An Economic Analysis of Peer-Disclosure in Online Social Communities

Online Supplement

Zike Cao†, Kai-Lung Hui‡, and Hong Xu‡

†Department of Technology and Operations Management, Rotterdam School of Management, Erasmus University, 3062 PA Rotterdam, The Netherlands. ‡Department of Information Systems, Business Statistics and Operations Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. Cao: cao@rsm.nl, Hui: klhui@ust.hk, Xu: hxu@ust.hk.

1. Omitted Proofs in the Main Model

Proof of Lemma 1. Note that $u^{sq,in}_{i|s} > u^{sq,out}_{i|s}$ is equivalent to $u^{sq,in-out}_{i|s} \equiv u^{sq,in}_{i|s} - u^{sq,out}_{i|s} > 0$. We can easily derive $u^{sq,in-out}_{i|s} = \frac{\gamma(1+\psi)v}{2c}[v - 2(1-\epsilon)s\lambda_i]$, $i \in \{L, H\}$. Note that $u^{sq,in-out}_{i|s}$ decreases in s and λ_i. Then any potential equilibrium can be characterized by one of the following 5 cases: (1) $u^{sq,in-out}_{H|s} < u^{sq,in-out}_{L|s} < 0$, (2) $u^{sq,in-out}_{H|s} < u^{sq,in-out}_{L|s} = 0$, (3) $u^{sq,in-out}_{H|s} < 0 < u^{sq,in-out}_{L|s}$, (4) $u^{sq,in-out}_{H|s} = 0 < u^{sq,in-out}_{L|s}$, or (5) $0 < u^{sq,in-out}_{H|s} < u^{sq,in-out}_{L|s}$. In the following, we derive the conditions for each equilibrium. Q denotes the total information, Π denotes the social welfare, ξ denotes the total *objective* privacy damage and ξ' denotes the total *perceived* privacy damage. The superscript sq indicates the status quo.

(1) If $u^{sq,in-out}_{H|s} < u^{sq,in-out}_{L|s} < 0$, then no uncommitted user will prefer to join, implying the equilibrium s is $s^{sq} = 1-\alpha$ (i.e., only committed users join). Substituting $s = 1-\alpha$ into $u^{sq,in-out}_{H|s} < u^{sq,in-out}_{L|s} < 0$, we get $v < 2(1-\epsilon)(1-\alpha)\lambda_L$, which is the sufficient and necessary condition for the existence of this equilibrium. Accordingly, $Q^{sq} = \frac{(1-\alpha)\gamma(1+\psi)v}{c}$, $\Pi^{sq} = \frac{(1-\alpha)\gamma(1+\psi)v}{2c}[v - 2(1-\alpha+\epsilon\alpha)\bar{\lambda}]$, $\xi^{sq} = \frac{(1-\alpha)\gamma\psi\bar{\theta}v}{c}$, and $\xi'^{sq} = \frac{(1-\alpha)\gamma\psi v}{c}[(1-\alpha+\epsilon\alpha)\bar{\theta}]$.

(2) If $u^{sq,in-out}_{H|s} < u^{sq,in-out}_{L|s} = 0$, then low-type uncommitted users are indifferent between joining and not joining while high-type uncommitted users prefer not to join, implying in equilibrium $s^{sq} \in [1-\alpha, 1-\alpha\beta]$. Meanwhile, the condition $u^{sq,in-out}_{L|s} = 0$ gives $s^{sq} = \frac{v}{2(1-\epsilon)\lambda_L}$. So $\frac{v}{2(1-\epsilon)\lambda_L} \in [1-\alpha, 1-\alpha\beta]$ leads to $2(1-\epsilon)(1-\alpha)\lambda_L \le v \le 2(1-\epsilon)(1-\alpha\beta)\lambda_L$, which is the sufficient and necessary condition for the existence of this equilibrium. In this equilibrium, a fraction s^{sq}_L of low-type uncommitted users choose to join while the other low-type uncommitted users choose not to join. Therefore $s^{sq}_L = \frac{s^{sq}-(1-\alpha)}{\alpha(1-\beta)} = \frac{v/(2(1-\epsilon)\lambda_L)-(1-\alpha)}{\alpha(1-\beta)}$. Accordingly, $Q^{sq} = \frac{\gamma(1+\psi)v^2}{2c(1-\epsilon)\lambda_L}$, $\Pi^{sq} = \frac{r(1+\psi)v^2}{2c}\left[(1-\alpha) - \frac{(1-\alpha+\epsilon\alpha)\bar{\lambda}}{(1-\epsilon)\lambda_L}\right]$, $\xi^{sq} = \frac{\gamma\psi\bar{\theta}v^2}{2c(1-\epsilon)\lambda_L}$, and $\xi'^{sq} = \frac{\gamma\psi v^2}{2c(1-\epsilon)\lambda_L}[(1-\alpha+\epsilon\alpha)\bar{\theta} + \alpha(1-\beta)(1-\epsilon)s^{sq}_L\theta_L]$.

(3) If $u^{sq,in-out}_{H|s} < 0 < u^{sq,in-out}_{L|s}$, then low-type uncommitted users prefer to join while high-type uncommitted users prefer not to join, implying in equilibrium $s^{sq} = 1-\alpha\beta$. Substituting $s = 1-\alpha\beta$ into $u^{sq,in-out}_{H|s} < 0 < u^{sq,in-out}_{L|s}$, we obtain $2(1-\epsilon)(1-\alpha\beta)\lambda_L < v < 2(1-\epsilon)(1-\alpha\beta)\lambda_H$, which is the sufficient and necessary condition for the existence of this equilibrium. Accordingly, $Q^{sq} = \frac{(1-\alpha\beta)\gamma(1+\psi)v}{c}$, $\Pi^{sq} = \frac{(1-\alpha\beta)\gamma(1+\psi)v}{2c}\{v - 2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H]\}$, $\xi^{sq} = \frac{(1-\alpha\beta)\gamma\psi\bar{\theta}v}{c}$, and $\xi'^{sq} = \frac{(1-\alpha\beta)\gamma\psi v}{c}[\bar{\theta} - (1-\epsilon)\alpha\beta\theta_H]$.

(4) if $u^{sq,in-out}_{H|s} = 0 < u^{sq,in-out}_{L|s}$, then low-type users prefer to join while high-type users are indifferent between joining and not joining, implying in equilibrium $s^{sq} \in [1-\alpha\beta, 1]$. Meanwhile, the condition $u^{sq,in-out}_{H|s} = 0$ gives $s^{sq} = \frac{v}{2(1-\epsilon)\lambda_H}$. So $\frac{v}{2(1-\epsilon)\lambda_H} \in [1-\alpha\beta, 1]$ leads to $2(1-\epsilon)(1-\alpha\beta)\lambda_H \le v \le 2(1-\epsilon)\lambda_H$, which is the sufficient and necessary condition for the existence of this equilibrium. In this equilibrium, a fraction s^{sq}_H of high-type uncommitted users choose to join while the other high-type uncommitted users choose not to join. Therefore $s^{sq}_H = \frac{s^{sq}-(1-\alpha\beta)}{\alpha\beta} = \frac{v/(2(1-\epsilon)\lambda_H)-(1-\alpha\beta)}{\alpha\beta}$. Accordingly, $Q^{sq} = \frac{\gamma(1+\psi)v^2}{2c(1-\epsilon)\lambda_H}$,

$\Pi^{sq} = \frac{r(1+\psi)v^2}{2c}\left[1 - \frac{\bar{\lambda}}{(1-\epsilon)\lambda_H}\right]$, $\xi^{sq} = \frac{\gamma\psi\bar{\theta}v^2}{2c(1-\epsilon)\lambda_H}$, and $\xi'^{sq} = \frac{\gamma\psi v^2}{2c(1-\epsilon)\lambda_H}[\bar{\theta} - (1-\epsilon)(1-s_H^{sq})\alpha\beta\theta_H]$.

(5) If $0 < u_{H|s}^{sq,in-out} < u_{L|s}^{sq,in-out}$, then all uncommitted users prefer to join, implying in equilibrium $s^{sq} = 1$. Substituting $s = 1$ into $0 < u_{H|s}^{sq,in-out} < u_{L|s}^{sq,in-out}$, we obtain $v > 2(1-\alpha\beta)\lambda_H$, which is the sufficient and necessary condition for the existence of this equilibrium. Accordingly, $Q^{sq} = \frac{\gamma(1+\psi)v}{c}$, $\Pi^{sq} = \frac{\gamma(1+\psi)v}{2c}(v - 2\bar{\lambda})$, $\xi^{sq} = \frac{\gamma\psi\bar{\theta}v}{c}$, and $\xi'^{sq} = \frac{\gamma\psi\bar{\theta}v}{c}$ (the same as ξ^{sq} since all users participate).

It is easy to check that $Q^{sq}(v)$, $\xi^{sq}(v)$, $\xi'^{sq}(v)$ are both continuous and increasing in v (to see that $\xi'^{sq}(v)$ increases in v, one should note $s_L^{sq}(v)$ and $s_H^{sq}(v)$ both increase in v).

Proof of Lemma 2. First, it is easy to check that $\Pi^{sq}(v)$ is continuous in v.

(1) If $0 < v < 2(1-\epsilon)(1-\alpha)\lambda_L$, $\Pi^{sq} = \frac{(1-\alpha)\gamma(1+\psi)v}{2c}[v - 2(1-\alpha+\epsilon\alpha)\bar{\lambda}] < 0$ because $v < 2(1-\epsilon)(1-\alpha)\lambda_L < 2(1-\alpha+\epsilon\alpha)\bar{\lambda}$. Π^{sq} is decreasing in v when $0 < v < \min\{2(1-\epsilon)(1-\alpha)\lambda_L, (1-\alpha+\epsilon\alpha)\bar{\lambda}\}$ and increasing in v when $(1-\alpha+\epsilon\alpha)\bar{\lambda} < v < 2(1-\epsilon)(1-\alpha)\lambda_L$.

(2) If $2(1-\epsilon)(1-\alpha)\lambda_L \leq v \leq 2(1-\epsilon)(1-\alpha\beta)\lambda_L$, $\Pi^{sq} = \frac{r(1+\psi)v^2}{2c}\left[(1-\alpha) - \frac{(1-\alpha+\epsilon\alpha)\bar{\lambda}}{(1-\epsilon)\lambda_L}\right] < 0$ because $(1-\alpha) - \frac{(1-\alpha+\epsilon\alpha)\bar{\lambda}}{(1-\epsilon)\lambda_L} < 0$. Obviously Π^{sq} is decreasing in v.

(3) If $2(1-\epsilon)(1-\alpha\beta)\lambda_L < v < 2(1-\epsilon)(1-\alpha\beta)\lambda_H$, $\Pi^{sq} = \frac{(1-\alpha\beta)\gamma(1+\psi)v}{2c}\{v - 2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H]\}$. First, note that $2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H] > 2(1-\epsilon)(1-\alpha\beta)\lambda_L$. If $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, then $2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H] \leq 2(1-\epsilon)(1-\alpha\beta)\lambda_H$, so $\Pi^{sq} < 0$ iff $2(1-\epsilon)(1-\alpha\beta)\lambda_L < v < 2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H]$; If $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, then $2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H] > 2(1-\epsilon)(1-\alpha\beta)\lambda_H$, so $\Pi^{sq} < 0$ iff $2(1-\epsilon)(1-\alpha\beta)\lambda_L < v < 2(1-\epsilon)(1-\alpha\beta)\lambda_H$. Π^{sq} is decreasing in v iff $2(1-\epsilon)(1-\alpha\beta)\lambda_L < v < \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H$ and increasing in v iff $\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H < v < 2(1-\epsilon)(1-\alpha\beta)\lambda_H$.

(4) If $2(1-\epsilon)(1-\alpha\beta)\lambda_H \leq v \leq 2(1-\epsilon)\lambda_H$, $\Pi^{sq} = \frac{r(1+\psi)v^2}{2c}\left[1 - \frac{\bar{\lambda}}{(1-\epsilon)\lambda_H}\right]$. If $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, then $1 - \frac{\bar{\lambda}}{(1-\epsilon)\lambda_H} > 0$, so $\Pi^{sq} > 0$ and is increasing in v; If $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, then $1 - \frac{\bar{\lambda}}{(1-\epsilon)\lambda_H} < 0$, so $\Pi^{sq} < 0$ and is decreasing in v.

(5) If $v > 2(1-\epsilon)\lambda_H$, $\Pi^{sq} = \frac{\gamma(1+\psi)v}{2c}(v - 2\bar{\lambda})$. If $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, then $2\bar{\lambda} \leq 2(1-\epsilon)\lambda_H$, so $\Pi^{sq} > 0$; If $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, then $2\bar{\lambda} > 2(1-\epsilon)\lambda_H$, so $\Pi^{sq} < 0$ iff $2(1-\epsilon)\lambda_H < v < 2\bar{\lambda}$. Π^{sq} is decreasing in v iff $2(1-\epsilon)\lambda_H < v < \bar{\lambda}$ and increasing in v iff $v > \max\{\bar{\lambda}, 2(1-\epsilon)\lambda_H\}$.

Lemma 2 summarizes (1) through (5) above. If $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H}) \Leftrightarrow \beta \leq \frac{(1-\epsilon)\lambda_H - \lambda_L}{\lambda_H - \lambda_L}$, condition (1), (2), and (3) together yield $\Pi^{sq} < 0$ iff $0 < v < 2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H]$; if $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H}) \Leftrightarrow \beta > \frac{(1-\epsilon)\lambda_H - \lambda_L}{\lambda_H - \lambda_L}$, condition (1), (2), (3), (4), and (5) together yield $\Pi^{sq} < 0$ iff $0 < v < 2\bar{\lambda}$. The two conditions are equivalent to conditions (i) and (ii) in Lemma 2.

It is also easy to check that Π^{sq} decreases in v only when $\Pi^{sq} < 0$.

Proof of Lemma 3. To obtain the equilibrium outcomes here, we follow the exact same approach in the proof of Lemma 1. Noting that the impact of a nudge is basically a linear reduction in v, we can easily obtain Lemma 3.

Proof of Proposition 1. Note that the impact of a nudge relative to the status quo is decreasing v. By Lemma 1, the participation size, s, weakly increases with v in the status quo, so decreasing v would (weakly) reduce the participation rate in the community. Similarly, by the proof of Lemma 1, the total amount of information and the total privacy damage both increase with v in the status quo, so decreasing v would reduce the total information and total privacy damage.

Proof of Proposition 2. By the the proof of Lemma 2, the social welfare in the status quo may decrease with v when it is negative, but the social welfare always increases with v when it is positive. So when the social welfare in the status quo is positive, imposing a positive nudge (i.e., decreasing v) would decrease the social welfare; when the social welfare in the status quo is negative, imposing a positive nudge (i.e., decreasing v) may increase the social welfare, but the maximum welfare is obviously zero, which is achieved by setting a sufficiently costly nudge, $\tau = v$.

Proof of Proposition 3. When user i makes decisions about the four types of postings, we can define a Lagrange function as the following (ignoring the externality terms since they are not user i's decisions)

$$\begin{aligned}\Phi(x_{ij}, y_{ij}, x_{ik}, y_{ik}, \kappa) =& n\left(vx_{ij} - \frac{cx_{ij}^2}{2} + vy_{ij} - \frac{cy_{ij}^2}{2\psi}\right) + (1-n)\left(vx_{ik} - \frac{cx_{ik}^2}{2\delta} + vy_{ik} - \frac{cy_{ik}^2}{2\delta\psi}\right) \\ &+ \kappa[n(x_{ij} + y_{ij}) + (1-n)(x_{ik} + y_{ik}) - \Lambda],\end{aligned}$$

where κ is the Lagrange multiplier. The FOCs are:

$$\begin{aligned}\frac{\partial}{\partial x_{ij}}\Phi &= n(v - cx_{ij}) + n\kappa = 0, \\ \frac{\partial}{\partial y_{ij}}\Phi &= n(v - \frac{c}{\psi}y_{ij}) + n\kappa = 0, \\ \frac{\partial}{\partial x_{ik}}\Phi &= (1-n)(v - \frac{c}{\delta}x_{ik}) + (1-n)\kappa = 0, \\ \frac{\partial}{\partial y_{ik}}\Phi &= (1-n)(v - \frac{c}{\delta\psi}y_{ik}) + (1-n)\kappa = 0, \\ \frac{\partial}{\partial \kappa}\Phi &= n(x_{ij} + y_{ij}) + (1-n)(x_{ik} + y_{ik}) - \Lambda = 0.\end{aligned}$$

Then we can easily derive: $x_{ij}^q = \frac{\Lambda}{(1+\psi)\gamma}$, $y_{ij}^q = \frac{\psi\Lambda}{(1+\psi)\gamma}$, $x_{ik}^q = \frac{\delta\Lambda}{(1+\psi)\gamma}$ and $y_{ik}^q = \frac{\delta\psi\Lambda}{(1+\psi)\gamma}$, where $\gamma = n+(1-n)\delta$. Since the utility function is concave, these quantities are obviously globally optimal ones.

Substituting the above equilibrium quantities back into equation (2), we can obtain user i's utility from participating in the community conditional on the participation size, s:

$$u_{i|s}^{q,in} = v\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)} - \Lambda s\lambda_i. \tag{A.1}$$

User i's utility from staying out conditional on s is

$$u_{i|s}^{q,out} = -\epsilon\Lambda s\lambda_i. \tag{A.2}$$

So user i's *net* utility from participating in the community is obtained by subtracting (A.2) from (A.1),

$$u_{i|s}^{q,in-out} = v\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)} - (1-\epsilon)\Lambda s\lambda_i = \Lambda\left[v - \frac{c\Lambda}{2\gamma(1+\psi)} - (1-\epsilon)s\lambda_i\right], i \in \{L, H\}. \tag{A.3}$$

Note that $u_{i|s}^{q,in-out}$ decreases in s and λ_i. Then any potential equilibrium can be characterized by one of the following five cases: (1) $u_{H|s}^{q,in-out} < u_{L|s}^{q,in-out} < 0$, (2) $u_{H|s}^{q,in-out} < u_{L|s}^{q,in-out} = 0$, (3) $u_{H|s}^{q,in-out} < 0 < u_{L|s}^{q,in-out}$, (4) $u_{H|s}^{q,in-out} = 0 < u_{L|s}^{q,in-out}$, or (5) $0 < u_{H|s}^{q,in-out} < u_{L|s}^{q,in-out}$. It is important to note that the participation rate increases with the numbering of the five cases: (5) > (4) > (3) > (2) > (1) in participation rate. We next derive the conditions (more importantly, the appropriate ranges for Λ) under which the *best* participation rate is (1), (2), (3), (4), or (5) respectively.

(1) If $u_{H|s}^{q,in-out} < u_{L|s}^{q,in-out} < 0$, then no uncommitted user will prefer to join, implying the equilibrium s is $s^q = 1-\alpha$ (i.e., only committed users join). Substituting $s = 1-\alpha$ into $u_{H|s}^{q,in-out} < u_{L|s}^{q,in-out} < 0$, we get $v < \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha)\lambda_L$, which is the sufficient and necessary condition for the existence of this equilibrium. Note that Λ is a decision variable and $\Lambda > 0$. To ensure this equilibrium is the best one, $v < \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha)\lambda_L$ must hold for any Λ, so $v \le (1-\epsilon)(1-\alpha)\lambda_L$. Therefore, when $v \le (1-\epsilon)(1-\alpha)\lambda_L$, no Λ can attract uncommitted users to join, so $s^q = 1-\alpha$. By Lemma 1, in this range of v, $s^{sq} = 1-\alpha$, so no Λ can improve the participation rate relative to the status quo.

(2) If $u_{H|s}^{q,in-out} < u_{L|s}^{q,in-out} = 0$, then low-type uncommitted users are indifferent between joining and not joining while high-type uncommitted users prefer not to join, implying in equilibrium $s^q \in [1-\alpha, 1-\alpha\beta]$. Meanwhile, the condition $u_{L|s}^{q,in-out} = 0$ gives $s^q = \frac{1}{(1-\epsilon)\lambda_L}\left[v - \frac{c\Lambda}{2\gamma(1+\psi)}\right]$. Noting the range of s^q, we obtain $\frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha)\lambda_L < v < \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha\beta)\lambda_L$, which is the sufficient and necessary condition for the existence of this equilibrium. To ensure this equilibrium is the best one, $v < \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha\beta)\lambda_L$ must hold for any Λ. Thus, $v \le (1-\epsilon)(1-\alpha\beta)\lambda_L$. Meanwhile, $\frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha)\lambda_L < v$ implies $v > (1-\epsilon)(1-\alpha)\lambda_L$ and $\Lambda < \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha)\lambda_L]}{c}$. Moreover, since $s^q = \frac{1}{(1-\epsilon)\lambda_L}\left[v - \frac{c\Lambda}{2\gamma(1+\psi)}\right]$, a smaller Λ yields a larger s^q. So, to maximize s^q, Λ should be set infinitesimally close to zero, $\Lambda^\star = \iota$, so the

maximum participation size $s^\star = \frac{v-\iota}{(1-\epsilon)\lambda_L}$. Therefore, when $(1-\epsilon)(1-\alpha)\lambda_L < v \le (1-\epsilon)(1-\alpha\beta)\lambda_L$, $\Lambda^\star = \iota$ can maximize the participation rate, and the fraction of uncommitted low-type users who participate is $f_L^\star = \frac{(v-\iota)/((1-\epsilon)\lambda_L)-(1-\alpha)}{\alpha(1-\beta)}$. By Lemma 1, in this range of v, $s^{sq} = \max\{1-\alpha, \frac{v}{2(1-\epsilon)\lambda_L}\} < s^\star$, so $\Lambda^\star$ improves the participation rate relative to the status quo.

(3) if $u_{H|s}^{q,in-out} < 0 < u_{L|s}^{q,in-out}$, then low-type uncommitted users prefer to join while high-type uncommitted users prefer not to join, implying in equilibrium $s^q = 1-\alpha\beta$. Substituting $s = 1-\alpha\beta$ into $u_{H|s}^{q,in-out} < 0 < u_{L|s}^{q,in-out}$, we obtain $\frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha\beta)\lambda_L < v < \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha\beta)\lambda_H$, which is the sufficient and necessary condition for the existence of this equilibrium. To ensure this equilibrium is the best one, $v < \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha\beta)\lambda_H$ must hold for any Λ, then we must have $v \le (1-\epsilon)(1-\alpha\beta)\lambda_H$. Meanwhile, $\frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha\beta)\lambda_L < v$ implies $v > (1-\epsilon)(1-\alpha\beta)\lambda_L$ and $\Lambda < \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_L]}{c}$. Therefore, when $(1-\epsilon)(1-\alpha\beta)\lambda_L < v \le (1-\epsilon)(1-\alpha\beta)\lambda_H$, $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_L]}{c}\right)$ can maximize the participation rate, and only all uncommitted low-type users participate ($s^\star = 1-\alpha\beta$). By Lemma 1, in this range of v, $s^{sq} = \min\{\frac{v}{2(1-\epsilon)\lambda_L}, 1-\alpha\beta\}$. So, when $(1-\epsilon)(1-\alpha\beta)\lambda_L < v \le \min\{(1-\epsilon)(1-\alpha\beta)\lambda_H, 2(1-\epsilon)(1-\alpha\beta)\lambda_L\}$, $s^{sq} = \frac{v}{2(1-\epsilon)\lambda_L} < s^\star$ and $\Lambda^\star$ improves the participation rate relative to the status quo; when $2(1-\epsilon)(1-\alpha\beta)\lambda_L < v \le (1-\epsilon)(1-\alpha\beta)\lambda_H$, $s^{sq} = 1-\alpha\beta = s^\star$, $\Lambda^\star$ does not change the participation rate relative to the status quo.

(4) If $u_{H|s}^{q,in-out} = 0 < u_{L|s}^{q,in-out}$, then low-type users prefer to join while high-type users are indifferent between joining and not joining, implying in equilibrium $s^{sq} \in [1-\alpha\beta, 1]$. Meanwhile, the condition $u_{H|s}^{q,in-out} = 0$ gives $s^q = \frac{1}{(1-\epsilon)\lambda_H}\left[v - \frac{c\Lambda}{2\gamma(1+\psi)}\right]$. Noting the range of s^q, we obtain $\frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha\beta)\lambda_H < v < \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)\lambda_H$, which is the sufficient and necessary condition for the existence of this equilibrium. To ensure this equilibrium is the best one, $v < \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)\lambda_H$ must hold for any Λ, so $v \le (1-\epsilon)\lambda_H$. Meanwhile, $\frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)(1-\alpha\beta)\lambda_H < v$ implies $v > (1-\epsilon)(1-\alpha\beta)\lambda_H$ and $\Lambda < \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_H]}{c}$. Moreover, since $s^q = \frac{1}{(1-\epsilon)\lambda_H}\left[v - \frac{c\Lambda}{2\gamma(1+\psi)}\right]$, a smaller Λ yields a larger s^q. So to maximize s^q, Λ should be set infinitesimally close to zero, i.e., $\Lambda^\star = \iota$, so the maximum participation size $s^\star = \frac{v-\iota}{(1-\epsilon)\lambda_H}$. Therefore, when $(1-\epsilon)(1-\alpha\beta)\lambda_H < v \le (1-\epsilon)\lambda_H$, $\Lambda^\star = \iota$ can maximize the participation rate: All uncommitted low-type users participate and the fraction of uncommitted high-type users who participate is $f_H^\star = \frac{(v-\iota)/((1-\epsilon)\lambda_H)-(1-\alpha\beta)}{\alpha\beta}$. By Lemma 1, in this range of v, $s^{sq} = \max\{1-\alpha\beta, \frac{v}{2(1-\epsilon)\lambda_H}\} < s^\star$, so $\Lambda^\star$ improves the participation rate relative to the status quo.

(5) If $0 < u_{H|s}^{q,in-out} < u_{L|s}^{q,in-out}$, then all uncommitted users prefer to join, implying in equilibrium $s^q = 1$. Substituting $s = 1$ into $0 < u_{H|s}^{q,in-out} < u_{L|s}^{q,in-out}$, we obtain $\frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)\lambda_H < v$, which is the sufficient and necessary condition for the existence of this equilibrium. To ensure we are able to achieve this equilibrium, we must have $v > (1-\epsilon)\lambda_H$ and $\Lambda < \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}$. Therefore, when $v > (1-\epsilon)\lambda_H$,

$\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}\right)$ can maximize the participation rate, and all uncommitted users participate ($s^\star = 1$). By Lemma 1, in this range of v, $s^{sq} = \min\{\frac{v}{2(1-\epsilon)\lambda_H}, 1\}$. So, when $(1-\epsilon)\lambda_H < v < 2(1-\epsilon)\lambda_H$, $s^{sq} = \frac{v}{2(1-\epsilon)\lambda_H} < s^\star$ and $\Lambda^\star$ improves the participation rate relative to the status quo; when $v \geq 2(1-\epsilon)\lambda_H$, $s^{sq} = 1 = s^\star$ and $\Lambda^\star$ does not change the participation rate relative to the status quo.

Proposition 3 summarizes the above five cases and highlights the conditions under which quota can be used to improve the participation rate relative to the status quo.

Proof of Proposition 4. With an effective quota, $\Lambda \in (0, \frac{\gamma(1+\psi)v}{c})$, all users will post up to the quota. And relative to the status quo, each individual posts less. So if an effective quota increases the total information posted in the community relative to the status quo, it must have attracted more users to participate in the community. However, this would never happen. We next show why it is the case.

First, given the participation size s, the total amount of information posted in the community under quota Λ is $Q^q(s) = s\Lambda$. So equation (A.3) (i.e., a representative user i's net utility from participating in the community under quota Λ) can be written as

$$u_{i|s}^{q,in-out} = \underbrace{v\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)}}_{i} - \underbrace{(1-\epsilon)Q^q(s)\lambda_i}_{ii}, i \in \{L, H\}. \tag{A.4}$$

Part (i) in (A.4) is user i's net posting benefit under Λ and is less than $\frac{r(1+\psi)v^2}{2c}$ due to $\Lambda < \frac{\gamma(1+\psi)v}{c}$, meaning that any effective quota would make each user enjoy less net posting benefit relative to the status quo. Now if $Q^q(s)$ is increased relative to the status quo, then Part (ii) in (A.4) is larger than in the status quo, meaning that each user will suffer more net privacy cost. Then each user will unambiguously enjoy less net utility from participating in the community due to the effective quota, Λ, relative to the status quo (since Part (i) is smaller and Part (ii) is larger), which is in contradiction with the prerequisite that more users should participate relative to the status quo for the total information to increase.

Proof of Proposition 5.

We now derive the optimal quota to maximize social welfare, Λ^*.

(1) When $0 < v \leq (1-\epsilon)(1-\alpha)\lambda_L$, by Proposition 3, no quota can change the participation rate in the status quo. So in any quota, no uncommitted users will participate (labeled as Case A). In Case A, the aggregate user welfare is ($\Lambda \in (0, \frac{\gamma(1+\psi)v}{c}]$):

$$\Pi_A^q(\Lambda) = (1-\alpha)\left\{[v - (1-\alpha+\epsilon\alpha)\bar{\lambda}]\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)}\right\}. \tag{A.5}$$

Note that $v - (1 - \alpha + \epsilon\alpha)\bar{\lambda} < 0$ since $(1 - \alpha + \epsilon\alpha)\bar{\lambda} > (1 - \epsilon)(1 - \alpha)\lambda_L$, and thus $\Pi_A^q(\Lambda)$ decreases in Λ. So the welfare-optimal quota is $\Lambda^* = \iota$ and accordingly $\Pi^{q*} = -\iota$. With Λ^* in place, no uncommitted users will participate and $s^{q*} = 1 - \alpha$.

(2) When $(1-\epsilon)(1-\alpha)\lambda_L < v \leq (1-\epsilon)(1-\alpha\beta)\lambda_L$, by Proposition 3, the best a quota can do is to attract a fraction of uncommitted low-type users to participate (labeled as Case B). In Case B, the requirement on Λ is $\Lambda \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha)\lambda_L]}{c}\right]$. In equilibrium, uncommitted low-type users are indifferent between participating and not participating. That is, $u_{L|s}^{q,in-out} = 0 \Leftrightarrow v = \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)s^q\lambda_L$, where s^q is the equilibrium participation rate and changes as Λ changes. Using this equality, the aggregate user welfare in Case B can be written as

$$\begin{aligned} \Pi_B^q(\Lambda) &= (1-\alpha\beta)\left\{\left\{v - \frac{s^q}{1-\alpha\beta}[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H]\right\}\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)}\right\} \\ &= (1-\alpha\beta)\left[1 - \frac{\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H}{(1-\epsilon)(1-\alpha\beta)\lambda_L}\right]\left[v\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)}\right]. \end{aligned} \tag{A.6}$$

Note that, in (A.6), $1 - \frac{\bar{\lambda}-(1-\epsilon)\alpha\beta\lambda_H}{(1-\epsilon)(1-\alpha\beta)\lambda_L} < 0$ since $\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H > (1-\epsilon)(1-\alpha\beta)\lambda_L$, and thus $\Pi_B^q(\Lambda)$ decreases in Λ. Case A is also achievable, but it obviously cannot do better than Case B.

So the welfare-optimal quota is also $\Lambda^* = \iota$ and accordingly $\Pi^{q*} = -\iota$. With Λ^* in place, only a fraction of uncommitted low-type users will participate and $s^{q*} = \frac{v-\iota}{(1-\epsilon)\lambda_L}$.

(3) When $(1-\epsilon)(1-\alpha\beta)\lambda_L < v \leq (1-\epsilon)(1-\alpha\beta)\lambda_H$, by Proposition 3, the best a quota can do is to attract all the uncommitted low-type users (labeled as Case C). In Case C, the requirement on Λ is $\Lambda \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_L]}{c}\right)$. The aggregate user welfare in Case C is

$$\Pi_C^q(\Lambda) = (1-\alpha\beta)\left\{[v - (\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H)]\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)}\right\}. \tag{A.7}$$

We can easily derive the welfare-optimal quota according to (A.7):

$$\Lambda_C^* = \begin{cases} \frac{\gamma(1+\psi)[v-\bar{\lambda}+(1-\epsilon)\alpha\beta\lambda_H]}{c}, & \text{if } v > \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H; \\ \iota, & \text{if } v \leq \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H. \end{cases}$$

Accordingly,

$$\Pi_C^{q*} = \begin{cases} \frac{(1-\alpha\beta)\gamma(1+\psi)[v-\bar{\lambda}+(1-\epsilon)\alpha\beta\lambda_H]^2}{2c}, & \text{if } v > \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H; \\ -\iota, & \text{if } v \leq \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H. \end{cases}$$

Case A or B is also achievable, but they obviously cannot do better than Case C. Note that $\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H > (1-\epsilon)(1-\alpha\beta)\lambda_L$, so the optimal quota in this case can be expressed as

$$\Lambda^* = \begin{cases} \iota, & \text{if } (1-\epsilon)(1-\alpha)\lambda_L < v \leq \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H; \\ \frac{\gamma(1+\psi)[v-\bar{\lambda}+(1-\epsilon)\alpha\beta\lambda_H]}{c}, & \text{if } \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H < v \leq (1-\epsilon)(1-\alpha\beta)\lambda_H. \end{cases} \tag{A.8}$$

The second case exists only if $\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H < (1-\epsilon)(1-\alpha\beta)\lambda_H \Leftrightarrow \epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$. With Λ^* in place, only all uncommitted low-type users will participate and $s^{q*} = 1-\alpha\beta$.

(4) When $(1-\epsilon)(1-\alpha\beta)\lambda_H < v \leq (1-\epsilon)\lambda_H$, by Proposition 3, the best a quota can do is to attract all the uncommitted low-type users and a fraction of uncommitted high-type users (labeled as Case D). In Case D, the requirement on Λ is $\Lambda \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_H]}{c}\right)$. In equilibrium, uncommitted high-type users are indifferent between participating and not participating. That is, $u^{q,in-out}_{H|s} = 0 \Leftrightarrow v = \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)s^q\lambda_H$, where s^q is the equilibrium participation rate and changes as Λ changes. Using this equality, the aggregate user welfare in Case D can be written as

$$\Pi^q_D(\Lambda) = (v - s\bar{\lambda})\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)} = \left[1 - \frac{\bar{\lambda}}{(1-\epsilon)\lambda_H}\right]\left[v\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)}\right]. \tag{A.9}$$

Note that $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H}) \Leftrightarrow \bar{\lambda} > (1-\epsilon)\lambda_H \Leftrightarrow 1 - \frac{\bar{\lambda}}{(1-\epsilon)\lambda_H} < 0$, so when $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, $\Pi^q_D(\Lambda)$ decreases in Λ; when $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, $\Pi^q_D(\Lambda)$ increases in $\Lambda \in (0, \frac{\gamma(1+\psi)v}{c}]$. Therefore, we can derive the welfare-optimal quota according to (A.9)

$$\Lambda^*_D = \begin{cases} \iota, & \text{if } \epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H}); \\ \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_H]}{c}, & \text{if } \epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H}). \end{cases} \tag{A.10}$$

• When $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, Case C, B, or A obviously cannot do better than Case D. So the optimal quota is $\Lambda^* = \iota$, and all uncommitted low-type and a fraction of uncommitted high-type users will participate and $s^{q*} = \frac{v-\iota}{(1-\epsilon)\lambda_H}$.

• When $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, the optimal quota Λ^*_D in equation (A.10) degenerates to Case C wherein only all the uncommitted low-type users participate (because Λ^*_D is the corner solution). Moreover, we need to consider when $v > 2(1-\epsilon)(1-\alpha\beta)\lambda_H \Leftrightarrow \frac{\gamma(1+\psi)v}{c} < \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_H]}{c}$, the corner solution is not achievable.

Under $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, consider the following.

(4.1) If $(1-\epsilon)(1-\alpha\beta)\lambda_H < v \leq \min\{(1-\epsilon)\lambda_H, 2(1-\epsilon)(1-\alpha\beta)\lambda_H\}$, we need to compare the corner solution with the interior solution in Case C to see when the interior solution is achievable. If yes, the interior solution is globally optimal; if not the corner solution is globally optimal. Cases B and A are obviously inferior. It is easy to obtain that

$$\Lambda^* = \begin{cases} \frac{\gamma(1+\psi)[v-\bar{\lambda}+(1-\epsilon)\alpha\beta\lambda_H]}{c}, & \text{if } v \le (1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda}; \\ \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_H]}{c}, & \text{if } v > (1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda}. \end{cases} \tag{A.11}$$

With Λ^* in place, only all uncommitted low-type users will participate and $s^{q*} = 1 - \alpha\beta$.

(4.2) If $2(1-\epsilon)(1-\alpha\beta)\lambda_H < v \le (1-\epsilon)\lambda_H$, Case C is not achievable now. Moreover, (A.9) is increasing in $\Lambda \in (0, \frac{\gamma(1+\psi)}{c}]$, so $\Lambda^* = \frac{\gamma(1+\psi)}{c}$ is ineffective. With Λ^* in place, all uncommitted low-type users and a fraction of uncommitted high-type users will participate and $s^{q*} = \frac{v}{2(1-\epsilon)\lambda_H}$.

(5) When $v > (1-\epsilon)\lambda_H$, by Proposition 3, the best a quota can do is attract all the uncommitted users (labeled as Case E). In Case E, the requirement on Λ is $\Lambda \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}\right]$. The aggregate user welfare in Case E is

$$\Pi_E^q(\Lambda) = (v - \bar{\lambda})\Lambda - \frac{c\Lambda^2}{2\gamma(1+\psi)}. \tag{A.12}$$

Note that $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H}) \Leftrightarrow \bar{\lambda} > (1-\epsilon)\lambda_H$. We can derive the welfare-optimal quota according to (A.12).

- When $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$,

$$\Lambda_E^* = \begin{cases} \iota, & \text{if } v \le \bar{\lambda}; \\ \frac{\gamma(1+\psi)(v-\bar{\lambda})}{c}, & \text{if } v > \bar{\lambda}. \end{cases} \tag{A.13}$$

It is easy to check Λ_E^* is indeed globally optimal. So Λ^* is expressed as in (A.13). With Λ^*, all uncommitted users will participate.

- When $\epsilon \le (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$,

$$\Lambda_E^* = \begin{cases} \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}, & \text{if } v \le 2(1-\epsilon)\lambda_H - \bar{\lambda}; \\ \frac{\gamma(1+\psi)(v-\bar{\lambda})}{c}, & \text{if } v > 2(1-\epsilon)\lambda_H - \bar{\lambda}. \end{cases} \tag{A.14}$$

We need to see if Case E is globally optimal. Under $\epsilon \le (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, consider the following.

(5.1) If $(1-\epsilon)\lambda_H < v < \min\{2(1-\epsilon)(1-\alpha\beta)\lambda_H, 2(1-\epsilon)\lambda_H - \bar{\lambda}\}$, Λ_E^* is the corner solution and can be easily proved to be inferior to Case C. Analysis as in (4.1) would yield

$$\Lambda^* = \begin{cases} \frac{\gamma(1+\psi)[v-\bar{\lambda}+(1-\epsilon)\alpha\beta\lambda_H]}{c}, & \text{if } v \le (1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda}; \\ \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_H]}{c}, & \text{if } v > (1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda}. \end{cases} \tag{A.15}$$

With Λ^* in place, only all uncommitted low-type users will participate and $s^{q*} = 1 - \alpha\beta$.

(5.2) If $\max\{(1-\epsilon)\lambda_H, 2(1-\epsilon)(1-\alpha\beta)\lambda_H\} \leq v \leq 2(1-\epsilon)\lambda_H - \bar{\lambda}$, Λ_E^* is the corner solution and can be easily proved to be inferior to the status quo (Case C is not achievable now). So $\Lambda^* = \frac{\gamma(1+\psi)}{c}$ and is ineffective. With Λ^* in place, all uncommitted low-type users and a fraction of uncommitted high-type users will participate and $s^{q*} = \frac{v}{2(1-\epsilon)\lambda_H}$.

(5.3) If $v > \max\{2(1-\epsilon)(1-\alpha\beta)\lambda_H, 2(1-\epsilon)\lambda_H - \bar{\lambda}\}$, Λ_E^* is the interior solution and is indeed globally optimal. That is, $\Lambda^* = \frac{\gamma(1+\psi)(v-\bar{\lambda})}{c}$. With Λ^* in place, all uncommitted users will participate and $s^{q*} = 1$.

Appendix A in the main manuscript summarizes the above five cases. We present the optimal quotas separately according to the value of ϵ. We have also combined adjacent ranges of v under which the optimal quotas are common.

Lastly, it is easy to verify that except for conditions (4.2) and (5.2), all the optimal quotas identified above are effective. That is, $\Lambda^* < \frac{\gamma(1+\psi)v}{c}$. Also note that when the quota is ineffective (i.e., $\Lambda = \frac{\gamma(1+\psi)v}{c}$), the quota is equivalent to the status quo. In the process of searching for the optimal quota above, ineffective quotas have also been considered. So when the optimal quota is effective, it will always improve the welfare relative to the status quo. In conditions (4.2) and (5.2), which can be combined as $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$ and $2(1-\epsilon)(1-\alpha\beta)\lambda_H < v < 2(1-\epsilon)\lambda_H - \bar{\lambda}$, the optimal quota is ineffective. That is, $\Lambda^* = \frac{\gamma(1+\psi)v}{c}$ and the optimal quota retains the status quo.

It is easy to verify that the participation rate in the community is weakly improved relative to the status quo. The aggregate privacy damage is $\xi = \frac{\psi\bar{\theta}}{1+\psi}Q$. When the optimal quota is effective, by Proposition 4, it will always reduce the total amount of information in the community (i.e., Q), so it will always reduce the aggregate privacy damage.

Proof of Proposition 6. By Proposition 1, a positive nudge weakly reduces the participation rate relative to the status quo. Meanwhile, by Proposition 3, a quota can be used to improve the participation rate relative to the status quo in some ranges. So obviously, a quota weakly dominates a nudge in increasing user participation.

Now consider welfare, we consider two cases according to the value of ϵ:

• When $\epsilon \leq (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, by Lemma 2 and Proposition 2, $\Pi^{n*} = 0$ when $0 < v < 2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H]$ and $\Pi^{n*} = \Pi^{sq}$ when $v \geq 2[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H]$. Meanwhile, by Proposition 5 and Appendix A, $\Pi^{q*} = 0$ when $0 < v < \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H$ and $\Pi^{q*} \geq \Pi^{sq}$ when $v \geq \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H$. So $\Pi^{q*} \geq \Pi^{n*}$.

• When $\epsilon > (1-\beta)(1-\frac{\lambda_L}{\lambda_H})$, by Lemma 2 and Proposition 2, $\Pi^{n*} = 0$ when $0 < v < 2\bar{\lambda}$ and $\Pi^{n*} = \Pi^{sq}$ when $v \geq 2\bar{\lambda}$. Meanwhile, by Proposition 5 and Appendix A, $\Pi^{q*} = 0$ when $0 < v < \bar{\lambda}$ and $\Pi^{q*} > \Pi^{sq}$ when $v \geq \bar{\lambda}$. So $\Pi^{q*} \geq \Pi^{n*}$.

Note a technical issue worth clarifying: We consider an extremely costly nudge (i.e., $\tau = v$ and thus $\Pi^n = 0$) equivalent to an extremely harsh quota (i.e., $\Lambda = \iota$ and thus $\Pi^q = -\iota$). This is purely because we have assumed $\Lambda \neq 0$ to avoid some trivial discussion.

Section 3.3.5. Formal results about overlapping of participation-optimal and welfare-optimal quotas. When the community owner's objective is to maximize participation, by Proposition 3, a quota can be used to achieve this goal. We now examine when the welfare-optimal quota (Λ^*) overlaps with the participation-optimal quota ($\Lambda^\star$). Referring to Proposition 3 and Appendix A in the main manuscript,

(i) $\epsilon \leq (1-\beta)\left(1 - \frac{\lambda_L}{\lambda_H}\right)$

(i.1) If $0 < v \leq (1-\epsilon)(1-\alpha)\lambda_L$, $\Lambda^* = \iota$, $\Lambda^\star \in \left(0, \frac{\gamma(1+\psi)v}{c}\right]$, and thus $\Lambda^* \in \Lambda^\star$.

(i.2) If $(1-\epsilon)(1-\alpha)\lambda_L < v \leq (1-\epsilon)(1-\alpha\beta)\lambda_L$, $\Lambda^* = \iota$, $\Lambda^\star = \iota$, and thus $\Lambda^* = \Lambda^\star$.

(i.3) If $(1-\epsilon)(1-\alpha\beta)\lambda_L < v \leq \bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H$, $\Lambda^* = \iota$, $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_L]}{c}\right)$, and thus $\Lambda^* \in \Lambda^\star$.

(i.4) If $\bar{\lambda}-(1-\epsilon)\alpha\beta\lambda_H < v \leq (1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda}$, $\Lambda^* = \frac{\gamma(1+\psi)[v-\bar{\lambda}+(1-\epsilon)\alpha\beta\lambda_H]}{c}$. $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_L]}{c}\right)$ when $v < \min\{(1-\epsilon)(1-\alpha\beta)\lambda_H, 2(1-\epsilon)(1-\alpha\beta)\lambda_L\}$, so $\Lambda^* \in \Lambda^\star$; $\Lambda^\star \in \left(0, \frac{\gamma(1+\psi)v}{c}\right)$ when $2(1-\epsilon)(1-\alpha\beta)\lambda_L < v < (1-\epsilon)(1-\alpha\beta)\lambda_H$, so $\Lambda^* \in \Lambda^\star$; $\Lambda^\star = \iota$ when $(1-\epsilon)(1-\alpha\beta)\lambda_H < v < (1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda}$, so $\Lambda^* \neq \Lambda^\star$.

(i.5) If $(1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda} < v \leq 2(1-\epsilon)(1-\alpha\beta)\lambda_H$, $\Lambda^* = \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_H]}{c}$. $\Lambda^\star = \iota$ when $(1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda} < v < (1-\epsilon)\lambda_H$, so $\Lambda^* \neq \Lambda^\star$; $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}\right]$ when $\max\{(1-\epsilon)(2-\alpha\beta)\lambda_H - \bar{\lambda}, (1-\epsilon)\lambda_H\} < v \leq 2(1-\epsilon)(1-\alpha\beta)\lambda_H$, so $\Lambda^* \neq \Lambda^\star$.

(i.6) If $2(1-\epsilon)(1-\alpha\beta)\lambda_H < v \leq 2(1-\epsilon)\lambda_H - \bar{\lambda}$, $\Lambda^* = \frac{\gamma(1+\psi)v}{c}$. $\Lambda^\star = \iota$ when $2(1-\epsilon)(1-\alpha\beta)\lambda_H < v < (1-\epsilon)\lambda_H$, so $\Lambda^* \neq \Lambda^\star$; $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}\right]$ when $\max\{2(1-\epsilon)(1-\alpha\beta)\lambda_H, (1-\epsilon)\lambda_H\} < v \leq 2(1-\epsilon)\lambda_H - \bar{\lambda}$, so $\Lambda^* \neq \Lambda^\star$.

(i.7) If $v > \max\{2(1-\epsilon)(1-\alpha\beta)\lambda_H, 2(1-\epsilon)\lambda_H - \bar{\lambda}\}$, $\Lambda^* = \frac{\gamma(1+\psi)(v-\bar{\lambda})}{c}$. $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}\right]$ when $\max\{2(1-\epsilon)(1-\alpha\beta)\lambda_H, 2(1-\epsilon)\lambda_H - \bar{\lambda}\} < v < 2(1-\epsilon)\lambda_H$, so $\Lambda^* \in \Lambda^\star$; $\Lambda^\star \in \left(0, \frac{\gamma(1+\psi)v}{c}\right]$ when $v > 2(1-\epsilon)\lambda_H$, so $\Lambda^* \in \Lambda^\star$.

(ii) $\epsilon > (1-\beta)\left(1 - \frac{\lambda_L}{\lambda_H}\right)$

(ii.1) If $0 < v \leq (1-\epsilon)(1-\alpha)\lambda_L$, $\Lambda^* = \iota$, $\Lambda^\star \in \left(0, \frac{\gamma(1+\psi)v}{c}\right]$, and thus $\Lambda^* \in \Lambda^\star$.

(ii.2) If $(1-\epsilon)(1-\alpha)\lambda_L < v \leq (1-\epsilon)(1-\alpha\beta)\lambda_L$, $\Lambda^* = \iota$, $\Lambda^\star = \iota$, and thus $\Lambda^* = \Lambda^\star$.

(ii.3) If $(1-\epsilon)(1-\alpha\beta)\lambda_L < v \le (1-\epsilon)(1-\alpha\beta)\lambda_H$, $\Lambda^* = \iota$, $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)(1-\alpha\beta)\lambda_L]}{c}\right)$, and thus $\Lambda^* \in \Lambda^\star$.

(ii.4) If $(1-\epsilon)(1-\alpha\beta)\lambda_H < v \le (1-\epsilon)\lambda_H$, $\Lambda^* = \iota$, $\Lambda^\star = \iota$, and thus $\Lambda^* = \Lambda^\star$.

(ii.5) If $(1-\epsilon)\lambda_H < v \le \bar{\lambda}$, $\Lambda^* = \iota$, $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}\right]$, and thus $\Lambda^* \in \Lambda^\star$.

(ii.6) If $v > \bar{\lambda}$, $\Lambda^* = \frac{\gamma(1+\psi)(v-\bar{\lambda})}{c}$. $\Lambda^\star \in \left(0, \frac{2\gamma(1+\psi)[v-(1-\epsilon)\lambda_H]}{c}\right]$ when $\bar{\lambda} < v < 2(1-\epsilon)\lambda$, and thus $\Lambda^* \in \Lambda^\star$; $\Lambda^\star \in \left(0, \frac{\gamma(1+\psi)v}{c}\right]$ when $v > 2(1-\epsilon)\lambda_H$, so $\Lambda^* \in \Lambda^\star$.

Proof of Proposition 7. Given a free allowance $\tilde{\Lambda} \in \left(0, \frac{\gamma(1+\psi)v}{c}\right]$, if user i does not post in excess of this $\tilde{\Lambda}$, then the equilibrium quantities will be the same as those in a pure quota. So, by Proposition 3, we have $x_{ij} = \frac{\Lambda}{(1+\psi)\gamma}$, $y_{ij} = \frac{\psi\Lambda}{(1+\psi)\gamma}$, $x_{ik} = \frac{\delta\Lambda}{(1+\psi)\gamma}$, and $y_{ik} = \frac{\delta\psi\Lambda}{(1+\psi)\gamma}$. If user i posts in excess of the free allowance, then she will face an additional nudging cost for the excessive information she posts.

Denote the amount of nonsensitive information user i would post about user j as x^c_{ij}. If user i decides to post in excess of $\tilde{\Lambda}$, she makes the following decision: $x^c_{ij} = \arg\max_{x_{ij} > \frac{\Lambda}{(1+\psi)\gamma}} \; vx_{ij} - \frac{cx^2_{ij}}{2} - \tilde{\tau}[x_{ij} - \frac{\Lambda}{(1+\psi)\gamma}] = \frac{v-\tilde{\tau}}{c}$. Similarly, we derive the optimal quantities of the other three types of information: $y^c_{ij} = \frac{\psi(v-\tilde{\tau})}{c}$, $x^c_{ik} = \frac{\delta(v-\tilde{\tau})}{c}$, and $y^c_{ik} = \frac{\psi\delta(v-\tilde{\tau})}{c}$. So the total amount of information each user would post is $Q^c = n(x^c_{ij} + y^c_{ij}) + (1-n)(x^c_{ik} + y^c_{ik}) = \frac{\gamma(1+\psi)(v-\tilde{\tau})}{c}$. So if $Q^c = \frac{\gamma(1+\psi)(v-\tilde{\tau})}{c} > \tilde{\Lambda}$, then users will post in excess of the free allowance, otherwise they just use up the free allowance.

We now see, when users post in excess of the free allowance (i.e., $\frac{\gamma(1+\psi)(v-\tilde{\tau})}{c} > \tilde{\Lambda} \Leftrightarrow \tilde{\tau} < v - \frac{c\tilde{\Lambda}}{\gamma(1+\psi)}$), whether the composite policy will outperform the pure quota in terms of participation and welfare. Substituting the equilibrium quantities of information derived above back into equations (2) and (3) in the main manuscript, we can derive a representative user i's net utility from participating in the community conditional on participation size s:

$$u^{c,in-out}_{i|s} = \frac{\gamma(1+\psi)(v-\tilde{\tau})}{2c}[v - \tilde{\tau} - 2(1-\epsilon)s\lambda_i] + \tilde{\tau}\tilde{\Lambda}. \tag{A.16}$$

We calculate $\partial u^{c,in-out}_{i|s}/\partial\tilde{\tau} = \frac{\gamma(1+\psi)}{c}[(1-\epsilon)s\lambda_i + \tilde{\tau} - v] + \tilde{\Lambda}$ for later use in the proof. We next prove that the composite policy cannot improve the participation rate more than a pure quota.

(i) When $0 < v \le (1-\epsilon)(1-\alpha)\lambda_L$, obviously, $\partial u^{c,in-out}_{i|s}/\partial\tilde{\tau} > 0$. Note that $\tilde{\tau} < v - \frac{c\tilde{\Lambda}}{\gamma(1+\psi)}$, so $u^{c,in-out}_{i|s} < u^{c,in-out}_{i|s,\tau=v-\frac{c\tilde{\Lambda}}{\gamma(1+\psi)}} = \tilde{\Lambda}\left[v - \frac{c\tilde{\Lambda}}{2\gamma(1+\psi)} - (1-\epsilon)s\lambda_i\right]$. The last term is user i's net utility in the pure quota. So each user would obtain less net utility from participating in the community under this composite policy than under the pure quota and thus this composite policy cannot improve the participation rate more than a pure quota.

(ii) When $(1-\epsilon)(1-\alpha)\lambda_L < v \leq (1-\epsilon)(1-\alpha\beta)\lambda_L$, by Proposition 3, the best a quota can do is to attract a fraction of uncommitted low-type users to participate and $u_{L|s}^{q,in-out} = 0 \Leftrightarrow v = \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)s\lambda_L$, which implies $\partial u_{L|s}^{c,in-out}/\partial\tilde{\tau} = \frac{\gamma(1+\psi)\tilde{\tau}}{c} + \frac{\tilde{\Lambda}}{2} > 0$. Logic similar to that in (i) yields that low-type users obtain less net utility from participating in the community in this composite policy than in the pure quota. However, for the composite policy to improve the participation rate more than a pure quota, it should make low-type users obtain at least the same net utilitiy, which is a contradiction. So the composite policy cannot improve the participation rate more than a pure quota.

(iii) When $(1-\epsilon)(1-\alpha\beta)\lambda_L < v \leq (1-\epsilon)(1-\alpha\beta)\lambda_H$, by Proposition 3, the best a quota can do is to attract all the uncommitted low-type users. Obviously, $\partial u_{H|s}^{c,in-out}/\partial\tilde{\tau} > 0$. Logic as in (i) yields that high-type users obtain less net utility from participating in the community under this composite policy than under the pure quota. However, for the composite policy to improve the participation rate more than a pure quota, it should make high-type users obtain more net utility, which is a contradiction. So the composite policy cannot improve the participation rate more than a pure quota.

(iv) When $(1-\epsilon)(1-\alpha\beta)\lambda_H < v \leq (1-\epsilon)\lambda_H$, by Proposition 3, the best a quota can do is to attract all the uncommitted low-type users and a fraction of uncommitted high-type users. Moreover, $u_{H|s}^{q,in-out} = 0 \Leftrightarrow v = \frac{c\Lambda}{2\gamma(1+\psi)} + (1-\epsilon)s\lambda_H$, which implies $\partial u_{H|s}^{c,in-out}/\partial\tilde{\tau} = \frac{\gamma(1+\psi)\tilde{\tau}}{c} + \frac{\tilde{\Lambda}}{2} > 0$. Logic as in (i) yields that high-type users obtain less net utility from participating in the community under this composite policy than under the pure quota. But for the composite policy to improve the participation rate more than a pure quota, it should at least make high-type users obtain the same net utility, which is a contradiction. So the composite policy cannot improve the participation rate more than a pure quota.

(v) When $v > (1-\epsilon)\lambda_H$, by Proposition 3, the best a quota can do is to attract all the uncommitted users. So there is no way for the composite policy to improve the participation rate more than the pure quota.

In summary, the composite policy cannot improve the participation rate better than the pure quota.

Now consider welfare. Given that the composite policy cannot improve the participation rate more than the pure quota, we just need to see if the composite policy can improve the welfare more than the pure quota when it produces a lower participation rate than the pure quota. Hhis is impossible because the welfare-optimal quota identified in Proposition 5, by construction, considers all the possible participation rates and then selects the participation rate that can produce the optimal welfare. In other words, if there indeed exists a lower participation rate that allows the composite policy to outperform the pure quota, the pure quota can also achieve it without the additional nudging and thus deliver a greater welfare.

We can also prove such composite policy will not increase the total quantity of information posted in the community when compared with the status quo. To do so, we only need to see, when users post in excess of the free allowance (i.e., $\frac{\gamma(1+\psi)(v-\tilde{\tau})}{c} > \tilde{\Lambda}$), whether the composite policy will increase total information relative to the status quo.

User i's net utility in (A.16) can be written as

$$u_{i|s}^{c,in-out} = \underbrace{\frac{\gamma(1+\psi)(v-\tilde{\tau})^2}{2c} + \tilde{\tau}\tilde{\Lambda}}_{i} - \underbrace{(1-\epsilon)Q^c(s)\lambda_i}_{ii}. \tag{A.17}$$

Part (i) in (A.17) is less than $\frac{\gamma(1+\psi)v^2}{c}$ due to $\tilde{\Lambda} < \frac{\gamma(1+\psi)(v-\tilde{\tau})}{c}$, implying each user still enjoys less net utility in this composite policy compared with in the status quo. Part (ii) in (A.17) is the net privacy cost user i suffers ($Q^c(s)$ is the total amount of information posted in the community in this composite policy). So the same contradiction seen in the Proof of Proposition 4 will occur here. Therefore, this composite policy cannot increase total information relative to the status quo either.

Proof of Proposition 8. Proposition 8 is evident from the discussion that accompanies it in the main manuscript.

2. Detailed Results in the Numerical Example and Extensions

2.1. Numerical Example

Here we explain how we derive users' posting quantities in the targeted nudge and quota mechanisms when "false positives" and "false negatives" are allowed. Denote the probability of "false positive" by p_x and the probability of "false negative" by p_y. Note that only detected "sensitive information" would be regulated by nudge or quota.

In targeted nudging, user i's net utility is

$$\begin{aligned} u_{i|s}^{tn,in-out} =& n\left[vx_{ij} - \frac{cx_{ij}^2}{2} + vy_{ij} - \frac{cy_{ij}^2}{2\psi} - \tau(p_x x_{ij} + (1-p_y)y_{ij})\right] + (1-n)\left[vx_{ik} - \frac{cx_{ik}^2}{2\delta} + vy_{ik} - \frac{cy_{ik}^2}{2\delta\psi}\right. \\ & \left. - \tau(p_x x_{ik} + (1-p_y)y_{ik})\right] + (1-\epsilon)(eQ_{-i} + \omega X_{\cdot i} - \theta_i Y_{\cdot i}). \end{aligned}$$

FOCs w.r.t this equation would give user i's optimal posting quantities under targeted nudging:

$$x_{ij}^{tn} = \frac{v - p_x\tau}{c},\ y_{ij}^{tn} = \frac{\psi[v-(1-p_y)\tau]}{c},\ x_{ik}^{tn} = \frac{\delta(v-p_x\tau)}{c},\ y_{ik}^{tn} = \frac{\psi\delta[v-(1-p_y)\tau]}{c}.$$

Under a targeted quota, user i's net utility is

$$\begin{aligned} u_{i|s}^{tq,in-out} =& n\left[vx_{ij} - \frac{cx_{ij}^2}{2} + vy_{ij} - \frac{cy_{ij}^2}{2\psi}\right] + (1-n)\left[vx_{ik} - \frac{cx_{ik}^2}{2\delta} + vy_{ik} - \frac{cy_{ik}^2}{2\delta\psi}\right] \\ & + (1-\epsilon)(eQ_{-i} + \omega X_{\cdot i} - \theta_i Y_{\cdot i}), \\ & \text{with the constraint } n[p_x x_{ij} + (1-p_y)y_{ij}] + (1-n)[p_x x_{ik} + (1-p_y)y_{ik}] = \Lambda \end{aligned}$$

As in the proof of Proposition 3, we can define a Lagrange function to derive user i's optimal posting quantities under a targeted quota,

$$x_{ij}^{tq} = \left[1 - \frac{p_x + \psi(1-p_y)}{p_x + \psi(1-p_y)^2/p_x}\right]\frac{v}{c} + \frac{\Lambda}{[p_x + \psi(1-p_y)^2/p_x]\gamma},$$

$$y_{ij}^{tq} = \left[1 - \frac{p_x + \psi(1-p_y)}{p_x^2/(1-p_y) + \psi(1-p_y)}\right]\frac{\psi v}{c} + \frac{\psi\Lambda}{[p_x^2/(1-p_y) + \psi(1-p_y)]\gamma},$$

$$x_{ik}^{tq} = \left[1 - \frac{p_x + \psi(1-p_y)}{p_x + \psi(1-p_y)^2/p_x}\right]\frac{\delta v}{c} + \frac{\delta\Lambda}{[p_x + \psi(1-p_y)^2/p_x]\gamma},$$

$$y_{ik}^{tq} = \left[1 - \frac{p_x + \psi(1-p_y)}{p_x^2/(1-p_y) + \psi(1-p_y)}\right]\frac{\delta\psi v}{c} + \frac{\delta\psi\Lambda}{[p_x^2/(1-p_y) + \psi(1-p_y)]\gamma}.$$

In the numerical example, we set $p_x = p_y = 0.1$ and fix all other exogenous parameters, including the values for τ under a targeted nudge and Λ under a targeted quota (see specific values in the main text). We choose six different values of v. For each v, using all the fixed parameters, we can compute each user's optimal posting quantities according to the equations derived above. Then we substitute the optimal posting quantities back into each user's net utility function to determine their participation decisions. Finally, all the aggregate measures (i.e., total quantity of information posted, social welfare, and total privacy damage) can be calculated once participation rate is determined.

2.2. Extensions

(i) Heterogeneity in friendship. We assume $n_i \in U[0,1]$ and $\theta_i = 1 - n_i$, then $\lambda_i = \frac{\psi\theta_i - e(1+\psi) - \omega}{1+\psi}$, $\gamma_i = n_i + (1-n_i)\delta = 1 - (1-\delta)\theta_i$. User i's net utility from participating in the community in the status quo (conditional on the total information posted in the community, Q^{sq}) is given by

$$u_i^{sq,in-out} = \frac{\gamma_i(1+\psi)v^2}{2c} - (1-\epsilon)Q^{sq}\lambda_i. \tag{A.18}$$

Note that $u_i^{sq,in-out}$ in equation (A.18) decreases in θ_i because γ_i decreases in θ_i and λ_i increases with θ_i. So less privacy-sensitive (more connected) users are more likely to participate. That means we can characterize the equilibrium participation rate by a cutoff user type (θ^o) such that users with privacy sensitivity below θ^o will participate in the community and the others will not. So θ^o indicates the equilibrium participation rate in the community.

Note also that the utility function in equation (A.18) increases with v. As v becomes sufficiently large, every user will get positive net utility and thus participate in the community, which implies $\theta^o = 1$. Now focus on $\theta^o < 1$, then the condition to solve for θ^o is that the cutoff type has a net utility of zero:

$$u_i^{sq,in-out}(\theta^o) = \frac{\gamma_i(\theta^o)(1+\psi)v^2}{2c} - (1-\epsilon)Q^{sq}(\theta^o)\lambda_i(\theta^o) = 0. \tag{A.19}$$

The total quantity of information posted in the community in equilibrium conditional on θ^o is given by

$$Q^{sq}(\theta^o) = \int_0^{\theta^o} \frac{\gamma_i(\theta_i)(1+\psi)v}{c} d\theta_i = \theta^o[1 - \frac{1-\delta}{2}\theta^o]\frac{(1+\psi)v}{c}. \tag{A.20}$$

Combining equations (A.19) and (A.20), we can derive the condition to solve for θ^o:

$$v = \frac{(1-\epsilon)[\psi\theta^o - e(1+\psi) - \omega]\theta^o[2 - (1-\delta)\theta^o]}{(1+\psi)[1-(1-\delta)\theta^o]}. \tag{A.21}$$

The RHS of equation (A.21) is a strictly increasing function of θ^o because $\psi\theta^o - e(1+\psi) - \omega > 0$ and increases with θ^o, $\theta^o[2-(1-\delta)\theta^o] > 0$ and increases with θ^o, and $1-(1-\delta)\theta^o > 0$ and decreases with θ^o. So as v increases, θ^o will increase. But note the condition of $\theta^o < 1$, which requires $v < (1-\epsilon)(1+\frac{1}{\delta})\frac{\psi-e(1+\psi)-\omega}{1+\psi}$. When $v \geq (1-\epsilon)(1+\frac{1}{\delta})\frac{\psi-e(1+\psi)-\omega}{1+\psi}$, all users will get positive net utility and participate in the community. So we obtain Lemma 4.

Nudge

Note that the impact of a nudge relative to the status quo is decreasing v. So according to Lemma 4, implementing a nudge will (weakly) decrease the participation rate relative to the status quo. According to equation (A.20), it is easy to verify that $Q^{sq}(\theta^o)$ increases with v and θ^o. That is, the total quantity of information posted in the community will decrease as posting benefit decreases and participation rate decreases, which are the impacts from a nudge. To examine the impact of a nudge on social welfare, we first derive the social welfare in the status quo:

$$\begin{aligned}\Pi^{sq} &= \int_0^{\theta^o}\left[\frac{\gamma_i(\theta_i)(1+\psi)v^2}{2c} - Q^{sq}(\theta^o)\lambda_i(\theta_i)\right]d\theta_i + \int_{\theta^o}^1[-\epsilon Q^{sq}(\theta^o)\lambda_i(\theta_i)]d\theta_i \\ &= \frac{(1+\psi)v}{c}\theta^o[1-\frac{1-\delta}{2}\theta^o]\left\{\left[\frac{v}{2} - \frac{\psi(\theta^o)^2}{2(1+\psi)} + \left(e+\frac{\omega}{1+\psi}\right)\theta^o\right] - \right. \\ &\left.\epsilon(1-\theta^o)\left[\frac{\psi(1+\theta^o)}{2(1+\psi)} - \left(e+\frac{\omega}{1+\psi}\right)\right]\right\}.\end{aligned} \tag{A.22}$$

Note the equality between θ^o and v in equation (A.21). So Π^{sq} can be viewed as a function of θ^o–the equilibrium participation rate. The function in equation (A.22) is obviously a highly nonmonotonic function. Figure A.1 illustrates this via an example with $\psi = 0.5$, $\epsilon = 0.01$, $e = 0.01$, $\omega = 0.01$, $\delta = 0.5$, and $c = 1$ (similar to Figure 3 in the main text). As participation increases, the social welfare may decrease. Therefore, we have shown numerically that, when a nudge is implemented, the participation rate decreases, which may increase social welfare. This result shows that the effects of a nudge are similar to that in the main model.

Quota

We consider a quota of the following form: $\Lambda_i = f \cdot \frac{\gamma_i(1+\psi)v}{c}$, where $f \in [0,1]$ and $\frac{\gamma_i(1+\psi)v}{c}$ is the total amount of information user i will post by her own choice in the status quo. User i's net utility from participating in the community under the quota (conditional on the total information posted in the community Q^q) is given by

$$u_i^{q,in-out} = v\Lambda_i - \frac{c\Lambda_i^2}{2\gamma_i(1+\psi)} - (1-\epsilon)Q^q\lambda_i. \tag{A.23}$$

$u_i^{q,in-out}$ in equation (A.23) decreases in θ_i because $v\Lambda_i - \frac{c\Lambda_i^2}{2\gamma_i(1+\psi)}$ decreases in θ_i (this in turn is because $v\Lambda_i - \frac{c\Lambda_i^2}{2\gamma_i(1+\psi)}$ increases with Λ_i and Λ_i decreases with θ_i) and λ_i increases with θ_i. So we can

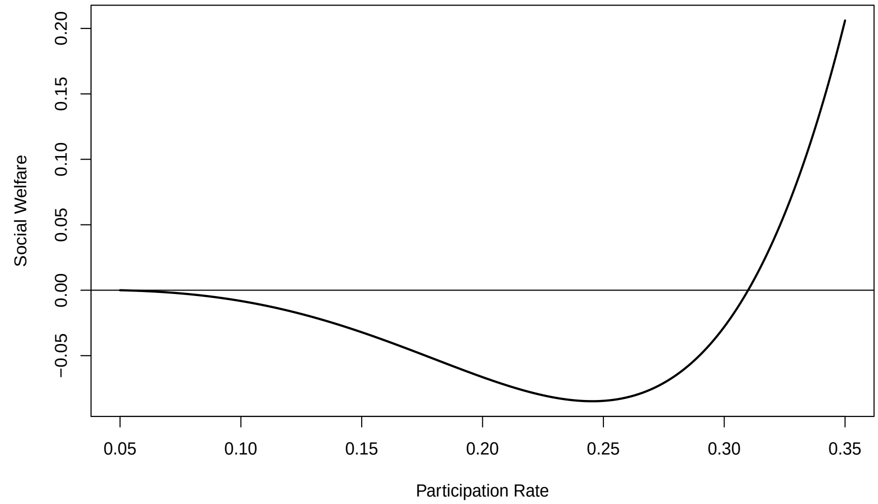

Figure A.1: Social welfare as a function of participation rate in the status quo

again characterize the equilibrium participation rate by a cutoff user type (θ^o) such that users with privacy sensitivity below θ^o will participate in the community and the others will not.

Focus on $\theta^o < 1$. The condition to solve for θ^o is that the cutoff type has a net utility of zero:

$$u_i^{q,in-out}(\theta^o) = v\Lambda_i(\theta^o) - \frac{c\Lambda_i(\theta^o)^2}{2\gamma_i(\theta^o)(1+\psi)} - (1-\epsilon)Q^q(\theta^o)\lambda_i(\theta^o) = 0. \tag{A.24}$$

The total quantity of information posted in the community in equilibrium conditional on θ^o is given by

$$Q^q(\theta^o) = \int_0^{\theta^o} f\frac{\gamma_i(\theta_i)(1+\psi)v}{c}d\theta_i = f\theta^o[1 - \frac{1-\delta}{2}\theta^o]\frac{(1+\psi)v}{c}. \tag{A.25}$$

Combining equations (A.24) and (A.25), we can derive the condition to solve for θ^o:

$$(2-f)v = \frac{(1-\epsilon)[\psi\theta^o - e(1+\psi) - \omega]\theta^o[2-(1-\delta)\theta^o]}{(1+\psi)[1-(1-\delta)\theta^o]}. \tag{A.26}$$

Now comparing equations (A.26) and (A.21), the RHSs of both equations are the same and increase with θ^o. Note that $(2-f)v \geq v$ (LHSs of the two equations), so the equilibrium θ^o under a quota will be larger than that in the status quo. In other words, adding a quota to the status quo will increase the participation rate. This holds as long as the participation rate in the status quo is not full and there is room for improvement (i.e., $v < (1-\epsilon)(1+\frac{1}{\delta})\frac{\psi - e(1+\psi)-\omega}{1+\psi}$). The impact of a quota on the total amount of information posted in the community can still be seen by the dilemma highlighted in Proposition 8. As in the main model, a quota will decrease the total amount of information and hence the total privacy damage relative to the status quo.

The social welfare under a quota is given by

$$\begin{aligned}\Pi^q &= \int_0^{\theta^o} \left[(2f - f^2)\frac{\gamma_i(\theta_i)(1+\psi)v^2}{2c} - Q^q(\theta^o)\lambda_i(\theta_i)\right] d\theta_i + \int_{\theta^o}^1 [-\epsilon Q^q(\theta^o)\lambda_i(\theta_i)]d\theta_i \\ &= f \cdot \frac{(1+\psi)v}{c} \cdot \theta^o \left[1 - \frac{1-\delta}{2}\theta^o\right] \left\{\left[\left(1 - \frac{f}{2}\right) v - \frac{\psi(\theta^o)^2}{2(1+\psi)} + \left(e + \frac{\omega}{1+\psi}\right)\theta^o\right] - \right. \\ &\left. \epsilon(1-\theta^o)\left[\frac{\psi(1+\theta^o)}{2(1+\psi)} - \left(e + \frac{\omega}{1+\psi}\right)\right]\right\}. \end{aligned} \tag{A.27}$$

Due to the complexity of the model, it is challenging to derive the closed-form optimal quota by comparing Π^q with Π^{sq}. We thus resort to numerical analysis to examine the impact of a quota. We set $f = 0.7$ and use the same set of parameter values as in Figure A.1. Figure A.2 shows that a quota exists that increases the social welfare relative to the status quo, which is consistent with Proposition 5 and Figure 4 in the main text. Note that, we also vary the values of the key parameters – ψ and δ from 0.1 to 0.9, and solve the numerical example under the different values. Under all these cases, we discover the consistent result that, with an appropriate quota, the social welfare can be increased relative to the status quo.

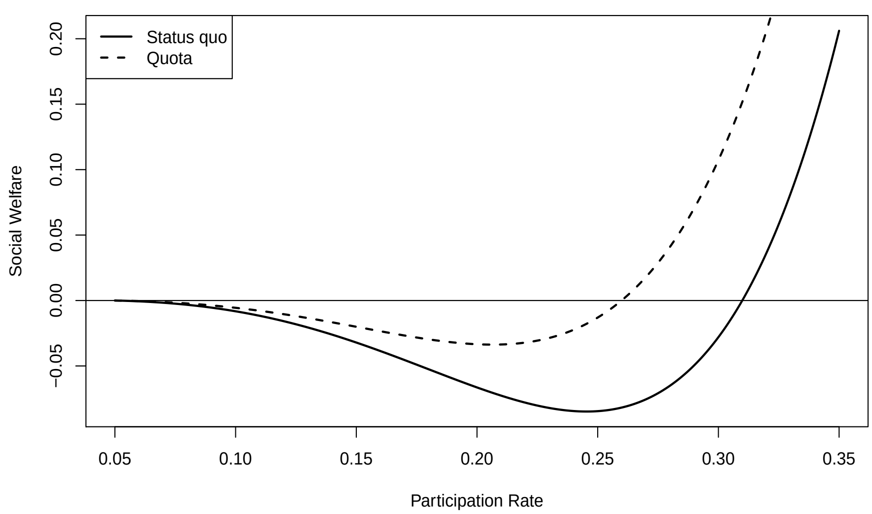

Figure A.2: Social welfare as a function of participation rate under the status quo and a quota

(ii) Nonlinear externality. Following the same approach in the main model, we can derive the equilibrium outcomes in the status quo as follows.

(1) When $0 < v < 2(1-\epsilon)(1-\alpha)^2\lambda_L$, no uncommitted users will participate. $s^{sq} = 1-\alpha$. $Q^{sq} = \frac{(1-\alpha)\gamma(1+\psi)v}{c}$. $\Pi^{sq} = \frac{(1-\alpha)\gamma(1+\psi)v}{2c}[v - 2(1-\alpha)(1-\alpha+\epsilon\alpha)\bar{\lambda}]$. $\xi^{sq} = \frac{(1-\alpha)^2\gamma\psi\bar{\theta}v}{c}$.

(2) When $2(1-\epsilon)(1-\alpha)^2\lambda_L \le v \le 2(1-\epsilon)(1-\alpha\beta)^2\lambda_L$, only a fraction of uncommitted low-type users will participate. $s^{sq} = \sqrt{\frac{v}{2(1-\epsilon)\lambda_L}}$. $Q^{sq} = \frac{\gamma(1+\psi)v}{c}\sqrt{\frac{v}{2(1-\epsilon)\lambda_L}}$. $\Pi^{sq} = \frac{\gamma(1+\psi)v^2}{2c}\left\{(1-\alpha) - \frac{(1-\alpha+\epsilon\alpha)\bar{\lambda}}{(1-\epsilon)\lambda_L}\right\}$. $\xi^{sq} = \frac{\gamma\psi\bar{\theta}v^2}{2c(1-\epsilon)\lambda_L}$

(3) When $2(1-\epsilon)(1-\alpha\beta)^2\lambda_L < v < 2(1-\epsilon)(1-\alpha\beta)^2\lambda_H$, only all uncommitted low-type users will participate. $s^{sq} = 1-\alpha\beta$. $Q^{sq} = \frac{(1-\alpha\beta)\gamma(1+\psi)v}{c}$. $\Pi^{sq} = \frac{(1-\alpha\beta)\gamma(1+\psi)v}{2c}\left\{v - 2(1-\alpha\beta)[\bar{\lambda} - (1-\epsilon)\alpha\beta\lambda_H]\right\}$. $\xi^{sq} = \frac{(1-\alpha\beta)^2\gamma\psi\bar{\theta}v}{c}$.

(4) When $2(1-\epsilon)(1-\alpha\beta)^2\lambda_H \le v \le 2(1-\epsilon)\lambda_H$, all uncommitted low-type users and a fraction of uncommitted high-type users will participate. $s^{sq} = \sqrt{\frac{v}{2(1-\epsilon)\lambda_H}}$. $Q^{sq} = \frac{\gamma(1+\psi)v}{c}\sqrt{\frac{v}{2(1-\epsilon)\lambda_H}}$. $\Pi^{sq} = \frac{\gamma(1+\psi)v^2}{2c}\left[1 - \frac{\bar{\lambda}}{(1-\epsilon)\lambda_H}\right]$. $\xi_{sq} = \frac{\gamma\psi\bar{\theta}v^2}{2c(1-\epsilon)\lambda_H}$.

(5) When $v > 2(1-\epsilon)\lambda_H$, all uncommitted users will participate. $s^{sq} = 1$. $Q^{sq} = \frac{\gamma(1+\psi)v}{c}$. $\Pi^{sq} = \frac{\gamma(1+\psi)v}{2c}(v - 2\bar{\lambda})$. $\xi^{sq} = \frac{\gamma\psi\bar{\theta}v}{c}$.

Compare the cutoff value of v when uncommitted low-type users start to participate here with that in the main model: $2(1-\alpha)^2(1-\epsilon)\lambda_L < 2(1-\alpha)(1-\epsilon)\lambda_L$. Compare the cutoff value of v when uncommitted high-type users start to participate here with that in the main model: $2(1-\alpha\beta)^2(1-\epsilon)\lambda_H < 2(1-\alpha\beta)(1-\epsilon)\lambda_H$. So we obtain Lemma 5.

As we can see, the corresponding impact is that the cutoff values of v which characterize different equilibrium outcomes change quantitatively. But all the results obtained in the main model can be verified to hold qualitatively.

(iii) Unintentional Disclosure of Sensitive Information. Users may not know the information they post is sensitive. We can modify the setup in line with this consideration. Assume users only make decisions about the total amount of information to be posted about friends and nonfriends, denoted by x_{ij} and x_{ik}, respectively. A fraction, $\phi \in (0,1)$, of the information posted by a user will cause privacy harm to others. Users may not know which information belongs to this fraction. All other settings and notations remain the same as in the main model. Then user i's utility from participating in the community conditional on participation size, s, is

$$u_{i|s}^{in} = n\left(vx_{ij} - \frac{cx_{ij}}{2}\right) + (1-n)\left(vx_{ik} - \frac{cx_{ik}^2}{2\delta}\right) + eQ_{-i} + \omega X_{\cdot i} - \theta_i Y_{\cdot i}. \quad \text{(A.28)}$$

Equation (A.28) is the counterpart of equation (2) in the main model. Solving the FOCs yields $x_{ij}^{sq} = \frac{v}{c}$ and $x_{ik}^{sq} = \frac{\delta v}{c}$. Using these equilibrium quantities, we can derive user i's net utility from participating in the community:

$$u_{i|s}^{in-out} = \frac{\gamma v}{2c}[v - 2(1-\epsilon)s\lambda_i],$$

where $\lambda_i = \phi\theta_i - e - (1-\phi)\omega$ has the same interpretation as in the main model. Then all the results in the main model can be easily replicated here. As we can see, the only difference is that $\frac{\psi}{1+\psi}$ in the main model is replaced by ϕ here. That is because we conveniently capture the differences in the quantities of nonsensitive and sensitive information posted by a user through the cost-ratio parameter, ψ, which is also exogenously given. As a result, an exogenous and fixed fraction of the total amount of information posted by a user is privacy infringing. Therefore, $\frac{\psi}{1+\psi}$ is equivalent to ϕ.