue rate (r), shape parameters of the distribution of user types (a, b) in Beta(a, b), and average content usage without ads (k). For each combination of parameters

(α, r, a, b, k), we compute the optimal revenue to the DCP under GAS* and BAS* and define the revenue gap between the two mechanisms by:

$$\Phi := \frac{\pi_{\mathrm{GAS}^*} - \pi_{\mathrm{BAS}^*}}{\pi_{\mathrm{BAS}^*}} \times 100.$$

Thus, Φ measures the percentage increase in expected revenue that the DCP obtains by employing the continuous GAS* mechanism over the two-tier BAS* mechanism. Since GAS* yields weakly higher revenue than BAS* by construction, our objective is to identify parameter instances in which this advantage is substantial. We proceed in two steps:

- Step 1: We use a classification tree to identify parameters that are linked with a large percentage gain of GAS* over BAS*. Accordingly, we classify an instance as exhibiting a "large" revenue improvement whenever the revenue gap satisfies $\Phi \geq 10\%$.

- Step 2: Using the parameters identified in the first step, we study the ranges over which GAS* outperforms BAS* by a large margin.

Step 1: Classification tree analysis: As we vary multiple parameters simultaneously, a comprehensive one-at-a-time sensitivity analysis is difficult to present in a meaningful way. To obtain a parsimonious and interpretable characterization of how the model parameters influence the gap Φ, we adopt a classification-tree approach that helps us identify important parameter thresholds and their interactions.

We construct the classification tree as follows. We treat each point on the parameter grid as an observation, with features (covariates) given by the model primitives (α, r, a, b, k). For each observation, we compute the revenue gap Φ and convert it into a binary outcome: $Y = \mathbb{1}_{\{\Phi \geq 10\%\}}$, so that $Y = 1$ indicates a "large" revenue improvement and $Y = 0$ otherwise. The classification tree then seeks simple threshold rules in the parameter space that best separate observations with $Y = 1$ from those with $Y = 0$. Each internal node of the tree corresponds to a binary split and the tree is grown by recursively selecting the split that maximizes classification accuracy at each node. The resulting tree summarizes the main drivers and their interactions as an ordered set of "if–then" statements. Thus, the tree provides a transparent description of which parameters and parameter ranges most strongly influence the gap between GAS* and BAS*.

Figure 18 displays the fitted classification tree. The first split is on ad-nuisance α: when $\alpha \geq 0.35$, the tree classifies nearly all instances as having a small revenue gap ($Y = 0$), regardless of the other parameters. This implies that, when users find ads highly annoying, the DCP cannot load users with ads under either mechanism, so the two mechanisms perform similarly. Conditional on low ad-nuisance ($\alpha < 0.35$), the subsequent splits involve the type-distribution shape parameter a and the ad-revenue rate r. Large revenue gaps arise when a is sufficiently large (indicating a user population with enough high-type users) and r is sufficiently high (indicating that each ad impression generates substantial revenue for the DCP). This is precisely the region where the continuous menu of GAS* screens users more effectively than the coarse two-tier menu in BAS*.

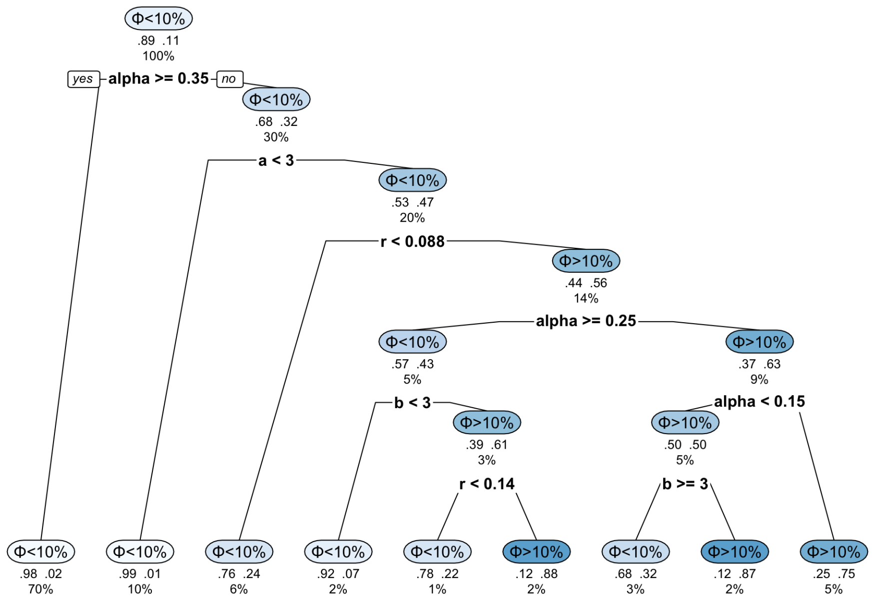


Figure 18 **Classification tree identifying regions of the parameter space in which GAS* yields a large revenue gap over BAS* (i.e., $\Phi \geq 10\%$). Each internal node shows a binary split on a model parameter. The leaf nodes report the predicted class ($Y = 0$ or $Y = 1$).**

Based on these split patterns, we identify two key drivers of the revenue gap: (i) ad attractiveness, captured jointly by the ad-revenue rate r and the ad-nuisance α, and (ii) the proportion of high-type users, captured by the distribution of user types $\text{Beta}(a, b)$. We now examine the effect of each driver in detail.

Step 2: Analysis of the revenue gap: Having identified ad attractiveness and the proportion of high-type users as the primary drivers, we now examine how the revenue gap Φ varies with these drivers. We organize the analysis into three parts: (i) the effect of ad attractiveness via the ad-revenue rate r and ad-nuisance α, (ii) the effect of the user-type distribution via the $\text{Beta}(a, b)$ shape parameters, and (iii) the joint effect of ad attractiveness and the user-type distribution through the composite ratios r/α and a/b.

Throughout this analysis, we employ the following averaging procedure. For each analysis, we fix the parameters of interest and compute the average revenue gap Φ over all remaining parameters. For instance, when studying the effect of ad attractiveness, we consider each pair (r, α) from Table 1 and, for each pair, average the revenue gap Φ across all combinations of the remaining parameters (a, b, k). This averaging smooths out the effects of the parameters not under consideration and isolates the marginal effect of the parameters of interest.

(i) Effect of ad attractiveness (r and α): Figure 19 presents a heatmap of the average revenue gap Φ (in %) as a function of the ad-revenue rate r and the ad-nuisance parameter α. We observe that the revenue gap Φ is largest in the upper-left region of the heatmap, where ads are simultaneously high-paying (large r) and less annoying to users (small α). Moving rightward along any row (i.e., increasing α while holding

r fixed) consistently decreases the revenue gap. This is because higher ad-nuisance reduces the effectiveness of advertising as a screening instrument under both mechanisms. This narrows the advantage of GAS* over BAS*.

Moving upward along any column (i.e., increasing r while holding α fixed) generally raises the revenue gap. The intuition is that a higher ad-revenue rate amplifies the value of tailored ad-intensity. This value is better captured by the continuous menu of GAS* mechanism compared to the two-tiered BAS* mechanism. An exception to this monotonicity arises when α is very low: in that scenario, the gap first increases and then decreases as r rises. Note that when α is very low, users can tolerate high ad intensities under both mechanisms. Starting with low r, an increase in the ad-revenue rate increases the revenue gap since GAS* can better screen users relative to BAS*. However, as r continues to rise, both mechanisms increasingly converge toward serving near-maximal ad-intensity to most user types, so the coarseness of the BAS* menu becomes less consequential and the gap Φ narrows.

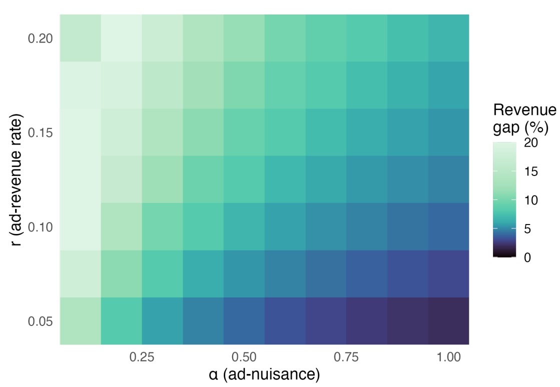


Figure 19 **Revenue gap Φ (in %) as a function of the ad-revenue rate r and ad-nuisance α. Each cell reports the average gap Φ computed over all remaining parameters (a, b, k). Lighter colors indicate a larger revenue gap.**

(ii) Effect of user-type distribution (a and b): The figure shows that the average gap increases with both a and b, but the effect of a is substantially more pronounced. To understand this effect, recall that user types are distributed as $\theta \sim \text{Beta}(a, b)$, so $\mathbb{E}[\theta] = a/(a+b)$. The ratio a/b can therefore be written as $\mathbb{E}[\theta]/(1-\mathbb{E}[\theta])$, which represents the odds of observing a high-valuation user relative to a low-valuation user. A higher value of a/b indicates a user population that is skewed toward higher valuations. We note here that a substantial revenue gap requires $a > 1$; when $a = 1$ (the leftmost column), the gap remains small for all values of b. This is because when a is small, the mass of high-type users is too thin for the GAS* mechanism to generate a meaningful revenue advantage. Conversely, as a increases (moving rightward), the population contains more high-type users among whom GAS* can better price-discriminate compared to BAS*.

The effect of b is more nuanced. Increasing b while holding a fixed shifts mass toward lower types (lowering $\mathbb{E}[\theta]$). Yet, we observe that the revenue gap still increases modestly. This is because increasing b (within the range considered) increases the dispersion of the user-type distribution, which in turn increases the proportion of intermediate type users that GAS* can screen more effectively than BAS*.

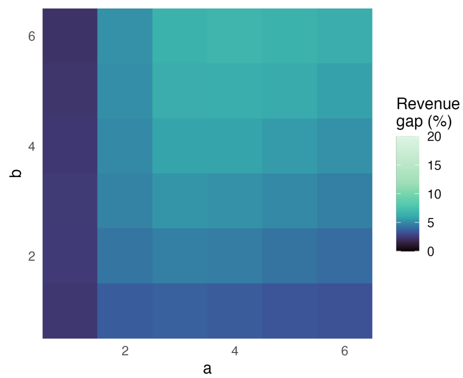


Figure 20 **Revenue gap Φ (in %) as a function of the user-type distribution modeled by $\mathrm{Beta}(a,b)$. Each cell reports the average gap Φ computed over all remaining parameters (α, r, k). Lighter colors indicate a larger revenue gap.**

(iii) Joint effect of ad attractiveness and user-type distribution (r/α and a/b): The classification tree in Step 1 suggests that ad attractiveness and the composition of user types interact: large revenue gaps arise when both drivers are favorable to GAS*. To examine this interaction directly, we construct two composite ratios that summarize each driver:

- The ratio r/α captures ad attractiveness: a higher value indicates that each ad impression generates more revenue (r) per unit of nuisance imposed on users (α). When r/α is large, advertising is an effective monetization lever, and the ability to fine-tune ad-intensity across user types becomes especially valuable.

- The ratio a/b captures the odds of high-valuation users: since $\theta \sim \mathrm{Beta}(a,b)$ implies $\mathbb{E}[\theta]/(1-\mathbb{E}[\theta]) = a/b$, a larger ratio indicates a population skewed toward higher valuations.

Figure 21 plots the average revenue gap Φ as a function of a/b (horizontal axis) and r/α (vertical axis). Since $a, b \in \{1, \ldots, 6\}$, the possible values of a/b lie in $[1/6, 6]$, with roughly half of these values concentrated below 1. The same holds for $r/\alpha \in [0.05, 2]$, because $r \in [0.05, 0.2]$ and $\alpha \in [0.1, 1]$ are each drawn from discrete grids. As a result, the observed grid points for both ratios are not evenly spaced. To avoid visually misleading spacing, we display the heatmap with cell dimensions proportional to the distance between adjacent grid values, so that the plot reflects a continuous scale.

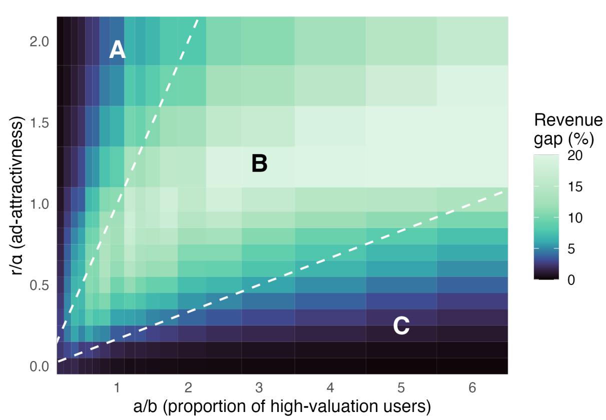
The heatmap reveals that the revenue gap is not simply increasing in both ratios. Instead, three distinct regions emerge, delineated by the dashed boundaries in the figure and summarized in the adjacent table. We describe each region in turn.

- **Region A** ($r/\alpha > a/b$; upper-left portion of the figure): When ad attractiveness is high relative to the proportion of high-type users, the revenue gap is moderate, ranging from $\approx 0\%$ to 13%. In this region, the ad revenue far exceeds the nuisance cost, so most user types face high ad intensities under both GAS* and BAS* mechanisms. This leaves limited room for the finer screening of GAS* to generate additional revenue. Thus, even if advertising is lucrative, the coarser BAS* menu extracts most of the available ad revenue, and the gap between the two mechanisms remains limited.

- **Region B** ($a/6b < r/\alpha \leq a/b$; the central band of the figure): This is the region where GAS* achieves its strongest advantage over BAS*, with the revenue gap ranging from $\approx 2\%$ to 25%. Here, ad attractiveness is high enough for advertising to serve as an effective monetization lever, yet not so high that both mechanisms converge to near-maximal ad intensities. Simultaneously, the user population contains a sufficient mass of high-type users for the continuous GAS* menu to screen effectively. The combination of these two conditions creates a setting in which the finer GAS* mechanism yields a substantial revenue advantage over the coarser BAS* mechanism.

- **Region C** ($r/\alpha \leq a/6b$; lower-right portion of the figure): When ad attractiveness is low relative to the user-type composition, the revenue gap is small, ranging from $\approx 0\%$ to 7%. In this region, ads are either low-paying or impose substantial nuisance on users. As a result, neither mechanism generates substantial ad revenue, and the continuous menu of GAS* offers little incremental value over BAS*.

Taken together, these three regions illustrate a clear complementarity between the two drivers. Neither ad attractiveness nor user-type composition alone is sufficient to generate a large revenue gap. The largest gaps arise when both r/α and a/b are moderate to high.



Region A ($\frac{r}{\alpha} > \frac{a}{b}$; gap ≈ 0–13%): Ad attractiveness is high enough for both mechanisms to push most users towards high ad intensities; this limits the scope for GAS* to screen better over BAS*.

Region B ($\frac{a}{6b} < \frac{r}{\alpha} \leq \frac{a}{b}$; gap ≈ 2–25%): Ads are effective and sufficient high-type users allow GAS* to better differentiate the users.

Region C ($\frac{r}{\alpha} \leq \frac{a}{6b}$; gap ≈ 0–7%): Ads are low-paying or impose substantial nuisance; neither mechanism generates substantial ad revenue.

Figure 21 **The left panel displays the revenue gap Φ (in %) as a joint function of ad attractiveness (r/α) and the user-type distribution (a/b). Each cell reports the average gap Φ computed over the remaining parameter k. Lighter colors indicate a larger revenue gap. The right panel describes the three regions delineated by the dashed boundaries in the figure.**

Appendix F: Proof of Theorem 2 (Capped Ad Intensity)

In this extension, we impose the additional constraint on the DCP that the ad intensities displayed to the users are capped at $n_{\max}$; that is, $0 \le n_{\text{GAS}}(\theta) \le n_{\max}$ for all $\theta \in [\theta_{\text{L}}, \theta_{\text{H}}]$ and some $n_{\max} \ge 0$. The revenue-maximization mechanism-design problem of the DCP can then be formulated as:

$$\max_{\substack{0 \le n_{\text{GAS}}(\cdot) \le n_{\max} \\ p_{\text{GAS}}(\cdot) \ge 0 \\ d_{\text{GAS}}(\cdot) \in \{0,1\}}} \int_{\theta_{\text{L}}}^{\theta_{\text{H}}} \left(\underbrace{p_{\text{GAS}}(\theta)}_{\text{subscription revenue}} + \underbrace{T\, r\, n_{\text{GAS}}(\theta)\, t^*_{\text{GAS}}(\theta)}_{\text{ad revenue}} \right) d_{\text{GAS}}(\theta) f(\theta) d\theta \qquad (\mathcal{P}\text{-GAS-C})$$

subject to:

$$V_{\text{GAS}}(\theta;\theta) \ge V_{\text{GAS}}(\hat{\theta};\theta) \;\; \forall \theta \in [\theta_{\text{L}}, \theta_{\text{H}}] \qquad \text{(incentive compatibility)}$$

$$d_{\text{GAS}}(\theta) = \begin{cases} 1 & \text{if } V_{\text{GAS}}(\theta;\theta) \ge 0, \\ 0 & \text{otherwise} \end{cases} \quad \forall \theta \in [\theta_{\text{L}}, \theta_{\text{H}}]. \qquad \text{(individual rationality)}$$

Following the arguments from the proof of Theorem 1, we first fix $d_{\text{GAS}}(\cdot)$ and optimize the relaxed objective point-wise for each value of θ by solving

$$\max_{0 \le n_{\text{GAS}}(\theta) \le n_{\max}} \left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2} \left(1 - \alpha n_{\text{GAS}}(\theta)_+\right)^2 + Trk\, n_{\text{GAS}}(\theta) \left(1 - \alpha n_{\text{GAS}}(\theta)\right)_+ \right) h_{\text{GAS}}(\theta)$$

Notice that if $n_{\max} > 1/\alpha$, then the objective function is 0 and the problem reduces to that in Theorem 1 and the original GAS* mechanism remains optimal. Thus, to rule out this trivial case, we will focus on the case where $0 \le n_{\max} \le 1/\alpha$. As before, let $g(n_{\text{GAS}}(\theta)) := \left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2} \left(1 - \alpha n_{\text{GAS}}(\theta)\right)^2 + Trk\, n_{\text{GAS}}(\theta) \left(1 - \alpha n_{\text{GAS}}(\theta)\right) \right) h_{\text{GAS}}(\theta)$ denote the objective function. Then, using the steps from Theorem 1, it is straightforward to see that the optimal solution to the above problem is:

$$n^*_{\text{GAS}}(\theta) = \begin{cases} \min\left\{ n_{\max}, \frac{r - \alpha\Psi_{\text{GAS}}(\theta)}{\alpha(2r - \alpha\Psi_{\text{GAS}}(\theta))} \right\} & \text{if } \Psi_{\text{GAS}}(\theta) < \frac{r}{\alpha} \\ 0 & \text{if } \Psi_{\text{GAS}}(\theta) \ge \frac{r}{\alpha}, \end{cases} \tag{55}$$

Let θ_0 denote user type for which $n_{\max} = \frac{r - \alpha\Psi_{\text{GAS}}(\theta_0)}{\alpha(2r - \alpha\Psi_{\text{GAS}}(\theta_0))}$. Then, $\Psi_{\text{GAS}}(\theta_0) = \frac{r(1 - 2\alpha n_{\max})}{\alpha(1 - \alpha n_{\max})} \le \frac{r}{\alpha}$. Thus, plugging the optimal value of $n^*_{\text{GAS}}(\theta)$ in the objective function, we get:

$$g(n^*_{\text{GAS}}(\theta)) = \begin{cases} \left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2} \left(1 - \alpha n_{\max}\right)^2 + Trk\, n_{\max} \left(1 - \alpha n_{\max}\right) \right) h_{\text{GAS}}(\theta) & \text{if } \Psi_{\text{GAS}}(\theta) < \Psi_{\text{GAS}}(\theta_0), \\ \frac{T\,k\,r^2\, h_{\text{GAS}}(\theta)}{2\alpha(2r - \alpha\Psi_{\text{GAS}}(\theta))} & \text{if } \Psi_{\text{GAS}}(\theta_0) \le \Psi_{\text{GAS}}(\theta) < \frac{r}{\alpha}, \\ \frac{T\,k\,\Psi_{\text{GAS}}(\theta) h_{\text{GAS}}(\theta)}{2} & \text{if } \Psi_{\text{GAS}}(\theta) \ge \frac{r}{\alpha}. \end{cases}$$

As before, let us define $G(\Psi_{\text{GAS}}(\theta)) := g(n^*_{\text{GAS}}(\theta))$. Note that when $\Psi_{\text{GAS}}(\theta) \ge \Psi_{\text{GAS}}(\theta_0)$, $G(\Psi_{\text{GAS}}(\theta)) \ge 0$. However, In contrast to the proof of Theorem 1, we note that $G(\Psi_{\text{GAS}}(\theta))$ may not always be greater than 0. In particular, when $\Psi_{\text{GAS}}(\theta) < \Psi_{\text{GAS}}(\theta_0)$, we have

$$G(\Psi_{\text{GAS}}(\theta)) = \left(\frac{Tk\Psi_{\text{GAS}}(\theta)}{2} \left(1 - \alpha n_{\max}\right)^2 + Trk\, n_{\max} \left(1 - \alpha n_{\max}\right) \right) h_{\text{GAS}}(\theta).$$

Note that $G(\cdot)$ is an affine function of its argument and let $\Psi^{\dagger}_{\text{GAS}}(\theta)$ be the point such that $G(\Psi^{\dagger}_{\text{GAS}}(\theta)) = 0$. Then,

$$\Psi^{\dagger}_{\text{GAS}}(\theta) = \frac{-2\, r\, n_{\max}}{1 - \alpha n_{\max}} < 0.$$

Thus, when $\Psi_{\text{GAS}}(\theta) < \Psi_{\text{GAS}}(\theta_0)$, we have $G(\Psi_{\text{GAS}}(\theta)) \geq 0$ for $\Psi_{\text{GAS}}(\theta) \geq \Psi^{\dagger}_{\text{GAS}}(\theta)$ and $G(\Psi_{\text{GAS}}(\theta)) < 0$ for $\Psi_{\text{GAS}}(\theta) < \Psi^{\dagger}_{\text{GAS}}(\theta)$. Summarizing, we have:

$$G(\Psi_{\text{GAS}}(\theta)) = \begin{cases} < 0 & \text{if } \Psi_{\text{GAS}}(\theta) < \frac{-2\, r\, n_{\max}}{1-\alpha n_{\max}} \\ \geq 0 & \text{if } \Psi_{\text{GAS}}(\theta) \geq \frac{-2\, r\, n_{\max}}{1-\alpha n_{\max}} \end{cases}.$$

Next, we substitute the optimal value of $n^*_{\text{GAS}}(\theta)$ and consider the optimization problem with respect to $d_{\text{GAS}}(\cdot)$

$$\max_{d_{\text{GAS}}(\cdot) \in \{0,1\}} \int_{\theta_{\text{L}}}^{\theta_{\text{H}}} G(\Psi_{\text{GAS}}(\theta)) d\theta$$

In the above problem, notice that when $\Psi_{\text{GAS}}(\theta) < \Psi^{\dagger}_{\text{GAS}}(\theta)$, then the integrand $G(\Psi_{\text{GAS}}(\theta)) < 0$ and excluding these users strictly raises the revenue. Thus, we have:

$$d^*(\theta) = \begin{cases} 0 & \text{if } \Psi_{\text{GAS}}(\theta) < \frac{-2\, r\, n_{\max}}{1-\alpha n_{\max}} \\ 1 & \text{if } \Psi_{\text{GAS}}(\theta) \geq \frac{-2\, r\, n_{\max}}{1-\alpha n_{\max}} \end{cases}$$

Notice that the set $\{\theta \mid \Psi_{\text{GAS}}(\theta) \geq \frac{-2\, r\, n_{\max}}{1-\alpha n_{\max}}\}$ need not be an interval and can therefore violate incentive compatibility.

Let us now construct a participation rule as follows. Define

$$\theta_{\text{C}} := \inf \left\{ \theta \mid \Psi_{\text{GAS}}(\theta) \geq \frac{-2\, r\, n_{\max}}{1 - \alpha n_{\max}} \right\}$$

and the participation rule as:

$$\hat{d}(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_{\text{C}} \\ 1 & \text{if } \theta \geq \theta_{\text{C}}. \end{cases}$$

Here, observe that (i) If $d^*(\theta) = 1$ and $\hat{d}(\theta) = 0$ (i.e., $\theta < \theta_{\text{C}}$), then we have $\Psi_{\text{GAS}}(\theta) < \frac{-2\, r\, n_{\max}}{1-\alpha n_{\max}}$ which implies $G(\Psi_{\text{GAS}}(\theta)) < 0$. Hence, removing that user (setting $\hat{d}(\theta) = 0$) increases the total revenue. (ii) For every type θ such that $d^*(\theta) = 0$ and $\hat{d}(\theta) = 1$ (i.e., $\theta \geq \theta_{\text{C}}$), we have $\Psi_{\text{GAS}}(\theta) \geq \frac{-2\, r\, n_{\max}}{1-\alpha n_{\max}}$ which implies $G(\Psi_{\text{GAS}}(\theta)) \geq 0$. Thus, adding that user cannot reduce total revenue. Since $\hat{d}(\theta)$ differs from $d^*(\theta)$ only in these two ways, the participation decision rule $\hat{d}(\theta)$ yields a weakly higher revenue and is thus optimal among all incentive-compatible participation decision rules.

For the participation rule $\hat{d}(\theta)$, we have:

$$h_{\text{GAS}}(\theta) = \hat{d}(\theta) f(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_{\text{C}} \\ f(\theta) & \text{if } \theta \geq \theta_{\text{C}}. \end{cases}$$

Consequently, for all participating users ($\theta \geq \theta_{\text{C}}$), we have

$$\Psi_{\text{GAS}}(\theta) = \theta - \frac{\int_{\theta}^{\theta_{\text{H}}} h_{\text{GAS}}(x) dx}{h_{\text{GAS}}(\theta)} = \theta - \frac{\int_{\theta}^{\theta_{\text{H}}} f(x) dx}{f(\theta)} = \psi(\theta).$$

Thus $\Psi_{\text{GAS}}(\theta) \geq \frac{-2\,r\,n_{\max}}{1-\alpha n_{\max}}$ is equivalent to $\psi(\theta) \geq \frac{-2\,r\,n_{\max}}{1-\alpha n_{\max}}$. Further, since $\psi(\theta)$ is monotone increasing and $\theta \in [\theta_{\text{L}}, \theta_{\text{H}}]$, we have $\theta_{\text{C}} = \max\left\{\theta_{\text{L}}, \psi^{-1}\left(\frac{-2\,r\,n_{\max}}{1-\alpha n_{\max}}\right)\right\}$.

Finally, notice that since the low-type users $\theta < \theta_{\text{C}}$ do not participate, their utility is 0. A revenue-maximizing mechanism will therefore have the individual-rationality constraint for indifferent type user θ_{C} to be $V_{\text{GAS}}(\theta_{\text{C}}) = 0$. Consequently, the subscription fee function can be obtained as $p^*_{\text{GAS}}(\theta) = \frac{\theta Tk\left(1-\alpha n^*_{\text{GAS}}(\theta)\right)^2}{2} - \int\limits_{\theta_{\text{C}}}^{\theta} \frac{Tk\left(1-\alpha n^*_{\text{GAS}}(x)\right)^2}{2} dx$.

Consider the mechanism GAS-C* defined as follows:

$$n^*_{\text{GAS-C}}(\theta) = \begin{cases} \min\left\{n_{\max}, \frac{r-\alpha\psi(\theta)}{\alpha(2r-\alpha\psi(\theta))}\right\} & \text{if } \theta < \psi^{-1}\left(\frac{r}{\alpha}\right) \\ 0 & \text{if } \theta \geq \psi^{-1}\left(\frac{r}{\alpha}\right), \end{cases} \tag{56}$$

$$p^*_{\text{GAS-C}}(\theta) = \frac{\theta Tk\left(1-\alpha n^*_{\text{GAS-C}}(\theta)\right)^2}{2} - \int\limits_{\theta_{\text{C}}}^{\theta} \frac{Tk\left(1-\alpha n^*_{\text{GAS-C}}(x)\right)^2}{2} dx, \tag{57}$$

$$d^*_{\text{GAS-C}}(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_{\text{C}} \\ 1 & \text{if } \theta \geq \theta_{\text{C}}. \end{cases} \tag{58}$$

Using arguments analogous to Lemma 1, it is straightforward to show that the mechanism $(n^*_{\text{GAS-C}}(\cdot), p^*_{\text{GAS-C}}(\cdot), d^*_{\text{GAS-C}}(\cdot))$ satisfies the incentive compatibility and individual rationality constraints. Consequently, GAS-C* is an optimal solution to problem ($\mathcal{P}$-GAS-C) in the setting with capped ad intensity.
■

Appendix G: Proof of Theorem 3 (Endogenous Ad-Revenue Rate)

Recall that the ad-revenue rate is modeled as a linear decreasing function of ad intensity: $r(n) = r_0 - \beta n$, where $0 \leq \beta \leq \alpha r_0$ captures the marginal decrease in per-unit ad revenue due to ad clutter. Further recall that with endogenous ad revenue, the DCP's problem of maximizing expected revenue under the GAS mechanism becomes:

$$\max_{\substack{n_{\text{GAS}}(\cdot)\geq 0\\ p_{\text{GAS}}(\cdot)\geq 0\\ d_{\text{GAS}}(\cdot)\in\{0,1\}}} \int_{\theta_{\text{L}}}^{\theta_{\text{H}}} \left(\underbrace{p_{\text{GAS}}(\theta)}_{\text{subscription revenue}} + \underbrace{T\,(r_0 - \beta n_{\text{GAS}}(\theta))\, n_{\text{GAS}}(\theta)\, t^*_{\text{GAS}}(\theta)}_{\text{ad revenue}}\right) d_{\text{GAS}}(\theta) f(\theta) d\theta \tag{$\mathcal{P}$-GAS-E}$$

subject to:

$$V_{\text{GAS}}(\theta;\theta) \geq V_{\text{GAS}}(\hat{\theta};\theta)\ \forall \theta \in [\theta_{\text{L}}, \theta_{\text{H}}] \tag{incentive compatibility}$$

$$d_{\text{GAS}}(\theta) = \begin{cases} 1 & \text{if } V_{\text{GAS}}(\theta;\theta) \geq 0, \\ 0 & \text{otherwise} \end{cases} \quad \forall \theta \in [\theta_{\text{L}}, \theta_{\text{H}}]. \tag{individual rationality}$$

Following the approach outlined in the proof of Theorem 1 in Appendix A, we relax the incentive compatibility and individual rationality constraints and express the problem in terms of the ad-intensity function $n_{\text{GAS}}(\theta)$. The relaxed problem now takes the following form:

$$\max_{\substack{0\le n_{\mathrm{GAS}}(\cdot)\le \frac{1}{\alpha}\\ d_{\mathrm{GAS}}(\cdot)\in\{0,1\}}} \int_{\theta_{\mathrm{L}}}^{\theta_{\mathrm{H}}} \left(\frac{Tk\Psi_{\mathrm{GAS}}(\theta)}{2}\left(1-\alpha n_{\mathrm{GAS}}(\theta)\right)^2 + Tk\left(r_0-\beta n_{\mathrm{GAS}}(\theta)\right) n_{\mathrm{GAS}}(\theta)\left(1-\alpha n_{\mathrm{GAS}}(\theta)\right)\right) h_{\mathrm{GAS}}(\theta)d\theta \tag{59}$$

To solve the above problem, we again apply point-wise maximization. Fixing θ and $\Psi_{\mathrm{GAS}}(\theta)$, define the objective function:

$$g(n_{\mathrm{GAS}}(\theta)) := \frac{\Psi_{\mathrm{GAS}}(\theta)Tk}{2}(1-\alpha n)^2 + T\,k\,n_{\mathrm{GAS}}(\theta)(1-\alpha n_{\mathrm{GAS}}(\theta))(r_0-\beta n),$$

where $n_{\mathrm{GAS}}(\cdot)\in[0,1/\alpha]$. We will use the shorthand $g(n)$ for $g(n_{\mathrm{GAS}}(\theta))$.
Taking the first derivative:

$$\begin{aligned}
g'(n) &= -\alpha\Psi_{\mathrm{GAS}}(\theta)Tk(1-\alpha n) + Tk\left((1-2\alpha n)(r_0-\beta n)-\beta n(1-\alpha n)\right)\\
&= Tk\left(-\alpha\Psi_{\mathrm{GAS}}(\theta)(1-\alpha n)+(r_0-\beta n)(1-2\alpha n)-\beta n(1-\alpha n)\right)\\
&= Tk\left(3\alpha\beta n^2+\left(\alpha^2\Psi_{\mathrm{GAS}}(\theta)-2\alpha r_0-2\beta\right)n+(r_0-\alpha\Psi_{\mathrm{GAS}}(\theta))\right).
\end{aligned}$$

The second-order derivative is:

$$\begin{aligned}
g''(n) &= Tk\left(\alpha^2\Psi_{\mathrm{GAS}}(\theta)+\frac{d}{dn}\left((r_0-\beta n)(1-2\alpha n)-\beta n(1-\alpha n)\right)\right)\\
&= Tk\left(\alpha^2\Psi_{\mathrm{GAS}}(\theta)-2\alpha(r_0-\beta n)-\beta(1-2\alpha n)-\beta(1-\alpha n)+\alpha\beta n\right)\\
&= Tk\left(\alpha^2\Psi_{\mathrm{GAS}}(\theta)-2\alpha r_0+2\alpha\beta n-\beta+2\alpha\beta n-\beta+\alpha\beta n\right)\\
&= Tk\left(\alpha^2\Psi_{\mathrm{GAS}}(\theta)-2\alpha r_0-2\beta+6\alpha\beta n\right).
\end{aligned}$$

Since the second-order derivative is affine in n, the function $g(\cdot)$ is at most uni-modal. Thus, to find the maximum value of $g(n)$, we first find all the stationary points by solving $g'(n)=0$. Then, we keep only those points that are local maximal ($g''(n)\le 0$). Finally, we compare their objective value with the boundary points 0 and $1/\alpha$.

The equation $g'(n)=0$ is a quadratic:

$$3\alpha\beta n^2+\left(\alpha^2\Psi_{\mathrm{GAS}}(\theta)-2\alpha r_0-2\beta\right)n+(r_0-\alpha\Psi_{\mathrm{GAS}}(\theta))=0. \tag{60}$$

The discriminant of the above quadratic is:

$$\Delta = \left(\alpha^2\Psi_{\mathrm{GAS}}(\theta)-2\alpha r_0-2\beta\right)^2-12\alpha\beta\left(r_0-\alpha\Psi_{\mathrm{GAS}}(\theta)\right).$$

Case 1: If $\Psi_{\mathrm{GAS}}(\theta)\ge\frac{r_0}{\alpha}$, then $r_0-\alpha\Psi_{\mathrm{GAS}}(\theta)\le 0$, and the quadratic has no positive root in $[0,1/\alpha]$. In this region, the slope at $n=0$ is non-positive ($g'(0)=r_0-\alpha\Psi_{\mathrm{GAS}}(\theta)\le 0$), so $g(\cdot)$ is (weakly) decreasing on $[0,1/\alpha]$, and its maximum is attained at the left boundary. Thus, we get

$$n^*_{\mathrm{GAS}}(\theta)=0. \tag{61}$$

Case 2: If $\Psi_{\text{GAS}}(\theta) < \frac{r_0}{\alpha}$, then $r_0 - \alpha\Psi_{\text{GAS}}(\theta) > 0$. In this case, notice that

$$g'(0) = Tk(r_0 - \alpha\Psi_{\text{GAS}}(\theta)) > 0,$$
$$g'(1/\alpha) = Tk\left(\frac{\beta}{\alpha} - r_0\right) \le 0 \quad (\text{since } \beta \le \alpha r_0).$$

Since g' is continuous, it changes sign on $[0, 1/\alpha]$. By the Intermediate Value Theorem, the equation $g'(n) = 0$ must have at least one root in the open interval $(0, 1/\alpha)$. Moreover, since $g'(n)$ is a quadratic function, it has two real roots if and only if its discriminant is strictly positive. Therefore, the discriminant Δ must be strictly positive whenever $\Psi_{\text{GAS}}(\theta) < \frac{r_0}{\alpha}$. Among the two roots, only one lies in $(0, 1/\alpha)$, and this is the root we select as the optimal solution. Solving, we get:

$$\begin{aligned} n^*_{\text{GAS}}(\theta) &= \frac{2\alpha r_0 + 2\beta - \alpha^2\Psi_{\text{GAS}}(\theta) - \sqrt{\Delta}}{6\alpha\beta} \\ &= \frac{2\alpha r_0 + 2\beta - \alpha^2\Psi_{\text{GAS}}(\theta) - \sqrt{\left(\alpha^2\Psi_{\text{GAS}}(\theta) - 2\alpha r_0 - 2\beta\right)^2 - 12\alpha\beta\left(r_0 - \alpha\Psi_{\text{GAS}}(\theta)\right)}}{6\alpha\beta}. \end{aligned}$$

At $n^*_{\text{GAS}}(\theta)$, the second derivative is:

$$g''(n^*_{\text{GAS}}(\theta)) = -\sqrt{\Delta} = \alpha^2\Psi_{\text{GAS}}(\theta) - 2\alpha r_0 - 2\beta + 6\alpha\beta n^*_{\text{GAS}}(\theta) < 0,$$

which shows that $n^*_{\text{GAS}}(\theta)$ is a local maximum. Because $g(1/\alpha) = 0$ and $g(0) = \frac{1}{2}\Psi_{\text{GAS}}(\theta) > 0$, direct evaluation shows $g(n^*_{\text{GAS}}(\theta)) \ge g(0)$; therefore, the global maximizer in $[0, 1/\alpha]$ is $n^*_{\text{GAS}}(\theta)$.

Combining the above two cases, we get:

$$n^*_{\text{GAS}}(\theta) = \begin{cases} \dfrac{2\alpha r_0 + 2\beta - \alpha^2\Psi_{\text{GAS}}(\theta) - \sqrt{\left(\alpha^2\Psi_{\text{GAS}}(\theta) - 2\alpha r_0 - 2\beta\right)^2 - 12\alpha\beta\left(r_0 - \alpha\Psi_{\text{GAS}}(\theta)\right)}}{6\alpha\beta}, & \Psi_{\text{GAS}}(\theta) < \frac{r_0}{\alpha}, \\ 0, & \Psi_{\text{GAS}}(\theta) \ge \frac{r_0}{\alpha}, \end{cases}$$

Next, we plug the value of $n^*_{\text{GAS}}(\theta)$ into the objective function and optimize it with respect to the participation rule $d_{\text{GAS}}(\cdot)$. The steps here proceed exactly as in Appendix A. By repeating the "flip the participation decision of an infinitesimal mass of users" argument, we can show that the firm can improve the revenue by making the users participate. Therefore, full participation, $d^*_{\text{GAS}}(\theta) = 1$ for all θ, is an optimal in this endogenous ad-revenue setting. Since $d^*_{\text{GAS}}(\theta) = 1$ for all θ and $\psi(\cdot)$ is monotonically increasing, we get $\Psi_{\text{GAS}}(\theta) = \psi(\theta)$. Consequently, $n^*_{\text{GAS}}(\theta)$ can be simplified to:

$$n^*_{\text{GAS}}(\theta) = \begin{cases} \dfrac{2\alpha r_0 + 2\beta - \alpha^2\psi(\theta) - \sqrt{\left(\alpha^2\psi(\theta) - 2\alpha r_0 - 2\beta\right)^2 - 12\alpha\beta\left(r_0 - \alpha\psi(\theta)\right)}}{6\alpha\beta}, & \theta < \psi^{-1}\left(\frac{r_0}{\alpha}\right), \\ 0, & \theta \ge \psi^{-1}\left(\frac{r_0}{\alpha}\right), \end{cases}$$

We now show that $n^*_{\text{GAS}}(\theta)$ is monotonically decreasing in θ. Consider the derivative:

$$\frac{d\,n^*_{\text{GAS}}(\theta)}{d\theta} = -\frac{\alpha^2\,\psi'(\theta)}{6\alpha\beta}\left(1 + \frac{\alpha^2\psi(\theta) - 2\alpha r_0 + 4\beta}{\sqrt{\left(\alpha^2\psi(\theta) - 2\alpha r_0 - 2\beta\right)^2 - 12\alpha\beta\left(r_0 - \alpha\psi(\theta)\right)}}\right).$$

Next, note that

$$(\alpha^2\psi - 2\alpha r_0 - 2\beta)^2 - 12\alpha\beta\,(r_0 - \alpha\psi) - (\alpha^2\psi - 2\alpha r_0 + 4\beta)^2 = 12\beta(\alpha r_0 - \beta) \geq 0,$$

since $0 \leq \beta \leq \alpha r_0$. Hence

$$\sqrt{(\alpha^2\psi - 2\alpha r_0 - 2\beta)^2 - 12\alpha\beta\,(r_0 - \alpha\psi)} \geq \left|\alpha^2\psi - 2\alpha r_0 + 4\beta\right|,$$

which implies the quantity in parentheses above is strictly positive. Because $\psi'(\theta) \geq 0$ and $\beta > 0$, we have $\frac{d\,n^*_{\text{GAS}}(\theta)}{d\theta} \leq 0$. Thus $n^*_{\text{GAS}}(\theta)$ is weakly decreasing in θ.

Consider the mechanism GAS-E* defined by:

$$n^*_{\text{GAS-E}}(\theta) = \begin{cases} \dfrac{2\alpha r_0 + 2\beta - \alpha^2\psi(\theta) - \sqrt{\left(\alpha^2\psi(\theta) - 2\alpha r_0 - 2\beta\right)^2 - 12\alpha\beta\,(r_0 - \alpha\psi(\theta))}}{6\alpha\beta}, & \theta < \psi^{-1}\left(\frac{r_0}{\alpha}\right), \\ 0, & \theta \geq \psi^{-1}\left(\frac{r_0}{\alpha}\right), \end{cases} \tag{62}$$

$$p^*_{\text{GAS-E}}(\theta) = \frac{Tk\theta}{2}\left(1 - \alpha n^*_{\text{GAS-E}}(\theta)\right)^2 - \int_{\theta_{\text{L}}}^{\theta} \frac{Tk}{2}\left(1 - \alpha n^*_{\text{GAS-E}}(x)\right)^2 dx, \tag{63}$$

$$d^*_{\text{GAS-E}}(\theta) = 1. \tag{64}$$

Using arguments analogous to Lemma 1, it is straightforward to show that the mechanism $(n^*_{\text{GAS-E}}(\cdot), p^*_{\text{GAS-E}}(\cdot), d^*_{\text{GAS-E}}(\cdot))$ satisfies the incentive compatibility and individual rationality constraints. Consequently, GAS-E* is an optimal solution to problem ($\mathcal{P}$-GAS-E) in the setting with endogenous ad-revenue rates. ■

Appendix H: An Approximation Bound for Two-Dimensional Heterogeneity

In this section, we extend the GAS mechanism to allow ad-nuisance α to be heterogeneous and privately known. Thus, each user is endowed with two-dimensional private information, (θ, α). Let $h(\theta, \alpha)$ and $H(\theta, \alpha)$ denote the joint density function and cumulative distribution function on the support $[\theta_{\text{L}}, \theta_{\text{H}}] \times [\alpha_{\text{L}}, \alpha_{\text{H}}]$, respectively. The DCP observes the joint distribution, but not individual realizations of (θ, α). With two-dimensional private information, the exact characterization of a revenue-maximizing GAS mechanism is intractable. To address this difficulty, we develop an approximation bound – inspired by the discretize-and-discount technique in (Madarász and Prat 2017) and its application in (Mehta et al. 2022) – for a simple posted-menu mechanism. In particular, our construction produces a posted-price menu consisting of a finite collection of ad-intensity–subscription-fee pairs (n, p) from which users self-select, and we provide an explicit additive bound on the resulting revenue loss relative to the optimal two-dimensional mechanism.

Extending the notation in (65) to the two-dimensional type space, let $V_\mu(\hat{\theta}, \hat{\alpha}; \theta, \alpha)$ denote the net utility of a user (conditional on participation) of true type (θ, α) who reports $(\hat{\theta}, \hat{\alpha})$ and consumes the corresponding optimal usage level. Here, $n_\mu(\hat{\theta}, \hat{\alpha})$ and $p_\mu(\hat{\theta}, \hat{\alpha})$ denote, respectively, the ad intensity and subscription fee assigned by mechanism μ when the report is $(\hat{\theta}, \hat{\alpha})$. Substituting the optimal usage decision into the user's utility function yields

$$V_\mu(\hat{\theta}, \hat{\alpha}; \theta, \alpha) = \tfrac{1}{2}T\theta k\big(1 - \alpha n_\mu(\hat{\theta}, \hat{\alpha})\big)_+^2 - p_\mu(\hat{\theta}, \hat{\alpha}). \tag{65}$$

Given a mechanism $(n_\mu(\theta,\alpha), p_\mu(\theta,\alpha), d_\mu(\theta,\alpha))$, a user of type (θ,α) who participates $(d_\mu(\theta,\alpha)=1)$ chooses optimal content consumption $t^*_\mu(\theta,\alpha) = k\big(1-\alpha n_\mu(\theta,\alpha)\big)_+$. If a user does not participate $(d_\mu(\theta,\alpha)=0)$, then the utility from outside option is normalized to zero. With this notation in place, the DCP's revenue-maximization problem in the two-dimensional type space can be written as

$$\max_{\substack{0\le n_\mu(\cdot,\cdot)\le \bar{n}\\ d_\mu(\cdot,\cdot)\in\{0,1\}}} \int_{\theta_{\mathrm{L}}}^{\theta_{\mathrm{H}}}\int_{\alpha_{\mathrm{L}}}^{\alpha_{\mathrm{H}}} \Bigg(\underbrace{p_\mu(\theta,\alpha)}_{\text{subscription revenue}} + \underbrace{T\,r\,n_\mu(\theta,\alpha)\,t^*_\mu(\theta,\alpha)}_{\text{ad revenue}} \Bigg) d_\mu(\theta,\alpha)\,h(\theta,\alpha)\,d\theta\,d\alpha \tag{$\mathcal{P}$-2D}$$

subject to:

$$d_\mu(\theta,\alpha)\,V_\mu(\theta,\alpha;\theta,\alpha) \ge d_\mu(\hat\theta,\hat\alpha)\,V_\mu(\hat\theta,\hat\alpha;\theta,\alpha) \quad \forall(\theta,\alpha),(\hat\theta,\hat\alpha), \qquad \text{(incentive compatibility)}$$

$$d_\mu(\theta,\alpha)\,V_\mu(\theta,\alpha;\theta,\alpha) \ge 0 \quad \forall(\theta,\alpha). \qquad \text{(individual rationality)}$$

In this formulation, we allow subscription fees $p_\mu(\cdot,\cdot)$ to take arbitrary real values and do not restrict them to be nonnegative. This enlarges the feasible set of mechanisms and therefore can only strengthen upper bound on the revenue that we want to obtain. The allowance of negative fees is purely technical and is convenient for the uniform discounting step in Step 3 of the approximation, where menu prices may become negative.

Let $\pi^*_{2\mathrm{D}}$ denote the optimal value of Problem ($\mathcal{P}$-2D). An exact solution to Problem ($\mathcal{P}$-2D) is analytically intractable, as the incentive compatibility constraints must hold simultaneously across both θ and α. For the ease of analysis, we impose a uniform upper bound on the ad intensity chosen by the DCP. Let $\bar{n}$ denote such an upper bound on $n_\mu(\hat\theta,\hat\alpha)$, so that $n_\mu(\hat\theta,\hat\alpha)\in[0,\bar{n}]$ for all reports $(\hat\theta,\hat\alpha)$. We will use this bound to obtain uniform Lipschitz constants for users' utilities and for the ad-revenue term, which will be useful in bounding the approximation error of the mechanism.

Since $\big(1-\alpha n_\mu(\hat\theta,\hat\alpha)\big)_+\in[0,1]$, we have for any θ,θ',

$$\begin{aligned}\big|V_\mu(\hat\theta,\hat\alpha;\theta,\alpha) - V_\mu(\hat\theta,\hat\alpha;\theta',\alpha)\big| &= \tfrac{1}{2}Tk\,\big|\theta-\theta'\big|\,\Big(1-\alpha n_\mu(\hat\theta,\hat\alpha)\Big)_+^2\\ &\le \tfrac{1}{2}Tk\,\big|\theta-\theta'\big|.\end{aligned}$$

Thus, the utility function is Lipschitz continuous in θ with constant $\kappa_\theta := \frac{1}{2}Tk$. For α,α', we compute

$$\begin{aligned}\big|V_\mu(\hat\theta,\hat\alpha;\theta,\alpha) - V_\mu(\hat\theta,\hat\alpha;\theta,\alpha')\big| &= \tfrac{1}{2}T\theta k\,\Big|\big(1-\alpha n_\mu(\hat\theta,\hat\alpha)\big)_+^2 - \big(1-\alpha' n_\mu(\hat\theta,\hat\alpha)\big)_+^2\Big|\\ &\le T\theta k\,n_\mu(\hat\theta,\hat\alpha)\,|\alpha-\alpha'|\\ &\le Tk\,\theta_{\mathrm{H}}\,\bar{n}\,|\alpha-\alpha'|.\end{aligned}$$

Hence the utility function is Lipschitz continuous in α with constant $\kappa_\alpha := Tk\,\theta_{\mathrm{H}}\,\bar{n}$. Finally, for α,α' and any $n\in[0,\bar{n}]$,

$$\begin{aligned}\big|Trkn(1-\alpha n)_+ - Trkn(1-\alpha' n)_+\big| &\le Trk\,n\,\big|(1-\alpha n)_+ - (1-\alpha' n)_+\big|\\ &\le Trk\,n^2\,|\alpha-\alpha'| \;\le\; Trk\,\bar{n}^2\,|\alpha-\alpha'|.\end{aligned}$$

This shows that the ad-revenue term is Lipschitz continuous in α with constant $\gamma_\alpha := Trk\bar{n}^2$. The κ_θ, κ_α, and γ_α quantify the sensitivity of utility and ad-revenue to the true values of θ and α, and will be useful in bounding the approximation error of the mechanism.

With these regularity conditions in place, we are now ready to develop the approximation bound. Before proceeding, we outline the three key steps of the procedure. First, the two-dimensional type space (θ, α) is discretized into a finite grid of representative cells, so that nearby types are grouped together. Second, the DCP solves a relaxed mechanism design problem defined only over these representative types, yielding a finite menu of ad-intensity–subscription-fee pairs that is incentive compatible within the grid. Finally, the menu prices are uniformly discounted to obtain a global deviation bound across the entire type space, which allows us to relate the discrete objective to the realized revenue under the posted menu.

H.1. Step 1: Discretization of the Type Space

The first step is to partition the continuous two-dimensional type space $[\theta_\mathrm{L}, \theta_\mathrm{H}] \times [\alpha_\mathrm{L}, \alpha_\mathrm{H}]$ into a finite grid of representative points. Let M_θ and M_α denote, respectively, the number of partitions along the θ and α dimensions. Let $\frac{\theta_\mathrm{H}-\theta_\mathrm{L}}{M_\theta}$ and $\frac{\alpha_\mathrm{H}-\alpha_\mathrm{L}}{M_\alpha}$ denote the grid intervals in the θ and α dimension, respectively. Next, we construct a partition of the type space into $M_\theta \times M_\alpha$ rectangular cells. For each cell (i, j), where $i \in \{1, \dots, M_\theta\}$ and $j \in \{1, \dots, M_\alpha\}$, we select the point

$$(\theta^i, \alpha^j) = \left(\theta_\mathrm{L} + (i-1)\cdot \frac{\theta_\mathrm{H}-\theta_\mathrm{L}}{M_\theta},\ \alpha_\mathrm{L} + (j-1)\cdot \frac{\alpha_\mathrm{H}-\alpha_\mathrm{L}}{M_\alpha}\right) \tag{66}$$

as the representative type[13]. By construction, every true type (θ, α) lies within distance at most $\frac{\theta_\mathrm{H}-\theta_\mathrm{L}}{M_\theta}$ in the θ dimension and $\frac{\alpha_\mathrm{H}-\alpha_\mathrm{L}}{M_\alpha}$ in the α dimension from its cell's representative point (θ^i, α^j). The Lipschitz continuity of V_μ then implies that the difference in utilities between any true type and its representative is uniformly bounded. Specifically, for any fixed report $(\hat{\theta}, \hat{\alpha})$,

$$\left|V_\mu(\hat{\theta}, \hat{\alpha}; \theta, \alpha) - V_\mu(\hat{\theta}, \hat{\alpha}; \theta^i, \alpha^j)\right| \le \kappa_\theta |\theta - \theta^i| + \kappa_\alpha |\alpha - \alpha^j|.$$

Since $|\theta - \theta^i| \le \frac{\theta_\mathrm{H}-\theta_\mathrm{L}}{M_\theta}$ and $|\alpha - \alpha^j| \le \frac{\alpha_\mathrm{H}-\alpha_\mathrm{L}}{M_\alpha}$ for all (θ, α) in the same cell, we get:

$$\left|V_\mu(\hat{\theta}, \hat{\alpha}; \theta, \alpha) - V_\mu(\hat{\theta}, \hat{\alpha}; \theta^i, \alpha^j)\right| \le \kappa_\theta \cdot \frac{\theta_\mathrm{H}-\theta_\mathrm{L}}{M_\theta} + \kappa_\alpha \cdot \frac{\alpha_\mathrm{H}-\alpha_\mathrm{L}}{M_\alpha}. \tag{67}$$

Define $\Delta_V := \kappa_\theta \cdot \frac{\theta_\mathrm{H}-\theta_\mathrm{L}}{M_\theta} + \kappa_\alpha \cdot \frac{\alpha_\mathrm{H}-\alpha_\mathrm{L}}{M_\alpha}$. This bound captures the maximum deviation in utility between a continuous type and its discrete representative.

H.2. Step 2: Solving the Discretized Problem for Representative Types

Next, we build a discrete model from the grid in Step 1 by assigning to each representative type (θ^i, α^j) the probability mass of its cell:

$$\rho^{ij} := \int_{\theta_\mathrm{L}+(i-1)\frac{\theta_\mathrm{H}-\theta_\mathrm{L}}{M_\theta}}^{\theta_\mathrm{L}+i\frac{\theta_\mathrm{H}-\theta_\mathrm{L}}{M_\theta}} \int_{\alpha_\mathrm{L}+(j-1)\frac{\alpha_\mathrm{H}-\alpha_\mathrm{L}}{M_\alpha}}^{\alpha_\mathrm{L}+j\frac{\alpha_\mathrm{H}-\alpha_\mathrm{L}}{M_\alpha}} h(\theta, \alpha)\, d\alpha\, d\theta. \tag{68}$$

[13] As in the discretization step in Madarász and Prat (2017), any point within a cell can serve as its representative; we fix the lower-left corner for notational convenience.

For notational simplicity, in this discrete problem we write n^{ij}, p^{ij}, and d^{ij} for the ad intensity, subscription fee, and participation decision assigned to type (θ^i, α^j), respectively. The DCP solves:

$$\max_{\substack{0\leq n^{ij}\leq \bar{n} \\ d^{ij}\in\{0,1\}}} \quad \sum_{i=1}^{M_\theta}\sum_{j=1}^{M_\alpha}\Big(p^{ij} + Trk\, n^{ij}\big(1-\alpha^j n^{ij}\big)_+\Big)\, d^{ij}\,\rho^{ij} \tag{$\mathcal{P}$-Rep}$$

$$\text{subject to} \quad d^{ij}\, V_\mu(\theta^i,\alpha^j;\theta^i,\alpha^j) \geq d^{i'j'}\, V_\mu(\theta^{i'},\alpha^{j'};\theta^i,\alpha^j) \quad \forall\, (i,j),(i',j') \in \{1,\ldots,M_\theta\}\times\{1,\ldots,M_\alpha\}, \tag{IC-Rep}$$

$$d^{ij}\, V_\mu(\theta^i,\alpha^j;\theta^i,\alpha^j) \geq 0 \quad \forall\, (i,j) \in \{1,\ldots,M_\theta\}\times\{1,\ldots,M_\alpha\}. \tag{IR-Rep}$$

Similar to that in Problem $\mathcal{P}$-2D, we allow subscription fees p^{ij} to be arbitrary real values, which enlarges the feasible set of Problem ($\mathcal{P}$-Rep) and is convenient for obtaining discounted menu in Step 3. An optimal solution to the discrete problem defines a finite menu indexed by grid cells. Let

$$\mathcal{M} := \Big\{(n_*^{ij}, p_*^{ij}) \,:\, d_*^{ij} = 1\Big\} \tag{69}$$

denote the set of offered ad-intensity–subscription-fee pairs. Thus, the mechanism resulting from Step 2 is a posted-menu mechanism: any user may either select one pair from $\mathcal{M}$ or opt out. Importantly, the incentive constraints in this step are imposed only within the discrete grid (among the representative types), and therefore this step alone does not guarantee global incentive compatibility for all continuous types (θ, α).

H.3. Step 3: Obtaining the Discounted Menu

In this step, we convert the grid-based solution from Step 2 into a posted menu that applies to all types in the continuous type space. To this end, we apply the discounting idea of Madarász and Prat (2017). Starting from the optimal discrete solution $(n_*^{ij}, p_*^{ij}, d_*^{ij})$ to Problem ($\mathcal{P}$-Rep), we keep the ad intensities n_*^{ij} fixed and discount each subscription fee by a constant fraction of the revenue generated by the corresponding menu item. Formally, for a discount parameter $\eta \in (0,1]$, we reduce each fee by η times its associated revenue.

This transformation has two effects. First, it gives up an η-fraction of revenue on the item that is ultimately chosen. Second, and more importantly, it increases the utility of every menu item by an amount proportional to the revenue that item generates, uniformly across all types. As a result, menu items that are more revenue-generating in the discrete problem become relatively more attractive to all users under the discounted menu. Consequently, if a user with a true type (θ, α) deviates from the contract intended for its grid cell, such a deviation must be toward another contract whose revenue performance in the discrete solution is similar. The discount parameter η therefore governs a tradeoff between the revenue lost directly through discounting and the degree to which off-grid deviations can be restricted to contracts with comparable revenue consequences.

We denote the outside option (no participation) by $\varnothing$, which yields zero utility to the user and zero revenue to the DCP. Fix any finite menu $\mathcal{M} := \{(n^{ij}, p^{ij})\}_{i,j} \cup \{\varnothing\}$ with $n^{ij} \in [0, \bar{n}]$ and $p^{ij} \in \mathbb{R}$ (i.e., we allow for negative subscription fees). For each (i,j), define the revenue from representative type (i,j) as

$$\pi^{ij} := p^{ij} + Trk\, n^{ij}\big(1-\alpha^j n^{ij}\big)_+, \tag{70}$$

and define $\pi^{\varnothing} := 0$. Let $\overline{R} := \frac{1}{2}Tk\theta_{\mathrm{H}} + Trk\bar{n}$ denote an upper bound such that $\pi^{ij} \le \overline{R}$ for all (i,j). The upper bound follows from individual rationality ($p^{ij} \le \frac{1}{2}Tk\theta_{\mathrm{H}}$) and the bound $n^{ij} \le \bar{n}$. Given $\eta \in (0,1]$, define discounted prices

$$\widetilde{p}^{ij} := p^{ij} - \eta\,\pi^{ij}, \tag{71}$$

and define $\widetilde{p}^{\varnothing} := 0$.

Define the discounted menu as $\widetilde{\mathcal{M}}(\eta) := \{(n^{ij}, \widetilde{p}^{ij})\}_{i,j} \cup \{\varnothing\}$ and let $\pi\big(\widetilde{\mathcal{M}}(\eta)\big)$ denote the corresponding revenue to the DCP obtained by implementing it. The menu $\widetilde{\mathcal{M}}(\eta)$ is the posted menu obtained from the discrete solution by keeping ad intensities fixed and uniformly discounting each subscription fee by a fraction η of its corresponding revenue, together with the outside option $\varnothing$.

The following lemma characterizes how uniform revenue-based discounting controls deviations from the intended grid menu items and bounds the resulting revenue loss.

Lemma 2 **(Discounting and deviations)**. *Fix any choice of points $\{(\theta^{ij,\circ}, \alpha^{ij,\circ})\}_{i,j}$ such that $(\theta^{ij,\circ}, \alpha^{ij,\circ})$ lies in grid cell (i,j). Suppose a direct mechanism assigns to each $(\theta^{ij,\circ}, \alpha^{ij,\circ})$ the menu item (n^{ij}, p^{ij}) (or the outside option $\varnothing$) and satisfies incentive compatibility and individual rationality on the finite set $\{(\theta^{ij,\circ}, \alpha^{ij,\circ})\}_{i,j}$, i.e.,*

$$V(\theta^{ij,\circ}, \alpha^{ij,\circ}; n^{ij}, p^{ij}) \ge V(\theta^{ij,\circ}, \alpha^{ij,\circ}; n^{i'j'}, p^{i'j'}) \quad \forall (i,j), (i',j'), \tag{72}$$

and

$$V(\theta^{ij,\circ}, \alpha^{ij,\circ}; n^{ij}, p^{ij}) \ge 0 \quad \forall (i,j). \tag{73}$$

For each (i,j), define the revenue of menu item (i,j) evaluated at $\alpha^{ij,\circ}$ by

$$\pi^{ij} := p^{ij} + Trk\,n^{ij}\big(1 - \alpha^{ij,\circ} n^{ij}\big)_+, \qquad \pi^{\varnothing} := 0, \tag{74}$$

and let

$$\pi_\circ := \sum_{i=1}^{M_\theta} \sum_{j=1}^{M_\alpha} \rho^{ij}\,\pi^{ij}. \tag{75}$$

Then the expected revenue obtained when types best-respond to the discounted menu $\widetilde{\mathcal{M}}(\eta)$ satisfies

$$\pi_\circ - \pi\big(\widetilde{\mathcal{M}}(\eta)\big) \le \eta\,\overline{R} + \frac{2\Delta_V}{\eta} + \gamma_\alpha\,(\alpha_H - \alpha_L). \tag{76}$$

Proof: Fix a grid cell (i,j) and any type (θ, α) in that cell (this can be either the representative type (θ^i, α^j) or any true type in the cell). Since both (θ, α) and $(\theta^{ij,\circ}, \alpha^{ij,\circ})$ lie in the same cell, we have $|\theta - \theta^{ij,\circ}| \le \frac{\theta_{\mathrm{H}} - \theta_{\mathrm{L}}}{M_\theta}$ and $|\alpha - \alpha^{ij,\circ}| \le \frac{\alpha_{\mathrm{H}} - \alpha_{\mathrm{L}}}{M_\alpha}$, and therefore Step 1 implies that for any menu item $(n^{i'j'}, p^{i'j'})$,

$$\big|V(\theta, \alpha; n^{i'j'}, p^{i'j'}) - V(\theta^{ij,\circ}, \alpha^{ij,\circ}; n^{i'j'}, p^{i'j'})\big| \le \Delta_V. \tag{77}$$

Using incentive compatibility at $(\theta^{ij,\circ}, \alpha^{ij,\circ})$ and applying the above deviation bound twice yields

$$V(\theta, \alpha; n^{ij}, p^{ij}) \ge V(\theta, \alpha; n^{i'j'}, p^{i'j'}) - 2\Delta_V \qquad \forall (i',j'). \tag{78}$$

Under discounting, the utility from menu item (i',j') increases by exactly $\eta\pi^{i'j'}$, independently of (θ,α). Hence, if (θ,α) selects an option from $\widetilde{\mathcal{M}}(\eta)$ (using the tie-breaking rule in favor of higher π^{ij}), then for some (i',j'),

$$V(\theta,\alpha;n^{i'j'},p^{i'j'})+\eta\pi^{i'j'} \geq V(\theta,\alpha;n^{ij},p^{ij})+\eta\pi^{ij}. \tag{79}$$

Combining with (78) implies

$$\eta\big(\pi^{ij}-\pi^{i'j'}\big) \leq 2\Delta_V, \tag{80}$$

and therefore

$$\pi^{i'j'} \geq \pi^{ij} - \frac{2\Delta_V}{\eta}. \tag{81}$$

The realized revenue from selling menu item (i',j') to type (θ,α) is

$$\begin{aligned}&\widetilde{p}^{i'j'} + Trk\, n^{i'j'}\big(1-\alpha n^{i'j'}\big)_+ \\ &= (1-\eta)\pi^{i'j'} + \Big(Trk\, n^{i'j'}\big(1-\alpha n^{i'j'}\big)_+ - Trk\, n^{i'j'}\big(1-\alpha^{i'j',\circ} n^{i'j'}\big)_+\Big).\end{aligned} \tag{82}$$

By Lipschitz continuity of the ad-revenue term in α,

$$\big|Trk\, n^{i'j'}\big(1-\alpha n^{i'j'}\big)_+ - Trk\, n^{i'j'}\big(1-\alpha^{i'j',\circ} n^{i'j'}\big)_+\big| \leq \gamma_\alpha\,|\alpha-\alpha^{i'j',\circ}| \leq \gamma_\alpha\,(\alpha_{\mathrm{H}}-\alpha_{\mathrm{L}}). \tag{83}$$

Combining (82), (81), and (83) yields

$$(1-\eta)\pi^{i'j'} + \Big(Trk\, n^{i'j'}\big(1-\alpha n^{i'j'}\big)_+ - Trk\, n^{i'j'}\big(1-\alpha^{i'j',\circ} n^{i'j'}\big)_+\Big) \geq (1-\eta)\left(\pi^{ij}-\frac{2\Delta_V}{\eta}\right) - \gamma_\alpha\,(\alpha_{\mathrm{H}}-\alpha_{\mathrm{L}}). \tag{84}$$

Note that when the selected option is $\varnothing$, the inequality holds trivially since $\pi^{\varnothing}=0$ and the realized revenue is 0. Averaging (84) over types in each cell and then summing with weights ρ^{ij} implies

$$\pi\big(\widetilde{\mathcal{M}}(\eta)\big) \geq (1-\eta)\pi_\circ - \frac{2\Delta_V}{\eta} - \gamma_\alpha\,(\alpha_{\mathrm{H}}-\alpha_{\mathrm{L}}). \tag{85}$$

Rearranging gives

$$\pi_\circ - \pi\big(\widetilde{\mathcal{M}}(\eta)\big) \leq \eta\,\pi_\circ + \frac{2\Delta_V}{\eta} + \gamma_\alpha\,(\alpha_{\mathrm{H}}-\alpha_{\mathrm{L}}). \tag{86}$$

Finally, using $\pi_\circ \leq \overline{R}$ yields (76). ■

We now obtain an approximation bound relative to the optimal two-dimensional mechanism by applying Lemma 2 twice, following Madarász and Prat (2017).

Theorem 5 **(Two-dimensional approximation bound)**. *Let π^*_{2D} denote the optimal expected revenue in the two-dimensional problem, and let $\pi\big(\widetilde{\mathcal{M}}(\eta)\big)$ denote the expected revenue when types best-respond to the discounted posted menu $\widetilde{\mathcal{M}}(\eta)$. Choosing*

$$\eta^* := \min\left\{1, \sqrt{\frac{2\Delta_V}{\overline{R}}}\right\} \tag{87}$$

the discounted menu $\widetilde{\mathcal{M}}(\eta^*)$ satisfies:

$$\pi^*_{2D} - \pi\big(\widetilde{\mathcal{M}}(\eta^*)\big) \leq 2\left(\eta^*\,\overline{R} + \frac{2\Delta_V}{\eta^*} + \gamma_\alpha(\alpha_H-\alpha_L)\right). \tag{88}$$

Proof: The proof proceeds in two steps: we first bound $\pi^*_{\mathrm{2D}} - \pi^*_{\mathrm{Rep}}$ (Step A), and then bound $\pi^*_{\mathrm{Rep}} - \pi\big(\widetilde{\mathcal{M}}(\eta)\big)$ (Step B).

Step A (Bounding $\pi^*_{\mathrm{2D}} - \pi^*_{\mathrm{Rep}}$): Let μ^* be an optimal solution to Problem ($\mathcal{P}$-2D), and let $\pi^*(\theta,\alpha)$ denote the revenue collected from type (θ,α) under μ^*. For each grid cell (i,j) with $\rho^{ij} > 0$, choose a point $(\theta^{ij,\circ}, \alpha^{ij,\circ})$ in that cell such that

$$\pi^*(\theta^{ij,\circ}, \alpha^{ij,\circ}) \geq \frac{1}{\rho^{ij}} \int_{\mathrm{cell}\ (i,j)} \pi^*(\theta,\alpha)\, h(\theta,\alpha)\, d\theta\, d\alpha. \tag{89}$$

Such a point exists because the cell average is bounded above by the cell supremum and $\pi^*(\theta,\alpha)$ is bounded above by $\overline{R}$.

Define a discrete environment that places probability mass ρ^{ij} on each selected point $(\theta^{ij,\circ}, \alpha^{ij,\circ})$. Restricting μ^* to this finite type set yields a finite direct mechanism that is incentive compatible and individually rational on $\{(\theta^{ij,\circ}, \alpha^{ij,\circ})\}_{i,j}$. For this restricted mechanism, define the reference revenues (at the selected points) by

$$\pi^{ij}_\circ := p_{\mu^*}(\theta^{ij,\circ}, \alpha^{ij,\circ}) + Trk\, n_{\mu^*}(\theta^{ij,\circ}, \alpha^{ij,\circ}) \Big(1 - \alpha^{ij,\circ}\, n_{\mu^*}(\theta^{ij,\circ}, \alpha^{ij,\circ})\Big)_+, \tag{90}$$

and let $\pi_\circ := \sum_{i,j} \rho^{ij} \pi^{ij}_\circ$. By (89) and summing over cells, we have $\pi_\circ \geq \pi^*_{\mathrm{2D}}$.

Now apply Lemma 2 to this restricted mechanism (model points $\{(\theta^{ij,\circ}, \alpha^{ij,\circ})\}$ with masses ρ^{ij}), and evaluate the resulting discounted menu in the representative environment (grid points (θ^i, α^j) with the same cell masses ρ^{ij}). Because the bound in Lemma 2 is derived type-by-type for every (θ,α) within each cell (i,j), it remains valid after taking expectations under any distribution that assigns mass ρ^{ij} to cell (i,j) and has support contained in that cell. In particular, it applies when evaluating under the representative grid distribution (placing mass ρ^{ij} on (θ^i, α^j)) and under the true continuous distribution $h(\theta,\alpha)$. This yields a discounted posted menu whose representative-environment revenue satisfies

$$\pi^*_{\mathrm{2D}} - \pi^*_{\mathrm{Rep}} \leq \eta\, \overline{R} + \frac{2\Delta_V}{\eta} + \gamma_\alpha\, (\alpha_{\mathrm{H}} - \alpha_{\mathrm{L}}), \tag{91}$$

where we used that the representative problem optimum π^*_{Rep} dominates the representative-environment revenue of any posted menu.

Step B (Bounding $\pi^*_{\mathrm{Rep}} - \pi\big(\widetilde{\mathcal{M}}(\eta)\big)$): Let $\{(n^{ij}_*, p^{ij}_*)\}_{i,j}$ be an optimal solution to Problem ($\mathcal{P}$-Rep) (including the outside option). Apply Lemma 2 again with model points given by the representative grid $\{(\theta^i, \alpha^j)\}_{i,j}$ (i.e., set $(\theta^{ij,\circ}, \alpha^{ij,\circ}) = (\theta^i, \alpha^j)$), and evaluate in the true continuous environment $h(\theta,\alpha)$. This yields that the discounted menu $\widetilde{\mathcal{M}}(\eta)$ constructed from the representative optimum satisfies

$$\pi^*_{\mathrm{Rep}} - \pi\big(\widetilde{\mathcal{M}}(\eta)\big) \leq \eta\, \overline{R} + \frac{2\Delta_V}{\eta} + \gamma_\alpha\, (\alpha_{\mathrm{H}} - \alpha_{\mathrm{L}}). \tag{92}$$

Combining (91) and (92) gives

$$\pi^*_{\mathrm{2D}} - \pi\big(\widetilde{\mathcal{M}}(\eta)\big) \leq 2\left(\eta\, \overline{R} + \frac{2\Delta_V}{\eta} + \gamma_\alpha\, (\alpha_{\mathrm{H}} - \alpha_{\mathrm{L}})\right), \tag{93}$$

which proves the first claim. For the second claim, substituting $\eta = \eta^*$ yields (88). ■

We now prove the corollary on menus of limited size.

A menu of size (at most) L corresponds to discretizing the type space into $M_\theta \times M_\alpha = L$ cells and offering one option per representative type, using the construction in Steps 1–3 of Section H. For a given grid (M_θ, M_α), recall that the Lipschitz argument in Step 1 yields the uniform discretization error

$$\Delta_V := \kappa_\theta \cdot \frac{\theta_{\rm H} - \theta_{\rm L}}{M_\theta} + \kappa_\alpha \cdot \frac{\alpha_{\rm H} - \alpha_{\rm L}}{M_\alpha}, \tag{94}$$

where $\kappa_\theta = \frac{1}{2}Tk$ and $\kappa_\alpha = Tk\,\theta_{\rm H}\,\bar{n}$.

We next choose (M_θ, M_α) (i.e., how to allocate the L menu slots across the two dimensions) to minimize Δ_V subject to $M_\theta M_\alpha = L$:

$$\min_{\substack{M_\theta, M_\alpha \geq 1 \\ M_\theta M_\alpha = L}} \left\{ \kappa_\theta \cdot \frac{\theta_{\rm H} - \theta_{\rm L}}{M_\theta} + \kappa_\alpha \cdot \frac{\alpha_{\rm H} - \alpha_{\rm L}}{M_\alpha} \right\}. \tag{95}$$

Treating (M_θ, M_α) as continuous in (95), the first-order condition yields

$$M_\theta^* = \sqrt{L}\sqrt{\frac{\kappa_\theta(\theta_{\rm H} - \theta_{\rm L})}{\kappa_\alpha(\alpha_{\rm H} - \alpha_{\rm L})}}, \qquad M_\alpha^* = \sqrt{L}\sqrt{\frac{\kappa_\alpha(\alpha_{\rm H} - \alpha_{\rm L})}{\kappa_\theta(\theta_{\rm H} - \theta_{\rm L})}}. \tag{96}$$

Substituting (96) into (94) gives the minimized (continuous) discretization error

$$\Delta_V^*(L) = 2\sqrt{\frac{\kappa_\theta \kappa_\alpha (\theta_{\rm H} - \theta_{\rm L})(\alpha_{\rm H} - \alpha_{\rm L})}{L}}. \tag{97}$$

Let $\overline{R} := \frac{1}{2}Tk\,\theta_{\rm H} + Trk\,\bar{n}$ and $\gamma_\alpha := Trk\,\bar{n}^2$. Given this choice of grid, Theorem 5 implies that for any $\eta \in (0,1]$,

$$\pi_{\rm 2D}^* - \pi\big(\widetilde{\mathcal{M}}_L(\eta)\big) \leq 2\eta\,\overline{R} + \frac{4\,\Delta_V^*(L)}{\eta} + 2\gamma_\alpha(\alpha_{\rm H} - \alpha_{\rm L}), \tag{98}$$

where $\widetilde{\mathcal{M}}_L(\eta)$ denotes the discounted menu constructed from the optimal representative mechanism on an $M_\theta \times M_\alpha$ grid with $M_\theta M_\alpha = L$ (Steps 1–3 above). Optimizing the right-hand side of (98) over η yields

$$\eta^* := \min\left\{1, \sqrt{\frac{2\,\Delta_V^*(L)}{\overline{R}}}\right\}. \tag{99}$$

Substituting (99) into (98) gives the desired bound.

$$\pi_{\rm 2D}^* - \pi\big(\widetilde{\mathcal{M}}_L(\eta_L^*)\big) \leq 2\left(\eta_L^*\,\overline{R} + \frac{2\Delta_V^*(L)}{\eta_L^*} + \gamma_\alpha(\alpha_{\rm H} - \alpha_{\rm L})\right). \tag{100}$$

■

Appendix I: Unary Ad-Supported (UAS) Mechanism

We organize this section as follows. We first characterize the parameter regions in which UAS* weakly revenue-dominates ADS* and SUB* (Section I.1). We then compare UAS* with GAS* through sensitivity analyses (Section I.2). Finally, we study the revenue gaps between GAS* and UAS* (Section I.3) and between BAS* and UAS* (Section I.4).

I.1. UAS* relative to SUB* and ADS*

Figure 22 shows how the expected revenue ranking of ADS*, SUB*, and UAS* changes with the ad-revenue rate r and ad-nuisance α. First, note that UAS* is a weakly revenue-dominant mechanism in that it always yields revenue greater than or equal to the maximum of the revenue under ADS* and SUB* mechanisms. When the ad-revenue rate is high and ads impose little nuisance (high r, low α), the ad-supported mechanism ADS* is attractive. Thus, UAS* reduces to ADS* (yellow region, $\pi_{\mathrm{UAS}^*} = \pi_{\mathrm{ADS}^*}$). When ads pay poorly and cause high nuisance (low r and high α), the subscription-only mechanism performs better. Thus, UAS* reduces to SUB* (purple region, $\pi_{\mathrm{UAS}^*} = \pi_{\mathrm{SUB}^*}$). Between these two extremes, charging a positive fee while keeping a moderate ad load yields higher revenue than both ADS* and SUB* mechanisms (green region, $\pi_{\mathrm{UAS}^*} > \max\{\pi_{\mathrm{ADS}^*}, \pi_{\mathrm{SUB}^*}\}$). Thus, in this region, UAS* extracts revenue from both sources (ads and subscription) without relying too heavily on either one.

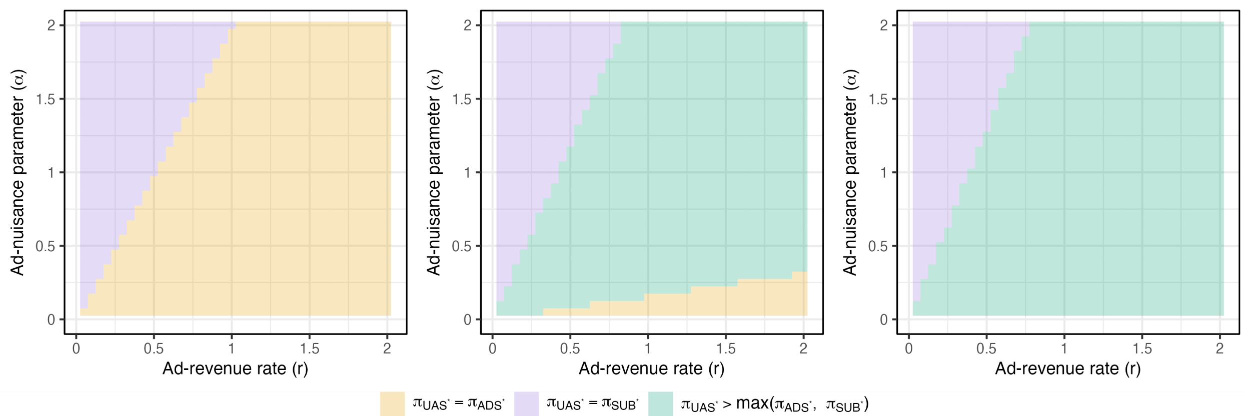


Figure 22 **Revenue dominance regions over the (r, α) parameter space for ADS*, SUB*, and UAS* mechanisms. Yellow indicates $\pi_{\mathrm{UAS}^*} = \pi_{\mathrm{ADS}^*}$ (i.e., $p^*_{\mathrm{UAS}} = 0$), purple indicates $\pi_{\mathrm{UAS}^*} = \pi_{\mathrm{SUB}^*}$ (i.e., $n^*_{\mathrm{UAS}} = 0$), and green indicates $\pi_{\mathrm{UAS}^*} > \max\{\pi_{\mathrm{ADS}^*}, \pi_{\mathrm{SUB}^*}\}$. The left panel assumes $\mathrm{Beta}(1,1)$, the middle panel assumes $\mathrm{Beta}(2,2)$, and the right panel assumes $\mathrm{Beta}(6,6)$; all other parameters are set to the nominal values in Table 1.**

I.2. Sensitivity Analysis of UAS*

Although UAS* weakly dominates ADS* and SUB*, it remains a more constrained mechanism than GAS*. We note that the UAS mechanism offers a single ad-intensity–subscription-fee pair $(n_{\mathrm{UAS}}, p_{\mathrm{UAS}})$ uniformly to all users. In contrast, the GAS mechanism offers a type-dependent menu, and can therefore better price-discriminate across users. We compare the UAS* mechanism with the GAS* mechanism in two steps: First, we conduct a sensitivity analysis of the revenues under the two mechanisms with respect to key model parameters. Then, we examine the joint effect of ad-attractiveness (r/α) and the composition of user-types (a/b) on the revenue gap between the two mechanisms. The nominal values of the parameters are as described in Table 1.

Sensitivity to Ad-Nuisance and Ad-Revenue Rates: Figure 23 illustrates how the expected revenue under GAS* and UAS* varies with the ad-nuisance parameter α (left panel) and the ad-revenue rate r (right panel). In both panels, the revenue gap is largest at intermediate values of α and r. When α is low (in the left panel) or r is high (in the right panel), both mechanisms rely more on advertising; when α is high or r is low, both mechanisms rely more on subscriptions. The revenue advantage of GAS* over UAS* is therefore greatest for intermediate values of α and r, where tailoring ad-intensities and subscription fees across user types yields the largest gains relative to UAS*.

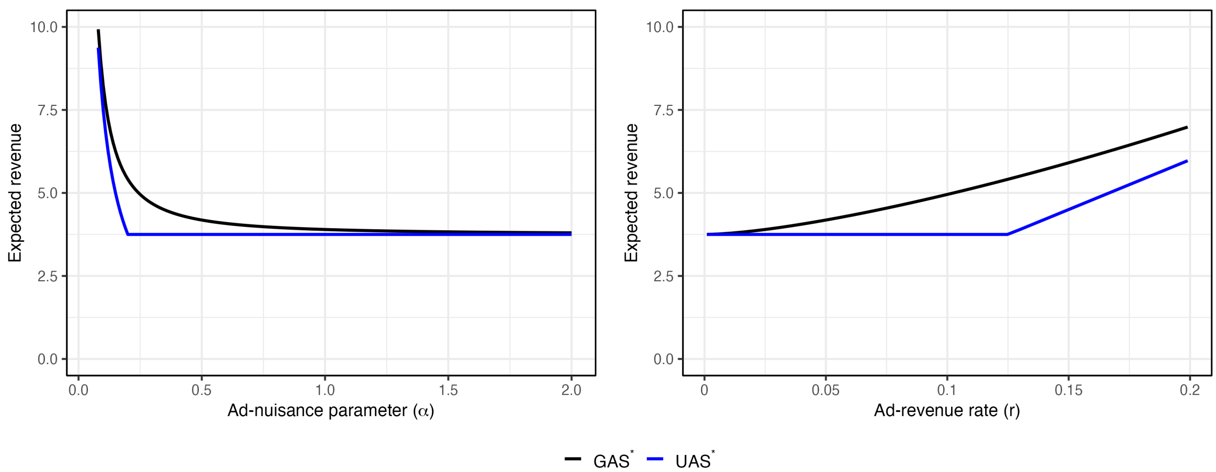


Figure 23 **Expected revenue under UAS* and GAS* mechanisms as a function of ad-nuisance parameter α (left) and ad-revenue rate r (right). Both panels use nominal and range values as described in Table 1.**

Sensitivity to the Distribution of User Types: Figure 24 shows how expected revenue under GAS* and UAS* varies with the shape of the user-type distribution. In the symmetric case $\theta \sim \mathrm{Beta}(m, m)$, revenues under both mechanisms increase modestly as the distribution becomes more concentrated around moderate types. In the asymmetric case $\theta \sim \mathrm{Beta}(a, 6)$, increasing a shifts mass from low types toward moderate and high types, which makes the subscription-fee lever more attractive for the DCP. In this case, the revenue gap widens since GAS* can better price-discriminate across user types, whereas UAS* is constrained to set a single ad-intensity–fee pair for all users.

Sensitivity to Baseline Usage: Figure 25 shows how expected revenue under GAS* and UAS* varies with baseline usage k for low and high values of ad-nuisance α. In both panels, revenues increase with k because higher baseline usage raises total content consumption and gives the DCP more usage to monetize through ads and subscription fees. The revenue gap also increases with k, since GAS* can tailor across user types while UAS* remains constrained to a single ad-intensity–fee pair.

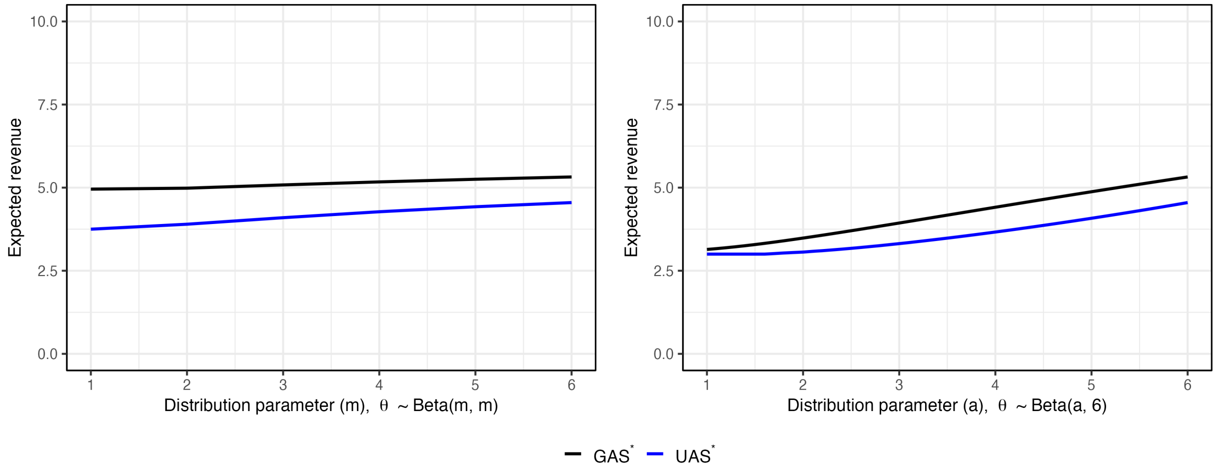


Figure 24 **Expected revenue as a function of the user-type distribution. Both panels are produced using nominal and range values from Table 1.**

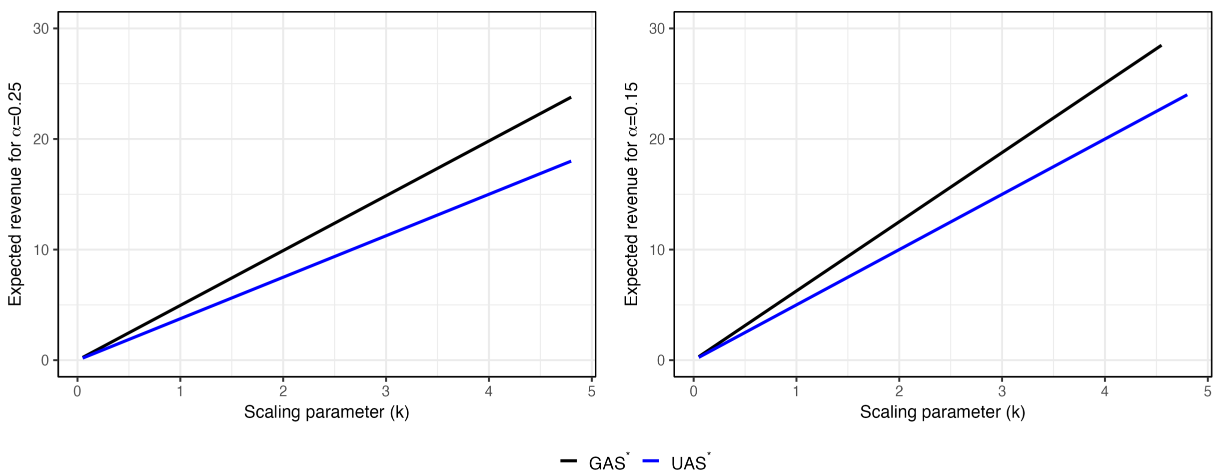


Figure 25 **Expected revenue as a function of the baseline usage (k). Both panels are produced using nominal and range values from Table 1.**

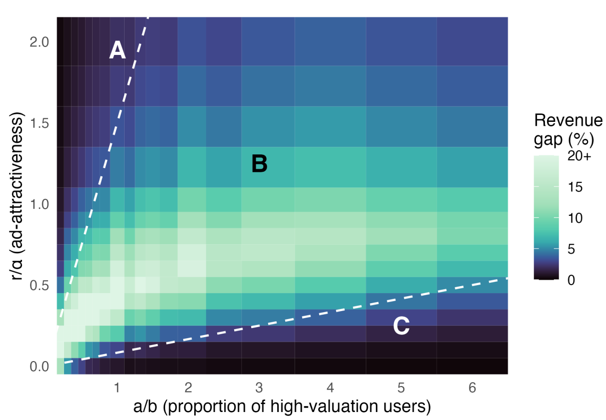
I.3. Numerical Study of the Revenue Gap Between GAS* and UAS*

Similar to Appendix E, we plot the heatmap of the revenue gap between GAS* and UAS* across the parameter grid of Table 1. We define the revenue gap (in %) as $(\pi_{\mathrm{GAS}^*} - \pi_{\mathrm{UAS}^*})/\pi_{\mathrm{UAS}^*} \times 100$, and average it over the remaining parameter k for each pair $(r/\alpha, a/b)$. Figure 26 displays the resulting heatmap, with the three regions delineated by the dashed boundaries and summarized in the adjacent table.

- **Region A** ($r/\alpha > 3a/2b$; upper-left portion of the figure): When ad-attractiveness is high but the share of high-type users is small, the gap is modest (ranging from $\approx 0\%$ to 9%). In this region, ad revenue dominates the ad-nuisance cost, so UAS* sets the ad-intensity close to its maximum. The UAS* mechanism extracts most of the available ad revenue, and the gap between the two mechanisms remains limited.

- **Region B** ($a/12b < r/\alpha \le 3a/2b$; the central band of the figure): This is the region where GAS* achieves its strongest advantage over UAS*, with the revenue gap ranging from $\approx 2\%$ to 36%. In this region, ads are attractive enough to serve as an effective monetization lever, but not so lucrative that a single uniform intensity captures most of the available surplus. Therefore, the type-dependent offering by GAS* yields a substantial revenue advantage over the UAS* mechanism.

- **Region C** ($r/\alpha \le a/12b$; lower-right portion of the figure): When ad attractiveness is low relative to the user-type composition, the gap collapses, ranging from $\approx 0\%$ to 13%. In this region, ads are not attractive (either too low-paying or too nuisance-heavy), so neither mechanism generates substantial ad revenue. As a result, the GAS* mechanism offers little incremental value over UAS*.



Region A ($\frac{r}{\alpha} > \frac{3a}{2b}$; gap ≈ 0–9%): Ad attractiveness is high enough that both mechanisms' ad intensity is already close to the maximum. This limits the scope for GAS* to extract additional revenue.

Region B ($\frac{a}{12b} < \frac{r}{\alpha} \le \frac{3a}{2b}$; gap ≈ 2–36%): Ads are effective but not dominant, and heterogeneity across user types allows GAS* to tailor ad intensities to each type, while UAS* is restricted to a single uniform ad-intensity.

Region C ($\frac{r}{\alpha} \le \frac{a}{12b}$; gap ≈ 0–13%): Ads are low-paying or impose substantial nuisance; neither mechanism generates substantial ad revenue. Thus, GAS*'s gain over UAS* is limited.

Figure 26 The left panel displays the revenue gap in % as a joint function of ad attractiveness (r/α) and the user-type distribution (a/b). Each cell reports the average gap computed over the remaining parameter k. Lighter colors indicate a larger revenue gap. The right panel describes the three regions delineated by the dashed boundaries in the figure.

Comparing the GAS*–UAS* heatmap in Figure 26 with the GAS*–BAS* heatmap in Figure 21 yields two key differences. First, Region B is markedly wider in the GAS*–UAS* heatmap. This is because BAS* screens users through a two-option menu while UAS* offers only a single (n, p) pair. Therefore, GAS* retains a sizable advantage over UAS* across a much broader range of $(r/\alpha, a/b)$. Second, while the revenue gap is largest in Region B in both figures, the location of the largest gap within the region is different. The GAS*–BAS* gap peaks in the upper-right of the region, whereas the GAS*–UAS* gap peaks in the lower-left of the region. To better understand this shift, we next plot the heatmap of the revenue gap between BAS* and UAS* over the same $(r/\alpha,\ a/b)$ space.

I.4. Numerical Study of the Revenue Gap Between BAS* and UAS*

As in the previous subsection, we define the revenue gap between BAS* and UAS* (in %) as $(\pi_{\text{BAS}^*} - \pi_{\text{UAS}^*})/\pi_{\text{UAS}^*} \times 100$ and average it over the remaining parameter k for each pair $(r/\alpha, a/b)$. Here, a positive revenue gap indicates that BAS* outperforms UAS*, whereas a negative gap indicates that UAS* outperforms BAS*. Figure 27 plots the proportion of high-type users a/b on the horizontal axis and ad-attractiveness r/α on the vertical axis. The dashed boundaries at $r/\alpha = 3/4$ and $a/b = 1$ partition the grid into four regions.

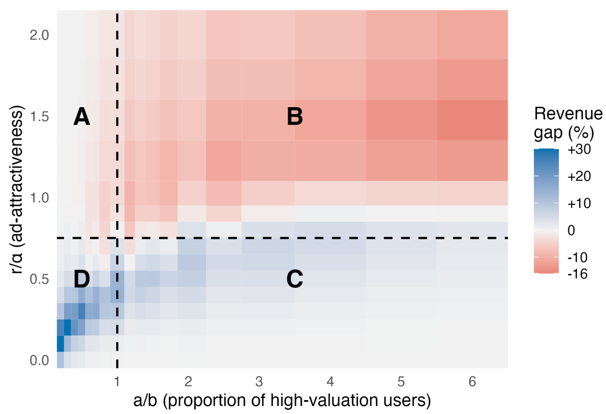
- **Region A** ($r/\alpha \geq 3/4$; $a/b < 1$; upper-left portion of the figure): In this region, ad attractiveness is high relative to the proportion of high-type users, so both mechanisms monetize mainly through advertising and yield comparable revenue. Here, the gap ranges from $\approx -5\%$ to 1%.

- **Region B** ($r/\alpha \geq 3/4$; $a/b \geq 1$; upper-right portion of the figure): In this region, UAS* attains its largest revenue advantage over BAS*, with the gap ranging from $\approx -16\%$ to 5%. Under BAS*, high-type users self-select into the ad-free subscription tier, so the DCP earns no ad revenue from them. Since this region is characterized by a high proportion of high-type users and high ad attractiveness, this forgone ad revenue is substantial and is precisely the source of UAS*'s advantage over BAS*. This is because UAS* levies both ads and a subscription fee on every participating user and thus monetizes the high-type users on both margins. Thus, in this region, the benefit from leveraging both the ad and subscription levers under UAS* outweighs the screening benefit from BAS*.

- **Region C** ($r/\alpha < 3/4$; $a/b \geq 1$; lower-right portion of the figure): In this region, ads either yield limited returns or impose substantial nuisance costs on users. As a result, neither mechanism generates substantial ad revenue, and both BAS* and UAS* primarily extract revenue through subscriptions. The revenue gap ranges from $\approx -3\%$ to 14%, and BAS*'s menu yields only a modest edge over the UAS* mechanism.

- **Region D** ($r/\alpha < 3/4$; $a/b < 1$; lower-left portion of the figure): In this region, BAS* attains its largest revenue advantage over UAS*, with the gap ranging from $\approx -3\%$ to 30%. Both ad attractiveness and the proportion of high-type users are low, so the market is dominated by low-type users and ads are not lucrative enough to monetize them. Since ads are weak, UAS* relies on its single subscription fee. This fee excludes many of these low-type users. On the other hand, BAS* instead offers them the free ad-supported tier and serves the few high types through the ad-free subscription tier. Thus, the two-option BAS* mechanism better screens the users and outperforms UAS* in this region.

In summary, neither BAS* nor UAS* uniformly revenue-dominates the other. Their ranking depends on ad-attractiveness r/α and the user-type composition a/b. UAS* is preferred when ads are attractive, and the high-type users predominate. In this case, it leverages both the ad and subscription levers, whereas BAS* forgoes ad revenue from users who self-select into its ad-free tier. On the other hand, BAS* is preferred when ads are not very attractive and low-type users predominate. Here, its free ad-supported tier serves the low-type users that the single UAS* fee excludes.



Region A ($r/\alpha \geq \frac{3}{4}$; $a/b < 1$; gap ≈ -5 to 1%): Ad attractiveness is high relative to the share of high-type users; both mechanisms monetize mainly through ads and earn comparable revenue.

Region B ($r/\alpha \geq \frac{3}{4}$; $a/b \geq 1$; gap ≈ -16 to 5%): UAS* attains its largest advantage: it levies both ads and a subscription fee on the many high-type users, whereas BAS* earns no ad revenue from them.

Region C ($r/\alpha < \frac{3}{4}$; $a/b \geq 1$; gap ≈ -3 to 14%): Ads are weak, so both mechanisms rely on subscriptions; BAS*'s menu yields a modest edge.

Region D ($r/\alpha < \frac{3}{4}$; $a/b < 1$; gap ≈ -3 to 30%): Low-type users predominate and ads are weak; BAS* attains its largest advantage by serving these users through its free ad-supported tier, whereas the single UAS* fee excludes them.

Figure 27 The left panel displays the revenue gap (in %) as a joint function of ad attractiveness (r/α) and the user-type distribution (a/b). Each cell reports the average gap computed over the remaining parameter k. The right panel describes the four regions delineated by the dashed boundaries in the figure.

Appendix J: GAS Mechanism under Competition

In this section, we model competition by introducing an outside option that provides users with a reservation utility $V^o \geq 0$ and obtain an optimal GAS mechanism under this setting. The derivation follows from the proof of Theorem 1 in Appendix A. Since most steps remain unchanged, we only highlight the key points where the presence of an outside option alters the analysis.

Recall that under a direct mechanism $\text{GAS} = (n_{\text{GAS}}(\cdot), p_{\text{GAS}}(\cdot), d_{\text{GAS}}(\cdot))$, the net utility to a user of type θ under an incentive-compatible mechanism is

$$V_{\text{GAS}}(\theta) := V_{\text{GAS}}(\theta;\theta) = \frac{1}{2}T\theta k\big((1-\alpha n_{\text{GAS}}(\theta))_+\big)^2 - p_{\text{GAS}}(\theta). \tag{101}$$

With an outside option, the participation condition becomes:

$$d_{\text{GAS}}(\theta) = \begin{cases} 1 & \text{if } V_{\text{GAS}}(\theta) \geq V^o, \\ 0 & \text{otherwise,} \end{cases} \tag{102}$$

instead of the baseline threshold 0.

The incentive-compatibility argument in Appendix A implies the envelope condition remains unchanged:

$$\frac{dV_{\text{GAS}}(\theta)}{d\theta} = \frac{Tk}{2}\left((1-\alpha\cdot n_{\text{GAS}}(\theta))_+\right)^2. \tag{103}$$

Let $\underline{\theta}^o \in [\theta_{\text{L}}, \theta_{\text{H}}]$ denote the lowest participating type under the optimal mechanism, i.e., $d_{\text{GAS}}(\theta) = 0$ for $\theta < \underline{\theta}^o$ and $d_{\text{GAS}}(\theta) = 1$ for $\theta \geq \underline{\theta}^o$. A revenue-maximizing mechanism must satisfy the binding condition

$$V_{\text{GAS}}(\underline{\theta}^o) = V^o, \tag{104}$$

since otherwise the DCP can increase payments for all participating type while preserving IC until (104) holds. Integrating (103) from $\underline{\theta}^o$ yields, for all $\theta \geq \underline{\theta}^o$,

$$V_{\mathrm{GAS}}(\theta) = V^{\mathrm{o}} + \int_{\underline{\theta}^o}^{\theta} \frac{Tk}{2} \left((1-\alpha \cdot n_{\mathrm{GAS}}(x))_+\right)^2 dx. \tag{105}$$

Thus, for all $\theta \geq \underline{\theta}^o$, the payment function becomes

$$p_{\mathrm{GAS}}(\theta) = \frac{\theta Tk}{2} \left((1-\alpha n_{\mathrm{GAS}}(\theta))_+\right)^2 - \int_{\underline{\theta}^o}^{\theta} \frac{Tk}{2} \left((1-\alpha \cdot n_{\mathrm{GAS}}(x))_+\right)^2 dx - V^{\mathrm{o}}. \tag{106}$$

Substituting (106) into the expected revenue expression and repeating the same change-of-order argument as in Appendix A, we obtain

$$\begin{aligned}
&\int_{\theta_{\mathrm{L}}}^{\theta_{\mathrm{H}}} \left(p_{\mathrm{GAS}}(\theta) + Trk\, n_{\mathrm{GAS}}(\theta)\left((1-\alpha n_{\mathrm{GAS}}(\theta))_+\right)\right) d_{\mathrm{GAS}}(\theta) f(\theta) d\theta \\
&= \int_{\theta_{\mathrm{L}}}^{\theta_{\mathrm{H}}} \left(\frac{Tk\Psi_{\mathrm{GAS}}(\theta)}{2}\left((1-\alpha n_{\mathrm{GAS}}(\theta))_+\right)^2 + Trk\, n_{\mathrm{GAS}}(\theta)\left((1-\alpha n_{\mathrm{GAS}}(\theta))_+\right)\right) h_{\mathrm{GAS}}(\theta) d\theta \;-\; V^{\mathrm{o}} \int_{\theta_{\mathrm{L}}}^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(\theta) d\theta,
\end{aligned} \tag{107}$$

where $h_{\mathrm{GAS}}(\theta) := d_{\mathrm{GAS}}(\theta) f(\theta)$ and

$$\Psi_{\mathrm{GAS}}(\theta) := \begin{cases} \theta - \dfrac{\int_{\theta}^{\theta_{\mathrm{H}}} h_{\mathrm{GAS}}(x) dx}{h_{\mathrm{GAS}}(\theta)} & \text{if } h_{\mathrm{GAS}}(\theta) > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{108}$$

Relative to Appendix A, the only change is the additional term $-V^{\mathrm{o}} \int h_{\mathrm{GAS}}(\theta) d\theta$.

Next, fix $d_{\mathrm{GAS}}(\cdot)$. Since the outside-option term in (107) does not depend on $n_{\mathrm{GAS}}(\cdot)$, the point-wise optimization over $n_{\mathrm{GAS}}(\theta)$ is identical to Appendix A. Hence the optimal ad intensity is still given by (37):

$$n^*_{\mathrm{GAS}}(\theta) = \begin{cases} \dfrac{r - \alpha \Psi_{\mathrm{GAS}}(\theta)}{\alpha\,(2r - \alpha \Psi_{\mathrm{GAS}}(\theta))} & \text{if } \Psi_{\mathrm{GAS}}(\theta) < \frac{r}{\alpha}, \\ 0 & \text{if } \Psi_{\mathrm{GAS}}(\theta) \geq \frac{r}{\alpha}. \end{cases} \tag{109}$$

Define the maximized per-user contribution (excluding the density term) as a function of a generic virtual value z:

$$\bar{G}(z) := \max_{0 \leq n \leq 1/\alpha} \left\{ \frac{Tkz}{2}(1-\alpha n)^2 + Trk\, n(1-\alpha n) \right\} = \begin{cases} \dfrac{Tk\, r^2}{2\alpha(2r - \alpha z)} & \text{if } z < \dfrac{r}{\alpha}, \\ \dfrac{Tk\, z}{2} & \text{if } z \geq \dfrac{r}{\alpha}. \end{cases} \tag{110}$$

Using (107) and the fact that $\bar{G}(\cdot)$ is increasing in z, under the monotone hazard rate assumption (so $\psi(\theta)$ is increasing), the optimal participation rule is a cutoff:

$$d^*_{\mathrm{GAS}}(\theta) = \mathbf{1}\{\theta \geq \underline{\theta}^o\}. \tag{111}$$

Moreover, for all $\theta \geq \underline{\theta}^o$ we have $d^*_{\mathrm{GAS}}(x) = 1$ for all $x \in [\theta, \theta_{\mathrm{H}}]$, and hence $\Psi_{\mathrm{GAS}}(\theta) = \psi(\theta)$ on the participating region. The cutoff $\underline{\theta}^o$ is characterized by the indifference condition

$$\bar{G}(\psi(\underline{\theta}^o)) = V^{\mathrm{o}}, \tag{112}$$

with the understanding that if $V^{\mathrm{o}} \le \bar{G}(\psi(\theta_{\mathrm{L}}))$ then $\underline{\theta}^o = \theta_{\mathrm{L}}$ (full coverage), and if $V^{\mathrm{o}} > \bar{G}(\psi(\theta_{\mathrm{H}}))$ then no type participates. Equivalently, letting $z^o := \bar{G}^{-1}(V^{\mathrm{o}})$, the cutoff type $\underline{\theta}^o$ is given by

$$z^o = \begin{cases} \dfrac{2r}{\alpha} - \dfrac{Tk\,r^2}{2\alpha^2 V^{\mathrm{o}}} & \text{if } 0 < V^{\mathrm{o}} < \dfrac{Tk\,r}{2\alpha}, \\ \dfrac{2V^{\mathrm{o}}}{Tk} & \text{if } V^{\mathrm{o}} \ge \dfrac{Tk\,r}{2\alpha}, \end{cases} \qquad \underline{\theta}^o = \max\left\{\theta_{\mathrm{L}},\, \psi^{-1}(z^o)\right\}. \tag{113}$$

Thus, the optimal GAS mechanism under competition—where users have an outside option that provides reservation utility V^{o}—is characterized by the cutoff $\underline{\theta}^o$ defined above and is given by:

$$n^{o*}_{\mathrm{GAS}}(\theta) = \begin{cases} \dfrac{r - \alpha\psi(\theta)}{\alpha\,(2r - \alpha\psi(\theta))} & \text{if } \underline{\theta}^o \le \theta < \psi^{-1}\left(\dfrac{r}{\alpha}\right), \\ 0 & \text{if } \theta \ge \max\left\{\underline{\theta}^o,\, \psi^{-1}\left(\dfrac{r}{\alpha}\right)\right\}, \end{cases}$$

$$p^{o*}_{\mathrm{GAS}}(\theta) = \frac{\theta Tk\,(1 - \alpha n^{o*}_{\mathrm{GAS}}(\theta))^2}{2} - \int_{\underline{\theta}^o}^{\theta} \frac{Tk\,(1 - \alpha n^{o*}_{\mathrm{GAS}}(x))^2}{2} dx - V^{\mathrm{o}}, \quad \forall \theta \in [\underline{\theta}^o, \theta_{\mathrm{H}}],$$

$$d^{o*}_{\mathrm{GAS}}(\theta) = \begin{cases} 0 & \text{if } \theta < \underline{\theta}^o, \\ 1 & \text{if } \theta \in [\underline{\theta}^o, \theta_{\mathrm{H}}]. \end{cases}$$

This completes the derivation. ■

Appendix K: Robustness to Estimation Error in Ad-Nuisance

Since ad-nuisance α is an important parameter of our model, a natural question is how robust the revenue of the GAS* mechanism is to estimation error in α. In this section, we discuss the revenue guarantee when the DCP deploys a GAS* mechanism computed under an estimated ad-nuisance parameter $\hat{\alpha}$ that differs from the true value α. The analytical structure underlying this guarantee draws on the Lipschitz continuity properties established in Appendix H.

Consider the one-dimensional setting in which ad-nuisance (α) is a global parameter. Suppose the DCP estimates $\hat{\alpha}$ and computes the optimal GAS* mechanism under $\hat{\alpha}$, i.e., the ad-intensity function $n^*_{\mathrm{GAS}}(\theta;\hat{\alpha})$ and the subscription-fee function $p^*_{\mathrm{GAS}}(\theta;\hat{\alpha})$ as characterized in Theorem 1. This mechanism is then deployed in a setting where the true ad-nuisance is α. In such a setting, there are two sources of revenue distortion: (i) users' actual usage times differ from those anticipated under $\hat{\alpha}$, because usage depends on the true ad-nuisance through $t^*(\theta) = k(1 - \alpha\, n(\theta))_+$, and (ii) ad revenue per user changes because the effective ad impressions depend on α rather than $\hat{\alpha}$.

Recall that the user's net utility function satisfies

$$\left|V_\mu(\hat{\theta}, \hat{\alpha}; \theta, \alpha) - V_\mu(\hat{\theta}, \hat{\alpha}; \theta, \alpha')\right| \le \kappa_\alpha\, |\alpha - \alpha'|, \tag{114}$$

where $\kappa_\alpha := Tk\,\theta_{\mathrm{H}}\,\bar{n}$, and the per-user ad-revenue term satisfies

$$\left|Trk\,n(1 - \alpha\, n)_+ - Trk\,n(1 - \alpha'\, n)_+\right| \le \gamma_\alpha\, |\alpha - \alpha'|, \tag{115}$$

where $\gamma_\alpha := Trk\,\bar{n}^2$. These constants quantify the sensitivity of utility and ad-revenue to the true value of α.

Let $\pi(\hat{\alpha};\alpha)$ denote the expected revenue when the DCP deploys the mechanism optimized under $\hat{\alpha}$ but the true ad-nuisance is α, and let $\pi(\alpha;\alpha)$ denote the expected revenue under the correctly specified mechanism. The Lipschitz constants κ_α and γ_α imply that if $|\hat{\alpha}-\alpha|\leq\epsilon$, then the per-user utility distortion is bounded by $\kappa_\alpha\,\epsilon$ and the per-user ad-revenue distortion is bounded by $\gamma_\alpha\,\epsilon$. Aggregating across the population of users, and following the discretize-and-discount steps from Appendix H, we obtain that for any $\hat{\alpha}$ with $|\hat{\alpha}-\alpha|\leq\epsilon$,

$$\pi(\alpha;\alpha)-\pi(\hat{\alpha};\alpha)\leq C\cdot\epsilon, \tag{116}$$

where C is a constant that depends on the model primitives (T, k, r, θ_{H}, $\bar{n}$) but not on the specific value of α. This bound formalizes the intuition that if the ad-nuisance estimate is close to the true value, the revenue loss from misspecification is proportionally small.

The bound in (116) is consistent with the numerical evidence from the sensitivity analysis of the GAS* mechanism (Figure 11, left panel, in Appendix C.1), where the revenue under GAS* varies smoothly and gradually with ad-nuisance α over a wide range of values. This flatness reflects the mechanism's ability to balance its two revenue levers – advertising and subscription fees – in response to changes in ad-nuisance.

In summary, the Lipschitz continuity properties established in our two-dimensional analysis provide a foundation for bounding the revenue loss from estimation error in α. Combined with the numerical evidence that the GAS* mechanism's revenue is relatively insensitive to α, these results suggest that the mechanism is robust to moderate misspecification of the ad-nuisance parameter.

Appendix L: Application to the Video-on-Demand Market

In this section, we provide further details on our data collection. We design an online survey to recover the advertising nuisance and the willingness to pay on the video-on-demand (VoD) Market. We have the following steps:

- In the first step, we ask users whether they are subscribed to any streaming services.
- In the second step, we ask them to report their willingness to pay for the service if there are no ads and how much time they report using it. We elicit honest valuations of the willingness to pay (WTP) using the Becker–DeGroot–Marschak (BDM) (Becker et al. 1964) incentive-compatible method. The BDM method is frequently employed in experimental economics to determine an individual's reservation price for a good, even when there are no established market prices for the item (List and Shogren 2002). A key feature of the BDM method is that it captures valuations not influenced by strategic considerations. This is particularly important for our set-up, in which treatments are supposed to alter participants' beliefs about advertising amounts. In line with the BDM, we ask users for their WTP twice using different phrasing. To check whether users' reporting is rational, we randomly draw a price above or below their stated willingness to pay and ask whether they would take the service at this price. We discard users who report non-rational choices.

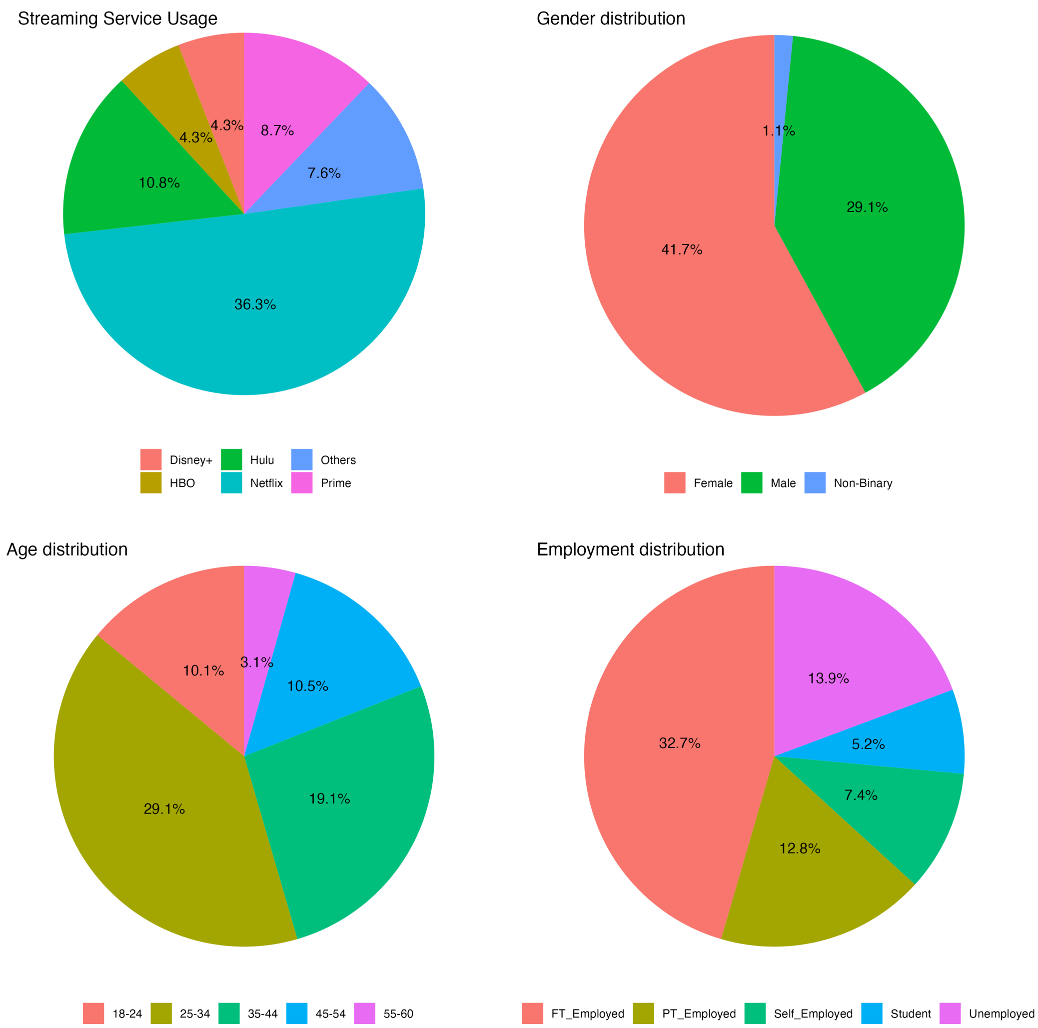


Figure 28 **Descriptive statistics**

- In the third step, we present users with four additional scenarios with different levels of ads: no ads, 1, 2, 3, or 4 ads per hour. We also randomized the announced ad duration at the user level, by breaks of 10 seconds between 30 seconds and 1 minute. We then ask them to report their usage time in minutes (without an upper limit) for each corresponding situation, while still paying the same subscription price as their initial willingness to pay. Overall, our database consists of 316 users, with 5 levels of ads, yielding 1580 observations.

Figure 28 summarizes the descriptive statistics of our sample. Because we observe users more than once, we capture variations in usage time within and between users and how they are affected by a change in advertising intensity. We display model-free evidence on our parameters in Figure 29.

Estimating advertising nuisance

As we have count data in minutes, we use a Poisson regression considering the level of advertising as continuous:

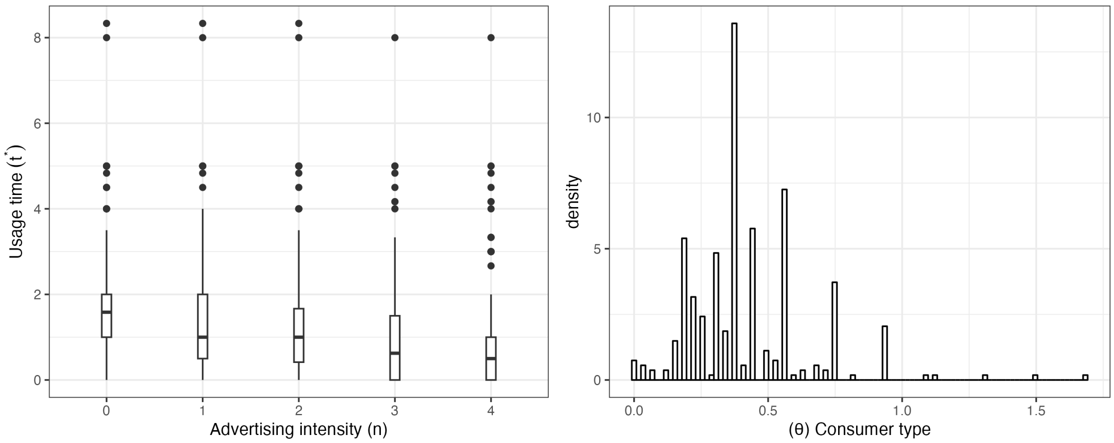


Figure 29 **The left panel displays box plots of usage time depending on the level of ad intensity. The right panel displays the distribution of user types.**

$$t_{ij} = \alpha_0 + \alpha n_j + \delta n_j^2 + \kappa_i + \varepsilon_{ji}, \tag{117}$$

Our coefficient of interest α captures whether the variation in advertising nuisance n impacts usage time t^* on the platform by user i. We also test for a non-linear relationship between time and ads, so we include a quadratic term n_j^2. We cluster the error at the user level.

	Dependent variable:					
	Usage time					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	4.675***		4.671***		4.567***	
	(0.0394)		(0.0379)		(0.0415)	
Ad intensity	-0.2084***	-0.2084***	-0.1964***	-0.1964***		
	(0.0150)	(0.0150)	(0.0259)	(0.0259)		
(Ad intensity)^2			-0.0032	-0.0032	-0.0512***	-0.0512***
			(0.0059)	(0.0059)	(0.0040)	(0.0040)
User FE	No	Yes	No	Yes	No	Yes
Observations	1580	1580	1580	1580	1580	1580

Clustered standard-errors at the user level
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 4 **Impact of advertising intensity on usage time - Poisson model.**

We find the quadratic term of advertising intensity not statistically significant when the linear term is included. Computing the effect of ads on usage time, we find that an increase in one ad decreases usage time by around $e^{-0.2084} - 1 = -18.8\%$ (see Table 4, Column (1)).

As a robustness check, we also specify our above strategy in a Tobit model with $t_{ij} = t^*_{ij}$ only if $t^*_{ij} > 0$ and $t_{ij} = 0$ otherwise, dropping the fixed effects. We find a similar effect, as displayed in Table 5.

	Dependent variable:		
	Usage time		
	(1)	(2)	(3)
Constant (Baseline)	104.63***	106.31***	92.23***
	(3.52)	(4.26)	(2.93)
Constant (Log SD)	4.38***	4.38***	4.39***
	(0.021)	(0.021)	(0.021)
Ad intensity	-19.63***	-23.06***	
	(1.47)	(5.11)	
(Ad intensity)^2		0.87	-4.48***
		(1.23)	(0.36)
Observations	1580	1580	1580

Standard errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 5 Impact of advertising intensity on usage time - Tobit model.

As a robustness check, we also provide a heterogeneity analysis based on the randomized ad durations in the experiment. Because we include user fixed effects in the regressions, we inherently control for ad duration when estimating how usage time varies with the number of ads people see. To make sure this is robust to the length of ads, we present a heterogeneity analysis in appendix in Table 6. We find no effect of ad duration, no matter the number of ads per hour, which leave our estimates unchanged.

Estimating user type distribution

As we have $V = WTP - p$, and we recover willingness-to-pay under a zero ad scenario, we obtain $WTP = \frac{1}{2}T\theta k$. Hence, for each user, we compute $\theta = \frac{2 \cdot WTP}{T \cdot k}$. We use $T = 30$ and the usage parameter k, calculated as the average daily viewing time under the ad-free subscription model, which equals $\hat{k} = 1.78$ hours (107 minutes) per user, which gives us $\theta = \frac{2 \cdot WTP}{30 \cdot 1.78}$

We estimate the distribution of user type θ, $F(\cdot)$. Our empirical results indicate that users' willingness to pay per hour (θ) follows a truncated exponential distribution $\hat{F}(\cdot)$ with parameter $\hat{\lambda} = 2.13$, supported on the interval $[0, 1.68]$.

	Dependent variable:					
	Usage time					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.5747***		0.5619***		0.4590***	
	(0.0760)		(0.0781)		(0.0792)	
Ad intensity	-0.2207***	-0.2207***	-0.1958***	-0.1958***		
	(0.0312)	(0.0312)	(0.0259)	(0.0259)		
Ad duration 40s	-0.0175		-0.0190		-0.0178	
	(0.1123)		(0.1172)		(0.1168)	
Ad duration 50s	-0.0099		0.0018		0.0006	
	(0.1099)		(0.1163)		(0.1155)	
Ad duration 60s	0.0516		0.0706		0.0682	
	(0.1085)		(0.1152)		(0.1146)	
Ad intensity × Ad duration 40s	-0.0140	-0.0140				
	(0.0433)	(0.0432)				
Ad intensity × Ad duration 50s	0.0201	0.0201				
	(0.0449)	(0.0449)				
Ad intensity × Ad duration 60s	0.0367	0.0367				
	(0.0405)	(0.0404)				
(Ad intensity)^2			-0.0060	-0.0060	-0.0541***	-0.0541***
			(0.0092)	(0.0092)	(0.0084)	(0.0084)
(Ad intensity)^2 × Ad duration 40s			-0.0047	-0.0047	-0.0050	-0.0050
			(0.0110)	(0.0110)	(0.0117)	(0.0117)
(Ad intensity)^2 × Ad duration 50s			0.0046	0.0046	0.0049	0.0049
			(0.0114)	(0.0114)	(0.0121)	(0.0121)
(Ad intensity)^2 × Ad duration 60s			0.0089	0.0089	0.0095	0.0095
			(0.0100)	(0.0100)	(0.0106)	(0.0106)
User FE	No	Yes	No	Yes	No	Yes
Observations	1580	1580	1580	1580	1580	1580

Clustered standard-errors at the user level
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 6 Impact of advertising intensity on usage time - Poisson model. The reference category is ad duration of 30 seconds. Hence, the *Constant* captures the reference category for no ads, and *Ad intensity* captures the average impact of one more ad for ad duration of 30 seconds.