MIS Quarterly

RESEARCH ARTICLE

HOW MUCH TO SHARE WITH THIRD PARTIES? USER PRIVACY CONCERNS AND WEBSITE DILEMMAS

Ram D. Gopal
Operations and Information Management, School of Business, University of Connecticut,
Storrs, CT 06269 U.S.A. {ram.gopal@business.uconn.edu}

Hooman Hidaji and Raymond A. Patterson
Haskayne School of Business, University of Calgary, Calgary, AB, CANADA T2N 1N4
{hooman.hidaji@haskayne.ucalgary.ca} {raymond.patterson@ucalgary.ca}

Erik Rolland
College of Business Administration, California State Polytechnic University,
Pomona, CA 91768 U.S.A. {erolland@cpp.edu}

Dmitry Zhdanov
Department of Computer Information Systems, J. Mack Robinson College of Business,
Georgia State University, Atlanta, GA 30302 U.S.A. {dzhdanov@gsu.edu}

Appendix A

Proofs

Proof for Lemma 1

The optimal publisher website royalties $R^*_{W_i}, i = 1,2$, and prices $P^*_{W_i}, i = 1,2$, satisfy the first order conditions:

$$\frac{\partial \Pi_{W_1}}{\partial R_{W_1}}\left(R^*_{W_1}, P^*_{W_1}, R^*_{W_2}, P^*_{W_2}\right) = \frac{\partial \Pi_{W_1}}{\partial P_{W_1}}\left(R^*_{W_1}, P^*_{W_1}, R^*_{W_2}, P^*_{W_2}\right) =$$
$$\frac{\partial \Pi_{W_2}}{\partial R_{W_2}}\left(R^*_{W_1}, P^*_{W_1}, R^*_{W_2}, P^*_{W_2}\right) = \frac{\partial \Pi_{W_2}}{\partial P_{W_2}}\left(R^*_{W_1}, P^*_{W_1}, R^*_{W_2}, P^*_{W_2}\right) = 0 \tag{A1.1}$$

By simultaneously solving these equations, $R^*_{W_1} = R^*_{W_2} = R^*_W$ and $P^*_{W_1} = P^*_{W_2} = P^*_W$ are calculated as given in Lemma 1. To ensure that profit is maximized, the second order conditions must hold:

$$\frac{\partial^2 \Pi_{W_i}}{\partial {P_{W_i}}^2} = -\frac{\Phi M_U(8\Phi t - M_U M_D (R_D - 3v)(R_D - v))}{2(2\Phi t + M_U M_D v (R_D - v))^2} < 0 \tag{A1.2}$$

$$\frac{\partial^2 \Pi_{W_i}}{\partial {R_{W_i}}^2} = -\frac{M_D {M_U}^2(4\Phi t + M_U M_D v\,(R_D - v\,))(4\Phi t + M_U M_D v\,(3R_D - v\,))}{8\Phi(2\Phi t + M_U M_D v\,(R_D - v\,))^2} < 0 \tag{A1.3}$$

$$det(Hessian) = \frac{\partial^2 \Pi_{W_i}}{\partial {R_{W_i}}^2}\frac{\partial^2 \Pi_{W_i}}{\partial {P_{W_i}}^2} - \left(\frac{\partial^2 \Pi_{W_i}}{\partial P_{W_i}\partial R_{W_i}}\right)^2 = \frac{4 M_U {M_U}^3 (2\Phi t + M_U M_D v (R_D - v))^2}{16(2\Phi t + M_U M_D v (R_D - v))^4} \tag{A1.4}$$

We also need the optimal number of users ${N_{U_i}}^* = {N_{U_i}}^*({R_{W_1}}^*, {P_{W_1}}^*)$ and number of third parties ${N_{D_i}}^* = {N_{D_i}}^*({R_{W_1}}^*, {P_{W_1}}^*)$ to be positive. So we need to have

$$N_{U_i} = M_U \frac{\Phi t + \Phi(P_{W_{-i}} - P_{W_i}) + M_U M_D v\,(R_D - R_{W_{-i}})}{2\Phi t + M_U M_D v((R_D - R_{W_i}) + (R_D - R_{W_{-i}}))} \geq 0$$

$$\Rightarrow 2\Phi t + M_U M_D v((R_D - R_{W_i}) + (R_D - R_{W_{-i}})) \geq 0 \qquad \forall i = 1,2 \tag{A1.5}$$

$$N_{D_i} = M_D \frac{M_U (R_D - R_{W_i})(\Phi t + \Phi(P_{W_{-i}} - P_{W_i}) + M_U M_D v\,(R_D - R_{W_{-i}}))}{\Phi(2\Phi t + M_U M_D v((R_D - R_{W_i}) + (R_D - R_{W_{-i}})))} \geq 0$$

$$\Rightarrow R_D - R_{W_i} \geq 0 \qquad \forall i = 1,2 \tag{A1.6}$$

Throughout the paper, we assume (A.1.2), (A.1.3), (A.1.5), and (A.1.6) to be true. (A.1.4) is always true. ■

Proposition 1

(i) By taking the derivatives of ${R_W}^*$ with respect to v and R_D we have

$$\frac{\partial\, {R_W}^*}{\partial\, v} = \frac{1}{2} > 0 \tag{A2.1}$$

$$\frac{\partial\, {R_W}^*}{\partial\, R_D} = \frac{1}{2} > 0 \tag{A2.2}$$

It is clear from the formula for ${R_W}^*$ in Lemma 1 that it is independent of the other parameters. ■

(ii) We have $R_D - v > 0$ and

$$\frac{\partial\, {P_W}^*}{\partial\, v} = \frac{M_D M_U (R_D - v)}{2\Phi} > 0 \tag{A2.3}$$

$$\frac{\partial\, {P_W}^*}{\partial t} = 1 > 0 \tag{A2.4}$$

$$\frac{\partial P_W^*}{\partial R_D} = -\frac{M_D M_U (R_D - v)}{2\Phi} < 0 \tag{A2.5}$$

$$\frac{\partial P_W^*}{\partial M_D} = -\frac{M_U (R_D - v)^2}{4\Phi} < 0 \tag{A2.6}$$

$$\frac{\partial P_W^*}{\partial M_U} = -\frac{M_D (R_D - v)^2}{4\Phi} < 0 \tag{A2.7}$$

■

Proposition 2

Using the formula for optimal number of third parties in (13), we have

$$\frac{\partial N_D^*}{\partial v} = -\frac{M_D M_U}{4\Phi} < 0 \tag{A3.1}$$

$$\frac{\partial N_D^*}{\partial \Phi} = -\frac{M_D M_U (R_D - v)}{4\Phi^2} < 0 \tag{A3.2}$$

$$\frac{\partial N_D^*}{\partial R_D} = \frac{M_D M_U}{4\Phi} > 0 \tag{A3.3}$$

$$\frac{\partial N_D^*}{\partial M_D} = \frac{M_U (R_D - v)}{4\Phi} > 0 \tag{A3.4}$$

$$\frac{\partial N_D^*}{\partial M_U} = \frac{M_D (R_D - v)}{4\Phi} > 0 \tag{A3.5}$$

■

Proposition 3

(i) Profit of each publisher website is calculated by substituting the optimal royalties and price equations (10) and (11) from Lemma 1 into the publisher website profit equation, and is given in equation (14) in Proposition 3. We have

$$\frac{\partial \Pi_W^*}{\partial v} = \frac{M_D M_U^2 (2R_D - 3v)}{8\Phi} \tag{A4.1}$$

which is positive when $v < \frac{2}{3} R_D$ and is negative when $\frac{2}{3} R_D < v$. ■

(ii) We calculate the user surplus from each publisher website as follows:

$$Z_{U_1} = \int_0^{(t+v(N_{D_2}-N_{D_1})+P_{W_2}-P_{W_1})/2t} (X - N_{D_1}\, v - P_{W_1} - ty)\, dy$$

$$Z_{U_2} = \int_{(t+v(N_{D_2}-N_{D_1})+P_{W_2}-P_{W_1})/2t}^{1} (X - N_{D_2}\, v - P_{W_2} - t(1-y))\, dy$$

$$Z_U = Z_{U_1} + Z_{U_2}$$

Solving the equation by substituting the optimal publisher website royalties and prices, we have

$$Z_U{}^* = \frac{M_U M_D (R_D - 2v)(R_D - v) + \Phi(4X - 5t)}{4\,\Phi} \tag{A4.2}$$

Taking the derivative with respect to v we have

$$\frac{\partial\, Z_U{}^*}{\partial\, v} = -\frac{M_U M_D (3R_D - 4v)}{4\,\Phi} \tag{A4.3}$$

which is negative when $v < \frac{3}{4} R_D$ and is positive when $\frac{3}{4} R_D < v$. ■

(iii) Here we calculate the third party surplus from each publisher website as follows:

$$Z_{D_1} = \int_0^{N_{U_1}(R_D - R_{W_1})} (N_{U_1}(R_D - R_{W_1}) - \varphi)\, d\varphi = \frac{1}{2} N_{U_1}{}^2 (R_D - R_{W_1})^2$$

$$Z_{D_2} = \int_0^{N_{U_2}(R_D - R_{W_2})} (N_{U_2}(R_D - R_{W_2}) - \varphi)\, d\varphi = \frac{1}{2} N_{U_2}{}^2 (R_D - R_{W_2})^2$$

$$Z_D = Z_{D_1} + Z_{D_2}$$

Solving the equation by substituting the optimal publisher website royalties and prices, we have

$$Z_D{}^* = \frac{1}{16} M_U{}^2 (R_D - v)^2 \tag{A4.4}$$

Taking the derivatives we have

$$\frac{\partial\, Z_D{}^*}{\partial\, v} = -\frac{1}{8} M_U{}^2 (R_D - v) < 0 \tag{A4.5}$$

Which is always negative. ■

Proposition 4

$$\frac{\partial R_W{}^*}{\partial T_{R_W}} = -\frac{R_D}{2(1+T_{R_W})^2} < 0 \tag{A5.1}$$

$$\frac{\partial R_W{}^*}{\partial T_{P_W}} = -\frac{v}{2(1+T_{P_W})^2} < 0 \tag{A5.2}$$

$$\begin{aligned}\frac{\partial P_W{}^*}{\partial T_{R_W}} &= \frac{M_U M_D (R_D{}^2(1+T_{P_W})^2 - v^2(1+T_{R_W})^2)}{4\Phi(1+T_{P_W})^2(1+T_{R_W})^2} \\ &= \frac{M_U M_D (R_D(1+T_{P_W}) + v(1+T_{R_W}))(R_D(1+T_{P_W}) - v(1+T_{R_W}))}{4\Phi(1+T_{P_W})^2(1+T_{R_W})^2} > 0\end{aligned} \tag{A5.3}$$

$$\frac{\partial P_W{}^*}{\partial T_{P_W}} = -\frac{2\Phi t(1+T_{P_W}) + M_U M_D v(R_D(1+T_{P_W}) - v(1+T_{R_W}))}{2\Phi(1+T_{P_W})^2} < 0 \tag{A5.4}$$

■

Propositions 5 and 6

(i) Using the transformations (17) and (18), the optimal profit for the website is calculated as

$$\Pi_W{}^* = \frac{M_U(8\,\Phi\,t\,(1+T_{P_W})(1+T_{R_W}) - M_U M_D(R_D(1+T_{P_W}) - 3v(1+T_{R_W}))(R_D(1+T_{P_W}) - v(1+T_{R_W})))}{16\,\Phi(1+T_{P_W})^2(1+T_{R_W})} \tag{A6.1}$$

Taking the derivative of profit with respect to the taxations we have

$$\frac{\partial \Pi_{W_i}{}^*}{\partial T_{R_W}} = \frac{M_U{}^2 M_D (R_D{}^2(1+T_{P_W})^2 - 3v^2(1+T_{R_W})^2)}{16\,\Phi(1+T_{P_W})^2(1+T_{R_W})^2} \tag{A6.2}$$

which is positive when $v < \frac{R_D}{\sqrt{3}}\frac{1+T_{P_W}}{1+T_{R_W}}$, and is negative when $\frac{R_D}{\sqrt{3}}\frac{1+T_{P_W}}{1+T_{R_W}} < v$.

$$\frac{\partial \Pi_{W_i}{}^*}{\partial T_{P_W}} = -\frac{M_U(4\Phi t(1+T_{P_W}) + M_U M_D v(2R_D(1+T_{P_W}) - 3v(1+T_{R_W}))}{8\Phi(1+T_{P_W})^3} < 0 \tag{A6.3}$$

(ii) User surplus when taxations are possible is calculated as

$$Z_U{}^* = \frac{\Phi(4X-5t)(1+T_{P_W})(1+T_{R_W}) + M_U M_D(R_D(1+T_{P_W}) - 2v(1+T_{R_W}))(R_D(1+T_{P_W}) - v(1+T_{R_W}))}{4\,\Phi(1+T_{P_W})(1+T_{R_W})} \tag{A6.4}$$

and we have

$$\frac{\partial Z_U{}^*}{\partial T_{R_W}} = -\frac{M_D M_U (R_D{}^2(1+T_{P_W})^2 - 2v^2(1+T_{R_W})^2)}{4\,\Phi(1+T_{P_W})\,(1+T_{R_W})^2} \tag{A6.5}$$

which is negative when $v < \frac{R_D}{\sqrt{2}}\frac{1+T_{P_W}}{1+T_{R_W}}$ and is positive when $\frac{R_D}{\sqrt{2}}\frac{1+T_{P_W}}{1+T_{R_W}} < v$.

$$\frac{\partial Z_U{}^*}{\partial T_{P_W}} = \frac{M_D M_U (R_D{}^2(1+T_{P_W})^2 - 2v^2(1+T_{R_W})^2)}{4\,\Phi\,(1+T_{P_W})^2\,(1+T_{R_W})} \tag{A6.6}$$

which is positive when $v < \frac{R_D}{\sqrt{2}}\frac{1+T_{P_W}}{1+T_{R_W}}$ and is negative when $\frac{R_D}{\sqrt{2}}\frac{1+T_{P_W}}{1+T_{R_W}} < v$.

(iii) Third party surplus when taxations are present is calculated as

$$Z_D{}^* = \frac{M_U{}^2}{16}(R_D(1+T_{P_W}) - v(1+T_{R_W}))^2 \tag{A6.7}$$

and we have

$$\frac{\partial Z_D{}^*}{\partial T_{R_W}} = -\frac{M_U{}^2\, v\,(R_D(1+T_{P_W}) - v(1+T_{R_W}))}{8\,(1+T_{P_W})^2} < 0 \tag{A6.8}$$

$$\frac{\partial Z_D{}^*}{\partial T_{P_W}} = \frac{M_U{}^2(1+T_{R_W})\,v\,(R_D(1+T_{P_W}) - v(1+T_{R_W}))}{8\,(1+T_{P_W})^3} > 0 \tag{A6.9}$$

■

Appendix B

Extension of Proposition 3

In Proposition 3 in the paper, we analyzed the effect of privacy concerns on publisher website profit, third party surplus, and user surplus. Here, we expand the analysis to consider other model parameters. Propositions B.1, B.2, and B.3 provide these results. We do not provide the proofs for these propositions as they are straightforward and can be calculated by taking the derivatives for equations (14), (15), and (16) for optimal website profit, user surplus, and third party surplus, respectively.

Proposition B.1: Effect of Parameters on Publisher Website Profit

(i) *When* $v < \frac{1}{2} R_D$ *profit of each publisher website* (Π_W^*) *decreases with third party revenue from user information* (R_D) *and when* $\frac{1}{2} R_D < v < R_D$ *it increases with third party revenue from user information* (R_D).

(ii) *When* $v < \frac{1}{3} R_D$ *profit of each publisher website* (Π_W^*) *increases with third party costs* (Φ) *and when* $\frac{1}{3} R_D < v < R_D$ *it decreases with third party costs* (Φ).

(iii) *Profit of each publisher website* (Π_W^*) *increases with differentiation between two publisher websites* (t).

(iv) *Profit of each publisher website* (Π_W^*) *increases with total number of potential users in the market* (M_U).

(v) *When* $v < \frac{1}{3} R_D$ *profit of each publisher website* (Π_W^*) *decreases with total number of potential third parties in the market* (M_D) *and when* $\frac{1}{3} R_D < v < R_d$ *it increases with total number of potential third parties in the market* (M_D).

Figure B1 summarizes Proposition B.1.

Proposition B.2: Effect of Parameters on User Surplus

(i) *When* $v < \frac{2}{3} R_D$ *user surplus* (Z_U^*) *increases with third party revenue from user information* (R_D) *and when* $\frac{2}{3} R_D < v < R_D$ *it decreases with third party revenue from user information* (R_D).

(ii) *When* $v < \frac{1}{2} R_D$ *user surplus* (Z_U^*) *decreases with third party fixed costs* (Φ) *and when* $\frac{1}{2} R_D < v < R_D$ *it increases with third party fixed costs* (Φ).

(iii) *User surplus* (Z_U^*) *decreases with publisher website differentiation* (t).

(iv) *When* $v < \frac{1}{2} R_D$ *user surplus* (Z_U^*) *increases with total number of users in the market* (M_U) *and total number of third parties in the market* (M_D), *and when* $\frac{1}{2} R_D < v < R_D$ *it decreases with total number of users in the market* (M_U) *and total number of third parties in the market* (M_D).

Figure B2 summarizes Proposition B.2.

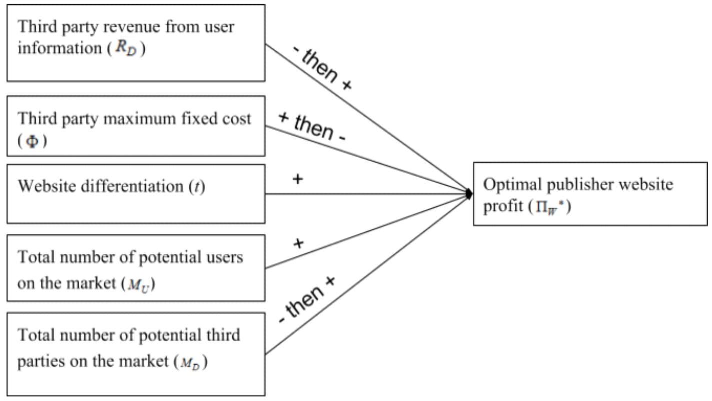

Figure B1. Effect of Model Parameters on Publisher Website Profit

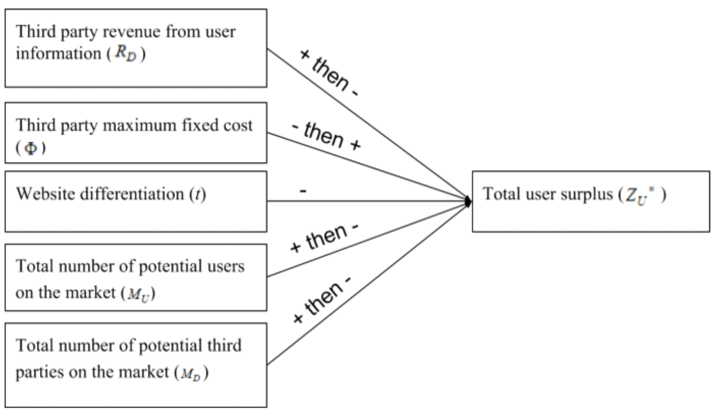

Figure B2. Effect of Model Parameters on User Surplus

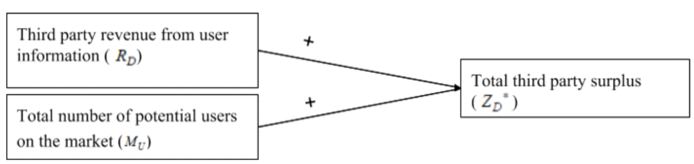

Figure B3. Effect of Model Parameters on Third Party Surplus

Proposition B.3: Effect of Parameters on Third Parties

(i) Third party surplus (Z_D^*) increases with third party revenue from user information (R_D)

(ii) Third party surplus (Z_D^*) increases with total number of users in the market (M_U)

Figure B3 summarizes Proposition B.3.

Appendix C

Asymmetric Model

In the asymmetric model, the two firms are asymmetric in terms of user privacy concerns. The user utility for websites in this case is as follows:

$$U_1(y) = u_1 - ty \qquad u_1 = X - N_{D_1} v_1 - P_{W_1} \tag{C1}$$

$$U_2(y) = u_2 - t(1-y) \qquad u_2 = X - N_{D_2} v_2 - P_{W_2} \tag{C2}$$

The user who is indifferent between websites 1 and 2 is calculated as

$$u_1 - t\hat{y} = u_2 - t(1-\hat{y}) \Rightarrow \hat{y} = \frac{t + \left(N_{D_2} v_2 - N_{D_1} v_1\right) + \left(P_{W_2} - P_{W_1}\right)}{2t} \tag{C3}$$

and the number of users for each publisher website is calculated as

$$N_{U_i} = M_U \frac{t + \left(N_{D_{-i}} - v_{-i} - N_{D_i} v_i\right) + \left(P_{W_{-i}} - P_{W_i}\right)}{2t} > 0 \tag{C4}$$

The third party profit and website profit equations as well as the equation for number of third parties in this case are similar to the base model. The number of users and third parties with respect to the parameters are calculated as

$$N_{U_i} = M_U \frac{\Phi t + \Phi\left(P_{W_{-i}} - P_{W_i}\right) + M_U M_D v_{-i}\left(R_D - R_{W_{-i}}\right)}{2\Phi t + M_U M_D\left(\left(R_D - R_{W_i}\right)v_i + \left(R_D - R_{W_{-i}}\right)v_{-i}\right)} \tag{C5}$$

$$N_{D_i} = M_D \frac{M_U\left(R_D - R_{W_i}\right)\left(\Phi t + \Phi\left(P_{W_{-i}} - P_{W_i}\right) + M_U M_D v_{-i}\left(R_D - R_{W_{-i}}\right)\right)}{\Phi\left(2\Phi t + M_U M_D\left(\left(R_D - R_{W_i}\right)v_i + \left(R_D - R_{W_{-i}}\right)v_{-i}\right)\right)} \tag{C6}$$

Using these equations along with the website profit function, the optimal royalties and prices of the websites can be calculated as follows:

$$R_{W_i}^* = \frac{R_D + v_i}{2} \tag{C7}$$

$$P_{W_i}^* = \frac{\left(2\Phi t + 2\Phi P_{W_{-i}} + M_D M_U v_{-i}\left(R_D - v_{-i}\right)\right)\left(4\Phi t + M_D M_U\left(R_D\left(v_i + v_{-i}\right) - v_{-i}^2 - R_D^2\right)\right)}{2\Phi\left(8\Phi t + M_D M_U\left(2R_D\left(v_i + v_{-i}\right) - 2v_{-i}^2 - v_i^2 - R_D^2\right)\right)} \geq 0 \tag{C8}$$

Note that the website prices are calculated based on the price from the other website, and the equilibrium price in analytically intractable when $P_{W_i}^*$ can differ from $P_{W_{-i}}^*$. It is clear from (C8) that the publisher website *i*'s price is nonlinear in v_i and v_{-i}. The results from the numerical analysis are provided in the body of the paper.

Appendix D

Effect of Privacy Concerns on Market Concentration

In studying the third party market concentration, we consider two cases: third parties with homogenous shares of the market and third parties with nonhomogenous shares of the market. We use the Herfindahl-Hirschman Index (HHI) as a recognized measure for market concentration. The HHI is generically calculated as follows:

$$HHI = \sum_{j=1}^{N_D} s_j^2 \tag{D1}$$

where s_j is the market share of j^{th} third party.

Third Parties with Homogenous Market Shares

In the symmetric duopoly model, because the publisher websites set identical royalties and prices, the third parties either participate in both publisher websites, or do not participate at all. In the homogeneous market share case, the total number of third parties on a particular publisher website i is $N_D = N_{D_1} = N_{D_2}$. When all of the third parties have an equal share of the market, the market share of each third party j is simply calculated as $s_j = 1/N_D$. The HHI is then calculated as

$$HHI = \sum_{j=1}^{N_D} (1/N_D)^2 = N_D (1/N_D)^2 = 1/N_D \tag{D2}$$

By inserting the optimal number of third parties from Proposition 2, we have

$$HHI = 1/N_D = 1\Big/\left[M_D \frac{M_U (R_D - v)}{4\Phi} \right] = \frac{4\Phi}{M_D M_U (R_D - v)} \tag{D3}$$

It can be seen that HHI is increasing in v. In other words, the market concentration is increasing in the user privacy concerns.

To include the effect of barriers to entry, we rewrite the total number of potential third parties, M_D to be as M_D/B, where B is the level of barrier. This means that higher barriers will reduce the number of potential third parties. We can rewrite the HHI formula as

$$HHI = 1/N_D = 1\Big/\left[(M_D / B) \frac{M_U (R_D - v)}{4\Phi} \right] = B \frac{4\Phi}{M_D M_U (R_D - v)} \tag{D4}$$

It can be seen that HHI is increasing in the entry barrier level, so the market concentration is increasing in the level of barrier to entry. The level of barrier to entry is higher for third parties that operate in areas with high privacy concerns and high information sensitivity. In practice, privacy is one reason that third parties need to invest more in information technology (IT) security. These IT investments lead to higher sunk cost of entry, and are a major barrier to entry.

Third Parties with Nonhomogeneous Market Shares

While previously we assumed the market shares to be homogenous for all third parties, this is not realistic in most cases. It results in a market concentration measure that is only dependent on the number of third parties utilized by the publisher websites. We now reconsider the asymmetric model described in the "Asymmetry in User Privacy Concerns" section of the paper, where v_1 varies and v_2 is held constant, using the number of third parties for the two publisher websites to calculate the third party market shares.

Let the number of third parties on publisher websites 1 and 2 be N_{D_1} and N_{D_2}, respectively. Note that since the third parties are differentiated only based on their costs, if a third party participates on the publisher website with higher privacy concern (and higher royalty), then it will also participate on the publisher website with lower privacy concerns (and lower royalty). Thus, there are a total of $Max\left\{N_{D_1}, N_{D_2}\right\}$ unique third parties active in the market. Out of these third parties, $Min\left\{N_{D_1}, N_{D_2}\right\}$ of them participate in both publisher websites, and the rest $Max\left\{N_{D_1}, N_{D_2}\right\}$ of them participate in only one publisher website (the one with lower privacy concerns). Let J_1 be the set of third parties who participate in only one publisher website, and J_2 be the set of third parties who participate in both publisher websites, where $J_1 \cap J_2 = \varnothing$. The size of J_1 is $|J_1| = Max\left\{N_{D_1}, N_{D_2}\right\} - Min\left\{N_{D_1}, N_{D_2}\right\}$, the size of J_2 is $|J_2| = Min\left\{N_{D_1}, N_{D_2}\right\}$, and $J_1 \cup J_2$ is the set of all third parties which has a size of $|J_1 \cup J_2| = Max\left\{N_{D_1}, N_{D_2}\right\}$. Third parties that are present on both publisher websites have a market size that is twice as much as those that participate in only one publisher website. For simplicity and without loss of generality, we assume the following market sizes for each third party. The market size of each third party j (q_j) depends on how many publisher websites they serve.

$$q_j = 1 \qquad \forall j \in J_1 \tag{D5}$$

$$q_j = 2 \qquad \forall j \in J_2 \tag{D6}$$

Let S be the total market size, which is calculated as the sum of relative market share for all third parties. We have

$$S = \sum_{j \in J_1, J_2} q_j = \sum_{j \in J_1} 1 + \sum_{j \in J_2} 2 \tag{D7}$$

The market share of each third party (s_j) is calculated as the ratio of their market size to the total market size, that is

$$s_j = \frac{1}{S} \qquad \forall j \in J_1 \tag{D8}$$

$$s_j = \frac{2}{S} \qquad \forall j \in J_2 \tag{D9}$$

Now that the total market size and share of each third party is known, we calculate the HHI as follows:

$$\begin{aligned} HHI &= \sum_{j \in J_1, J_2} s_j^2 = \sum_{j \in J_1} \left(\frac{1}{S}\right)^2 + \sum_{j \in J_2} \left(\frac{2}{S}\right)^2 \\ &= \left(Max\left\{N_{D_1}, N_{D_2}\right\} - Min\left\{N_{D_1}, N_{D_2}\right\}\right)\left(\frac{1}{S}\right)^2 + Min\left\{N_{D_1}, N_{D_2}\right\}\left(\frac{2}{S}\right)^2 \end{aligned} \tag{D10}$$

which can be calculated as

$$HHI = \frac{Max\left\{N_{D_1}, N_{D_2}\right\} + 3\,Min\left\{N_{D_1}, N_{D_2}\right\}}{\left(Max\left\{N_{D_1}, N_{D_2}\right\} + Min\left\{N_{D_1}, N_{D_2}\right\}\right)^2} \tag{D11}$$

Without loss of generality, let's assume that $N_{D_2} > N_{D_1}$. It can be shown that HHI will be maximized when $N_{D_1}^{Max} = \frac{N_{D_2}}{3}$. Figure D1 provides the effect of change in number of third parties on HHI values for a numerical example.

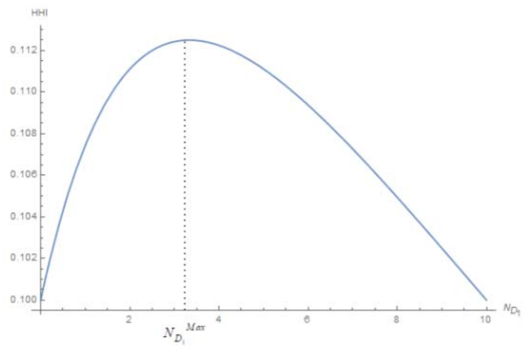

Figure D1. HHI Values with Respect to N_{D_1} when N_{D_2}= 10

It can be seen in the example that the HHI is not maximized where $v_1 = v_2$, where the two number of third parties are equal $\left(N_{D_1} = N_{D_2} = 10\right)$, but at $N_{D_1}^{Max} = \frac{10}{3} \cong 3.33$. Thus market concentration is at its highest when the number of third parties in two publisher websites are different from each other. Next we will see how this factor directs the way market concentration is affected by user privacy concerns.

As we saw in Figure 5d, $N_{D_1}^*$ decreases as user privacy concerns for publisher 1's website increases. Even though publisher 2's website user privacy concerns are held steady at $v_2 = 4$, $N_{D_2}^*$ will be affected by changes in v_1. Here we look at how these changes in number of third parties determines the market concentration. Figure D2 provides the effect of user privacy concerns of one publisher website on HHI when the user privacy concerns of the other publisher website is fixed.

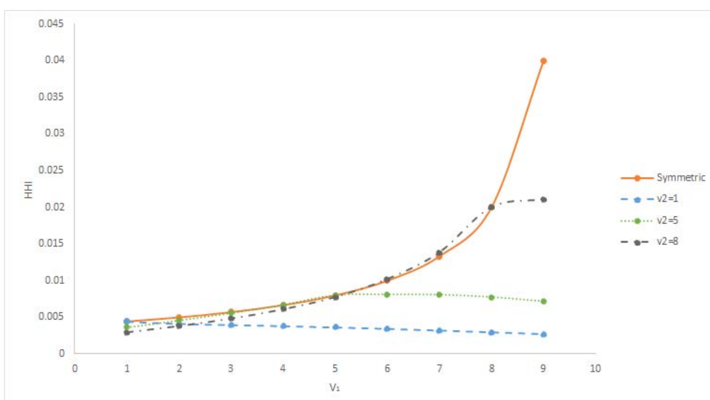

Figure D2. HHI Values with Respect to v_1 when v_2 Is Fixed

The orange line in Figure D2 represents the homogeneous case where both publisher websites are symmetric, and $N_{D_1} = N_{D_2}$. The other lines represent asymmetric cases where v_2 is fixed while v_1 varies. By comparing the homogeneous case to any nonhomogeneous case, it can be seen that the HHI is initially higher for the symmetric case (or equal to in when $v_2 = 1$). As v_1 increases while less than v_2, then market concentration for the asymmetric cases can become higher than the symmetric HHI. At $v_1 = v_2$, the two lines cross, for $v_1 > v_2$, again the HHI for the symmetric case is higher than the asymmetric case.

Appendix E

Collusion

The calculations for collusion are provided for the case where one publisher sets the equilibrium royalties ($R_{W_{Eq.}}$) for its website and the other sets royalties for its website to R_W. The profit in this case is maximized when the the royalty is set to its equilibrium point, $R_{W_{Eq.}}$. However, if the publishers can collude and set an identical royalty ($R_{W_{Col.}}$) for their websites, then they can increase their profits to the collusion equilibrium ($\Pi_{W_{R_{W_{Col.}}}}$), where both firms make higher profits. From the Lemma 1, we know $R_{W_{Eq.}} = R_W^* = \frac{R_D+v}{2}$ and the equilibrium profit (Proposition 3) is given as

$$\Pi_{W_{R_{W_{Eq.}}}} = \Pi_W^* = \frac{M_U\left(8\Phi t - M_U M_D\left(R_D-3v\right)\left(R_D-v\right)\right)}{16\Phi} \tag{E1}$$

For the collusion case, in the publisher website profit equation (9), the prices are set to their optimal values, and both publisher's website royalties are set to R_W. The profit is calculated as

$$\Pi_{W_{R_W}} = \frac{M_U\left(4\Phi t - M_U M_D\left(R_D{}^2 + 2R_W{}^2 + v^2 - 2R_D\left(R_W+v\right)\right)\right)}{8\Phi} \tag{E2}$$

which is maximized at $R_{W_{Col.}} = \frac{R_D}{2}$ for which the profit is

$$\Pi_{W_{R_{W_{Col.}}}} = \frac{M_U\left(8\Phi t - M_U M_D\left(R_D{}^2 - 4R_D v + 2v^2\right)\right)}{16\Phi} \tag{E3}$$

It can easily be shown using the formulae above that the phenomena that collusion royalties are lower than equilibrium royalties and collusion profits are higher than equilibrium profits are analytical results and hold irrespective of the parameters. In other words, the following hold:

$$R_{W_{Eq.}} > R_{W_{Col.}} \tag{E4}$$

$$\Pi_{W_{R_{W_{Eq.}}}} > \Pi_{W_{R_{W_{Col.}}}} \tag{E5}$$

This collusion results in setting lower royalties overall, and thus is also beneficial for the third parties, as presented in the third party surplus curve in Figure E1.

When collusion is possible, the following formula provides the effect of R_W on the user surplus when prices are set to their equilibrium values (we do not consider firms colluding on price but rather only royalties), and publisher websites set equal royalties $\left(CS_{R_W}\right)$:

$$CS_{R_W} = \frac{M_D M_U\left(R_D{}^2 - 4R_D v + v\left(2R_W+v\right)\right) + \Phi\left(4X - 5t\right)}{4\Phi} \tag{E6}$$

Taking the partial derivative of the user surplus with respect to R_W we have

$$\frac{\partial CS_{R_W}}{\partial R_W} = \frac{M_U M_D v}{2\Phi} \tag{E8}$$

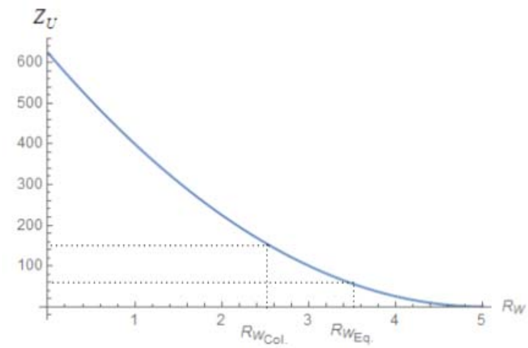

Figure E1. Third Party Surplus with and Without Collusion with Respect to Royalties

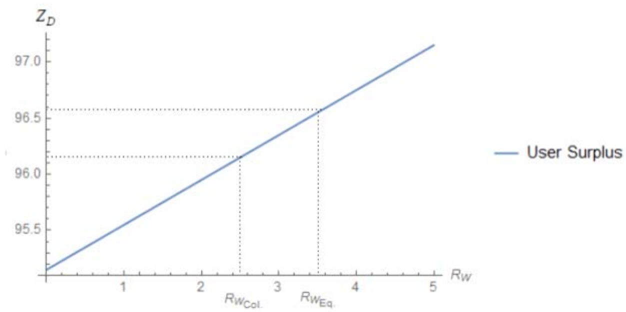

Figure E2. User Surplus with and Without Collusion with Respect to Royalties

Te collusion is thus not beneficial for the users, as they will be exposed to more third parties due to decrease in R_W This can be seen for a numerical example in Figure E2.

Appendix F

Duopoly with Nonlinear Utility Function

For the duopoly with nonlinear utility function (NL Duopoly), the transformations (23) and (24) are made in the base model. We then look at the behavior of the model variables with respect to the different variables. Tables F1, F2, and F3 present the comparison of the behavior of parameters and variables between the base duopoly model versus the duopoly model with nonlinear utility function.

In Table F1, it can be seen that while the behavior of some of the parameters are different in the duopoly model with nonlinear utility compared to the base model, the main results of the model in terms of user privacy concerns (v) are consistent with the base model. In Table F2, we can see that the behavior of the number of users and third parties are entirely consistent between the two duopoly models. As described in Table F3, the NL Duopoly model mostly picks up the effect of higher range user privacy concerns seen in the duopoly model for the publisher website profit. While we see some discrepancy among the two models, the overall conclusion is that the results for the duopoly and NL duopoly models are consistent. This is especially true for the key results with respect to user privacy concerns (v).

Table F1. Publisher Website Decision Variables				
	Changes With Respect To		Duopoly	NL Duopoly
Publisher Website Royalty R_W^*	v	$\left(\frac{\partial R_W^*}{\partial v}\right)$	+	+
	R_D	$\left(\frac{\partial R_W^*}{\partial R_D}\right)$	+	+
	Φ	$\left(\frac{\partial R_W^*}{\partial \Phi}\right)$	Independent	–
	M_U	$\left(\frac{\partial R_W^*}{\partial M_U}\right)$	Independent	+
	M_D	$\left(\frac{\partial R_W^*}{\partial M_D}\right)$	Independent	+
	t	$\left(\frac{\partial R_W^*}{\partial t}\right)$	Independent	Independent
	T_{R_W}	$\left(\frac{\partial R_W^*}{\partial T_{R_W}}\right)$	–	–
	T_{P_W}	$\left(\frac{\partial R_W^*}{\partial T_{P_W}}\right)$	–	–
Publisher Website Price P_W^*	v	$\left(\frac{\partial P_W^*}{\partial v}\right)$	+	+
	R_D	$\left(\frac{\partial P_W^*}{\partial R_D}\right)$	–	–
	Φ	$\left(\frac{\partial P_W^*}{\partial \Phi}\right)$	+	–
	M_U	$\left(\frac{\partial P_W^*}{\partial M_U}\right)$	–	+
	M_D	$\left(\frac{\partial P_W^*}{\partial M_D}\right)$	–	+
	t	$\left(\frac{\partial P_W^*}{\partial t}\right)$	+	+
	T_{R_W}	$\left(\frac{\partial P_W^*}{\partial T_{R_W}}\right)$	+	+
	T_{P_W}	$\left(\frac{\partial P_W^*}{\partial T_{P_W}}\right)$	–	–

Table F2. Impact on Number of Users and Third Parties

	Changes With Respect To	Duopoly	NL Duopoly
Number of Users N_D^*	v $\left(\frac{\partial N_U^*}{\partial v}\right)$	Independent	Independent
	R_D $\left(\frac{\partial N_U^*}{\partial R_D}\right)$	Independent	Independent
	Φ $\left(\frac{\partial N_U^*}{\partial \Phi}\right)$	Independent	Independent
	M_U $\left(\frac{\partial N_U^*}{\partial M_U}\right)$	Independent	Independent
	M_D $\left(\frac{\partial N_U^*}{\partial M_D}\right)$	Independent	Independent
	t $\left(\frac{\partial N_U^*}{\partial t}\right)$	Independent	Independent
	T_{R_W} $\left(\frac{\partial N_U^*}{\partial T_{R_W}}\right)$	Independent	Independent
	T_{P_W} $\left(\frac{\partial N_U^*}{\partial T_{P_W}}\right)$	Independent	Independent
Number of Third Parties N_D^*	v $\left(\frac{\partial N_D^*}{\partial v}\right)$	–	–
	R_D $\left(\frac{\partial N_D^*}{\partial R_D}\right)$	+	+
	Φ $\left(\frac{\partial N_D^*}{\partial \Phi}\right)$	–	–
	M_U $\left(\frac{\partial N_D^*}{\partial M_U}\right)$	+	+
	M_D $\left(\frac{\partial N_D^*}{\partial M_D}\right)$	+	+
	t $\left(\frac{\partial N_D^*}{\partial t}\right)$	Independent	Independent
	T_{R_W} $\left(\frac{\partial N_D^*}{\partial T_{R_W}}\right)$	–	–
	T_{P_W} $\left(\frac{\partial N_D^*}{\partial T_{P_W}}\right)$	+	+

Table F3. Publisher Website Profit, User Surplus, and Third Party Surplus

	Changes With Respect To		Duopoly	NL Duopoly
Publisher Website Profit Π_W^*	v	$\left(\frac{\partial \Pi_W^*}{\partial v}\right)$	+ then –	–
	R_D	$\left(\frac{\partial \Pi_W^*}{\partial R_D}\right)$	– for low *v* + for high *v*	+
	Φ	$\left(\frac{\partial \Pi_W^*}{\partial \Phi}\right)$	– for low *v* + for high *v*	–
	M_U	$\left(\frac{\partial \Pi_W^*}{\partial M_U}\right)$	+	+
	M_D	$\left(\frac{\partial \Pi_W^*}{\partial M_D}\right)$	– for low *v* + for high *v*	+
	t	$\left(\frac{\partial \Pi_W^*}{\partial t}\right)$	+	+
	T_{R_W}	$\left(\frac{\partial \Pi_W^*}{\partial T_{R_W}}\right)$	– for low *v* + for high *v*	–
	T_{P_W}	$\left(\frac{\partial \Pi_W^*}{\partial T_{P_W}}\right)$	–	–
User Surplus Z_U^*	v	$\left(\frac{\partial Z_U^*}{\partial v}\right)$	– then +	+
	R_D	$\left(\frac{\partial Z_U^*}{\partial R_D}\right)$	– for low *v* + for high *v*	–
	Φ	$\left(\frac{\partial Z_U^*}{\partial \Phi}\right)$	– for low *v* + for high *v*	+
	M_U	$\left(\frac{\partial Z_U^*}{\partial M_U}\right)$	– for low *v* + for high *v*	+
	M_D	$\left(\frac{\partial Z_U^*}{\partial M_D}\right)$	– for low *v* + for high *v*	–
	t	$\left(\frac{\partial Z_U^*}{\partial t}\right)$	–	–
	T_{R_W}	$\left(\frac{\partial Z_U^*}{\partial T_{R_W}}\right)$	– for low *v* + for high *v*	+
	T_{P_W}	$\left(\frac{\partial Z_U^*}{\partial T_{P_W}}\right)$	– for low *v* + for high *v*	–

Table F3. Publisher Website Profit, User Surplus, and Third Party Surplus (Continued)				
	Changes With Respect To		Duopoly	NL Duopoly
Third Party Surplus Z_D^*	v	$\left(\frac{\partial Z_D^*}{\partial v}\right)$	–	–
	R_D	$\left(\frac{\partial Z_D^*}{\partial R_D}\right)$	+	+
	Φ	$\left(\frac{\partial Z_D^*}{\partial \Phi}\right)$	Independent	+
	M_U	$\left(\frac{\partial Z_D^*}{\partial M_U}\right)$	+	+
	M_D	$\left(\frac{\partial Z_D^*}{\partial M_D}\right)$	Independent	–
	t	$\left(\frac{\partial Z_D^*}{\partial t}\right)$	Independent	–
	T_{R_W}	$\left(\frac{\partial Z_D^*}{\partial T_{R_W}}\right)$	–	–
	T_{P_W}	$\left(\frac{\partial Z_D^*}{\partial T_{P_W}}\right)$	+	+

Appendix G

Monopoly Model

The following tables compare the effect of model parameters on the key variables in the model, as well as on the publisher website profit, user and third party surplus. Tables G1, G2, and G3 present the comparison of the behavior of parameters and variables between the base duopoly model versus the monopoly model.

In Table G1, it can be seen that the decision variables of royalties and prices behave similarly in the monopoly and duopoly models. Addition of the Hotelling's parameter in the duopoly model enables us to see the effect of competition on the prices. The higher the differentiation between the two publisher websites (higher), the higher the prices. In other words, competition would decrease the prices for the publisher websites.

In Table G2, we can see that the behavior of the number of users in the monopoly model is different from the duopoly model, because the key assumption in the duopoly model is that the market is covered. Thus, the number of users in the duopoly model is independent of the parameters. For the number of third parties, we see that the behavior of the monopoly and duopoly models are similar.

In Table G3, the publisher website profit, user surplus, and third party surplus are presented. The monopoly model picks up the effect of higher range user privacy concerns seen in the duopoly model for the publisher website profit. For user surplus, the monopoly model picks up the effect of the lower range of user privacy concerns seen in the duopoly model. While we see two different effects in the duopoly model, the pattern of results is consistent between the two models. Thus, our overall conclusion is that the results for the duopoly and monopoly models are not inconsistent. This is especially true for the key results with respect to user privacy concerns (v).

Table G1. Publisher Website Decision Variables				
	Changes With Respect To		Duopoly	Monopoly
Publisher Website Royalty R_W^*	v	$\left(\frac{\partial R_W^*}{\partial v}\right)$	+	+
	R_D	$\left(\frac{\partial R_W^*}{\partial R_D}\right)$	+	+
	Φ	$\left(\frac{\partial R_W^*}{\partial \Phi}\right)$	Independent	Independent
	M_U	$\left(\frac{\partial R_W^*}{\partial M_U}\right)$	Independent	Independent
	M_D	$\left(\frac{\partial R_W^*}{\partial M_D}\right)$	Independent	Independent
	t	$\left(\frac{\partial R_W^*}{\partial t}\right)$	Independent	N/A
	T_{R_W}	$\left(\frac{\partial R_W^*}{\partial T_{R_W}}\right)$	–	–
	T_{P_W}	$\left(\frac{\partial R_W^*}{\partial T_{P_W}}\right)$	–	–
Publisher Website Price P_W^*	v	$\left(\frac{\partial P_W^*}{\partial v}\right)$	+	+
	R_D	$\left(\frac{\partial P_W^*}{\partial R_D}\right)$	–	–
	Φ	$\left(\frac{\partial P_W^*}{\partial \Phi}\right)$	+	+
	M_U	$\left(\frac{\partial P_W^*}{\partial M_U}\right)$	–	–
	M_D	$\left(\frac{\partial P_W^*}{\partial M_D}\right)$	–	–
	t	$\left(\frac{\partial P_W^*}{\partial t}\right)$	+	N/A
	T_{R_W}	$\left(\frac{\partial P_W^*}{\partial T_{R_W}}\right)$	+	+
	T_{P_W}	$\left(\frac{\partial P_W^*}{\partial T_{P_W}}\right)$	–	–

Table G2. Impact on Number of Users and Third Parties

	Changes With Respect To	Duopoly	Monopoly
Number of Users N_D^*	v $\left(\frac{\partial N_U^*}{\partial v}\right)$	Independent	–
	R_D $\left(\frac{\partial N_U^*}{\partial R_D}\right)$	Independent	+
	Φ $\left(\frac{\partial N_U^*}{\partial \Phi}\right)$	Independent	–
	M_U $\left(\frac{\partial N_U^*}{\partial M_U}\right)$	Independent	+
	M_D $\left(\frac{\partial N_U^*}{\partial M_D}\right)$	Independent	+
	t $\left(\frac{\partial N_U^*}{\partial t}\right)$	Independent	N/A
	T_{R_W} $\left(\frac{\partial N_U^*}{\partial T_{R_W}}\right)$	Independent	–
	T_{P_W} $\left(\frac{\partial N_U^*}{\partial T_{P_W}}\right)$	Independent	+
Number of Third Parties N_D^*	v $\left(\frac{\partial N_D^*}{\partial v}\right)$	–	–
	R_D $\left(\frac{\partial N_D^*}{\partial R_D}\right)$	+	+
	Φ $\left(\frac{\partial N_D^*}{\partial \Phi}\right)$	–	–
	M_U $\left(\frac{\partial N_D^*}{\partial M_U}\right)$	+	+
	M_D $\left(\frac{\partial N_D^*}{\partial M_D}\right)$	+	+
	t $\left(\frac{\partial N_D^*}{\partial t}\right)$	Independent	N/A
	T_{R_W} $\left(\frac{\partial N_D^*}{\partial T_{R_W}}\right)$	–	–
	T_{P_W} $\left(\frac{\partial N_D^*}{\partial T_{P_W}}\right)$	+	+

Table G3. Publisher Website Profit, User Surplus, and Third Party Surplus

	Changes With Respect To		Duopoly	Monopoly
Publisher Website Profit Π_W^*	v	$\left(\frac{\partial \Pi_W^*}{\partial v}\right)$	+ then –	–
	R_D	$\left(\frac{\partial \Pi_W^*}{\partial R_D}\right)$	– for low *v* + for high *v*	+
	Φ	$\left(\frac{\partial \Pi_W^*}{\partial \Phi}\right)$	+ for low *v* – for high *v*	–
	M_U	$\left(\frac{\partial \Pi_W^*}{\partial M_U}\right)$	+	+
	M_D	$\left(\frac{\partial \Pi_W^*}{\partial M_D}\right)$	– for low *v* + for high *v*	+
	t	$\left(\frac{\partial \Pi_W^*}{\partial t}\right)$	+	N/A
	T_{R_W}	$\left(\frac{\partial \Pi_W^*}{\partial T_{R_W}}\right)$	+ for low *v* – for high *v*	–
	T_{P_W}	$\left(\frac{\partial \Pi_W^*}{\partial T_{P_W}}\right)$	–	+
User Surplus Z_U^*	v	$\left(\frac{\partial Z_U^*}{\partial v}\right)$	– then +	–
	R_D	$\left(\frac{\partial Z_U^*}{\partial R_D}\right)$	+ for low *v* – for high *v*	+
	Φ	$\left(\frac{\partial Z_U^*}{\partial \Phi}\right)$	– for low *v* + for high *v*	–
	M_U	$\left(\frac{\partial Z_U^*}{\partial M_U}\right)$	+ for low *v* – for high *v*	+
	M_D	$\left(\frac{\partial Z_U^*}{\partial M_D}\right)$	+ for low *v* – for high *v*	+
	t	$\left(\frac{\partial Z_U^*}{\partial t}\right)$	–	N/A
	T_{R_W}	$\left(\frac{\partial Z_U^*}{\partial T_{R_W}}\right)$	– for low *v* + for high *v*	–
	T_{P_W}	$\left(\frac{\partial Z_U^*}{\partial T_{P_W}}\right)$	+ for low *v* – for high *v*	+

Table G3. Publisher Website Profit, User Surplus, and Third Party Surplus (Continued)			
	Changes With Respect To	**Duopoly**	**Monopoly**
Third Party Surplus Z_D^*	v $\left(\frac{\partial Z_D^*}{\partial v}\right)$	–	–
	R_D $\left(\frac{\partial Z_D^*}{\partial R_D}\right)$	+	+
	Φ $\left(\frac{\partial Z_D^*}{\partial \Phi}\right)$	Independent	–
	M_U $\left(\frac{\partial Z_D^*}{\partial M_U}\right)$	+	+
	M_D $\left(\frac{\partial Z_D^*}{\partial M_D}\right)$	Independent	+
	t $\left(\frac{\partial Z_D^*}{\partial t}\right)$	Independent	N/A
	T_{R_W} $\left(\frac{\partial Z_D^*}{\partial T_{R_W}}\right)$	–	–
	T_{P_W} $\left(\frac{\partial Z_D^*}{\partial T_{P_W}}\right)$	+	+

Appendix H

Empirical Analysis

We find partial support for the model by empirically examining the number of third party participants utilized by publisher websites, as well as the industry concentration of third parties. Alexa Internet provides rankings for publisher websites within 17 different subject categories.[1] We carry out an exploratory validation study on the 100 most-visited publisher websites from seven of these subject categories (news, arts, shopping, kids and teens, health, business, and adult) provided and ranked by Alexa website rankings. These seven categories were selected with the intention of finding subject categories for which users might reasonably be expected to have different intentions to disclose personal information and browsing behavior due to the nature of the subject content. For the study, an automated browser accessed the home page of a publisher's website, and the connections made from the publisher's website to third parties were recorded. We used page loading time plus a 3-second window to collect data gathered using a residential internet plan and using Lightbeam for Firefox (Windows) to record these connections.

To better capture the structure of the industry, we profile the third parties and separate them based on the industry sectors as classified by Cookiepedia.co.uk. The three industry sectors are targeting/advertising (*T/A*), functionality (*F*), and performance (*P*). For those third parties that are not profiled in Cookiepedia.co.uk, we make a judgment using available information. A total of 1,893 third party websites are identified, with 568 classified as *T/A*, 487 classified as *F*, 627 classified as *P*, and 211 classified as unknown (*U*). Using different domain finder services,[2] multiple third party websites in each sector owned by the same company are treated as a single third party for analysis, entailing 1,066 unique

[1]Alexa.com/topsites/category. The Alexa list of website categories is consistent with the Open Directory Project categories found at rdf.DMOZ.org/rdf/categories.txt.

[2]This paper uses whois.domaintools.com, whois.net, and who.is.

owner companies comprising 442 classified as *T/A*, 336 classified as *F*, and 340 classified as *P*, with some owner companies providing services in multiple categories. The number of connections made and number of cookies used follow a similar pattern to number of third parties, and so we provide the analysis based on number of third parties only. Table H1 provides a summary of descriptive statistics of the data on number of third parties.

Table H1. Descriptive Statistics for Number of Third Parties Used Among Websites

Subject Category	N	Targeting/Advertising				Functionality				Performace			
		Min	*Max*	*Avg.*	*StDv*	*Min*	*Max*	*Mean*	*StDv*	*Min*	*Max*	*Mean*	*StDv*
News	100	1	65	16.8	11.7	1	10	4.5	2.3	1	10	3.4	2.0
Arts	100	1	55	11.5	10.0	1	10	3.9	2.1	1	7	2.8	1.5
Shopping	100	1	51	9.4	9.2	1	8	2.9	1.6	1	8	2.7	1.6
Kids and teens	100	1	58	8.8	11.5	1	7	2.7	1.4	1	7	2.3	1.5
Health	100	1	70	8.2	11.3	1	7	2.9	1.6	1	7	2.4	1.5
Business	100	1	60	6.9	9.2	1	7	2.4	1.6	1	7	2.4	1.4
Adult	100	1	34	3.2	5.2	1	9	2.2	1.5	1	7	1.7	1.0

Observations

Noting that information sensitivity and user privacy concerns likely vary among different publisher websites, we expect the sharing behavior to differ for publisher websites with different subjects. Figure H1 provides the sharing behavior for the top 100 publisher websites in each subject category and industry sector.

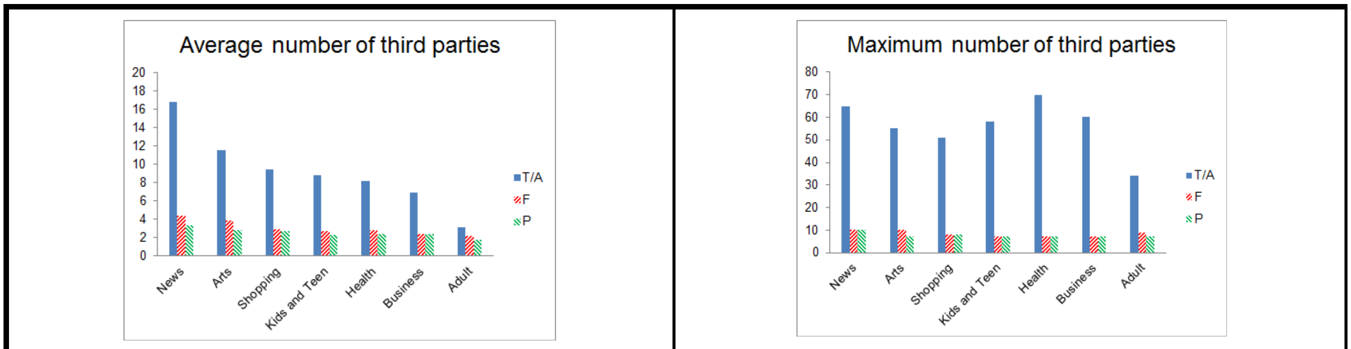

Figure H1. Third Party Usage by Subject Categories and Industry Sectors

Table H2 provides the statistical test results for number of third parties on different categories of websites. Since the variances are different among the categories, we use the Welch's two-tailed t-test for testing if the means are different among these websites. It can be seen from Table H2 that the number of third parties are significantly different for most of the categories. Especially, in the *T/A* sector, news and adult categories are statistically different from other categories.

Table H3 provides the statistical test results for number of third parties on different sectors of the industry. It can be seen from Table H3 that the number of third parties used in the *T/A* industry sector is significantly higher than for both *F* and *P*.

We also examine the third party market concentration measure, using the Herfindahl-Hirschman index (HHI) based on the average monthly unique visitors to the publisher's website in the United States for a single year period ending in March 2014 as provided by compete.com. The *T/A* sector has the lowest HHI concentrations, followed by *P*, and then by *F*. In terms of publisher website categories, we see that news and arts have the lowest industry concentration, with adult having the highest industry concentration. The HHI results are provided in Figure H2.

Table H2. P-Values for Testing if Number of Third Parties Used in Different Categories of Websites are Statistically Different

		News	Arts	Shopping	Kids & Teens	Health	Business
T/A	News						
	Arts	0.001					
	Shopping	0.000	0.125				
	Kids & Teens	0.000	0.082	0.700			
	Health	0.000	0.027	0.390	0.670		
	Business	0.000	0.001	0.056	0.191	0.393	
	Adult	0.000	0.000	0.000	0.000	0.000	0.001
F	News						
	Arts	0.064					
	Shopping	0.000	0.001				
	Kids & Teens	0.000	0.000	0.299			
	Health	0.000	0.000	0.704	0.522		
	Business	0.000	0.000	0.012	0.106	0.032	
	Adult	0.000	0.000	0.002	0.024	0.006	0.551
P	News						
	Arts	0.019					
	Shopping	0.006	0.568				
	Kids & Teens	0.000	0.023	0.099			
	Health	0.000	0.061	0.210	0.690		
	Business	0.000	0.036	0.147	0.796	0.877	
	Adult	0.000	0.000	0.000	0.001	0.000	0.000

Table H3. P-Values for Testing If Number of Third Parties Used in Different Industry Sectors Are Statistically Different

		T/A	F
News	T/A		
	F	0.000	
	P	0.000	0.001
Arts	T/A		
	F	0.000	
	P	0.000	0.000
Shopping	T/A		
	F	0.000	
	P	0.000	0.286
Kids and Teens	T/A		
	F	0.000	
	P	0.000	0.067
Health	T/A		
	F	0.000	
	P	0.000	0.051
Business	T/A		
	F	0.000	
	P	0.000	0.925
Adult	T/A		
	F	0.080	
	P	0.007	0.008

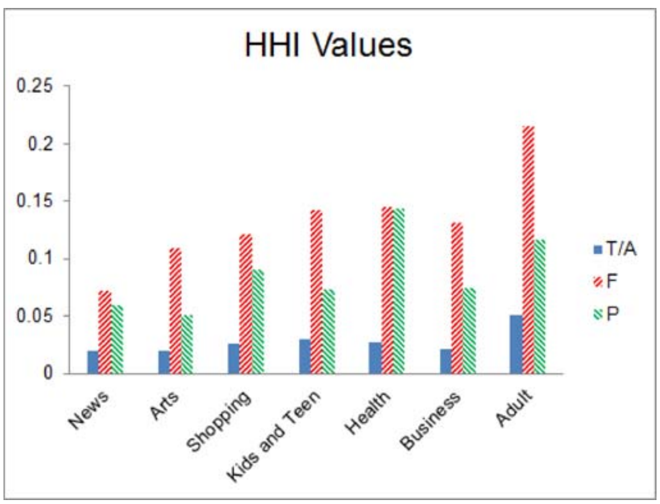

Figure H2. HHI for Third Party Industry by Subject Categories and Industry Sectors