Technical Appendix for "Link to Success: How Blogs Build an Audience by Promoting Rivals"

Dina Mayzlin Hema Yoganarasimhan *

July 3, 2012

1 Lemma 1

Let $\lambda = Pr\left[U_A^R > U_j^R | \mu\left(\Omega^R\right)\right]$ be the probability that a reader R will revisit blog A at Stage 2a.

Lemma 1 *The total expected utility of a blog A in period 2, $E[u(N^R + N^I + N^A)]$, is increasing in probabilty λ.*

Proof: Suppose that n readers visited blog A in period 1. We need to show that $\frac{dE\left[u\left(N^R+N^I+N^A\right)|n\right]}{d\lambda}$ for all n. First, we expand $E\left[u\left(N^R + N^I + N^A\right) \mid n\right]$ as follows:

$$\begin{aligned}
E\left[u\left(N^R + N^I + N^A\right) \mid n\right] &= E\left[u\left(N^R + N^I + n\right)\right]\binom{n}{0}\lambda^n \\
&\quad + E\left[u\left(N^R + N^I + n - 1\right)\right]\binom{n}{1}\lambda^{n-1}(1-\lambda) + \cdots \\
&\quad + E\left[u\left(N^R + N^I + n - k\right)\right]\binom{n}{k}\lambda^{n-k}(1-\lambda)^k + \cdots \\
&\quad + E\left[u\left(N^R + N^I + 0\right)\right]\binom{n}{n}(1-\lambda)^n
\end{aligned}$$

Differentiating this $w.r.t$ to λ, we have:

$$\frac{dE\left[u\left(N^R + N^I + N^A\right) \mid n\right]}{d\lambda} = \sum_{k=0}^{n} E\left[u\left(N^R + N^I + n - k\right)\right]\binom{n}{k}\left[\begin{array}{c}(n-k)\lambda^{n-k-1}(1-\lambda)^k \\ -k\lambda^{n-k}(1-\lambda)^{k-1}\end{array}\right]$$

*Please address all correspondence to: hyoganarasimhan@ucdavis.edu.

This in turn can be rewritten as:

$$\frac{dE\left[u\left(N^R+N^I+N^A\right)\mid n\right]}{d\lambda} = \sum_{k=0}^{n-1}\lambda^{n-k-1}(1-\lambda)^k\left[\begin{array}{c} E\left[u\left(N^R+N^I+n-k\right)\right]\binom{n}{k}(n-k) \\ -E\left[u\left(N^R+N^I+n-k-1\right)\right]\binom{n}{k+1}(k+1)\end{array}\right] \quad (1)$$

If we show that each term in Equation (1) is positive, then we would have show that $\frac{dE\left[u\left(N^R+N^I+N^A\right)|n\right]}{d\lambda} > 0$. So consider the k^{th} term in (1). First, note that $\binom{n}{k}(n-k) = \binom{n}{k+1}(k+1)$ because

$$\begin{aligned}\binom{n}{k}(n-k) &= \frac{n!}{(n-k)!\,k!}(n-k) = \frac{n!}{(n-k-1)!\,k!} \\ \binom{n}{k+1}(k+1) &= \frac{n!}{(n-k-1)!\;(k+1)!}(k+1) = \frac{n!}{(n-k-1)!\,k!}\end{aligned}$$

Second, note that $E\left[u\left(N^R+N^I+n-k\right)\right] > E\left[u\left(N^R+N^I+n-k-1\right)\right]$ by assumption. Hence, the k^{th} term in Equation (1) is always positive. Since each single term in (1) satisfies this property, we have:

$$\frac{dE\left[u\left(N^R+N^I+N^A\right)\mid n\right]}{d\lambda} = \sum_{k=0}^{n-1}\lambda^{n-k-1}(1-\lambda)^k\left[\begin{array}{c} E\left[u\left(N^R+N^I+n-k\right)\right]\binom{n}{k}(n-k) \\ -E\left[u\left(N^R+N^I+n-k-1\right)\right]\binom{n}{k+1}(k+1)\end{array}\right] > 0$$

Q.E.D

2 Analysis of Benchmark Cases

We now present the detailed analysis of the benchmark cases discussed in Section 4.1 in the main text.

2.1 No Heterogeneity on the Ability to Break News

In this case, there is no heterogeneity among bloggers on the ability to break news, i.e., $v = w > 0$, but there is heterogeneity in news-breaking ability, i.e., $p > q > 0$. Hence R's posterior beliefs on blog j's ability to break news remain at α_0, regardless of whether j posted information or not. That is, $\alpha_U = \alpha_D = \alpha_0$. Consequently, the inequalities (6) and (7) simplify to

$$EU_A(\alpha_0, \beta_U) \geq EU_A(\alpha_0, \beta_D) \quad (2)$$

Similarly, the inequalities (8) and (9) simplify to

$$EU_A(\alpha_0, \beta_U) < EU_A(\alpha_0, \beta_D) \tag{3}$$

Recall that $EU(\alpha, \beta) = \alpha u + (1 - \alpha)\beta(u - c)$. Since $EU_A(\alpha, \beta)$ is increasing in β and $\beta_U > \beta_D$, (2) holds in this case, while (3) does not. This in turn implies that the only equilibrium that exists is the (L, L) equilibrium (since all other equilibria require that (3) is satisfied). Hence, when there is no heterogeneity in bloggers ability to break news, (L, L) is the unique equilibrium.

2.2 No Heterogeneity on the Ability to Find News in Other Blogs

In this case there is no heterogeneity among bloggers on the ability to find news, i.e., $p = q > 0$, but there is heterogeneity in news-breaking ability, i.e., $v > w > 0$. Hence, R's posterior beliefs on blog j's ability to find news remain at β_0, regardless of whether j linked or not. That is, $\beta_U = \beta_D = \beta_0$. Consequently, the inequalities (6) and (7) simplify to

$$EU_B(\alpha_U, \beta_0) \leq EU_C(\alpha_0, \beta_0) \tag{4}$$

Since $EU(\alpha, \beta)$ is increasing in α and $\alpha_U > \alpha_0$, (4) does not hold. This rules out (L, L), (L, DL), and (DL, L) equilibria since they all require for (4) to hold. The fact that (4) is not satisfied also implies that (DL, DL) exists for all parameter values. Hence, when there is no heterogeneity in bloggers ability to find news, (DL, DL) is the unique equilibrium.

2.3 No Heterogeneity on Any Dimension

In this case, $v = w > 0$ and $p = q > 0$. Given the homogeneity across types, there is no updating of beliefs. This implies that (6) and (7) reduce to

$$EU_B(\alpha_0, \beta_0) \leq EU_C(\alpha_0, \beta_0) \tag{5}$$

while (8) and (9) reduce to

$$EU_B(\alpha_0, \beta_0) > EU_C(\alpha_0, \beta_0) \tag{6}$$

Clearly, while (5) is true, (6) is not. Hence, (L,L) is the unique equilibrium.

3 Relative Positions of Iso-curves

Proposition A1. *Consider the different equilibria regions as pictured in Figure 3.*

- $6^=$ *and* $9^=$ *are increasing in* v*,* $8^=$ *and* $7^=$ *are either increasing in* v *or increasing and then decreasing in* v*.*

- *For all* $0 \leq q < p \leq 1, 0 \leq w < v \leq 1$, $9^=$ *lies above* $6^=$, $8^=$, *and* $7^=$; $7^=$ *lies below* $6^=$, $8^=$, *and* $9^=$.

Proof: $6^=$, $7^=$, $8^=$, and $9^=$ are the iso-curves of the inequalities (6)-(9) in the $v-p$ plane. That is:

$$\begin{aligned} f_6 &= (1-\alpha_U)(\beta_U-\beta_D)(u-c)-(\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c] \\ f_7 &= (1-\alpha_D)(\beta_U-\beta_D)(u-c)-(\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c] \\ f_8 &= (1-\alpha_D)(\beta_U-\beta_0)(u-c)-(\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c] \\ f_9 &= (1-\alpha_U)(\beta_U-\beta_0)(u-c)-(\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c] \end{aligned}$$

$6^=$ is defined as $f_6 = 0$; $7^=$ is defined as $f_7 = 0$, and so on. The slopes of the iso-curve are obtained using the implicit function theorem and are:

$$\left.\frac{dp}{dv}\right|_6 = -\frac{\frac{df_6}{dv}}{\frac{df_6}{dp}} = \frac{(\beta_U-\beta_D)(u-c)\frac{d\alpha_U}{dv}+[u-\beta_0(u-c)]\frac{d(\alpha_U-\alpha_0)}{dv}}{(1-\alpha_U)(u-c)\frac{d(\beta_U-\beta_D)}{dp}+(\alpha_U-\alpha_0)(u-c)\frac{d\beta_0}{dp}}$$

$$\left.\frac{dp}{dv}\right|_7 = -\frac{\frac{df_7}{dv}}{\frac{df_7}{dp}} = \frac{(\beta_U-\beta_D)(u-c)\frac{d\alpha_D}{dv}+[u-\beta_0(u-c)]\frac{d(\alpha_U-\alpha_0)}{dv}}{(1-\alpha_D)(u-c)\frac{d(\beta_U-\beta_D)}{dp}+(\alpha_U-\alpha_0)(u-c)\frac{d\beta_0}{dp}}$$

$$\left.\frac{dp}{dv}\right|_8 = -\frac{\frac{df_8}{dv}}{\frac{df_8}{dp}} = \frac{(\beta_U-\beta_0)(u-c)\frac{d\alpha_D}{dv}+[u-\beta_0(u-c)]\frac{d(\alpha_U-\alpha_0)}{dv}}{(1-\alpha_D)(u-c)\frac{d(\beta_U-\beta_0)}{dp}+(\alpha_U-\alpha_0)(u-c)\frac{d\beta_0}{dp}}$$

$$\left.\frac{dp}{dv}\right|_9 = -\frac{\frac{df_9}{dv}}{\frac{df_9}{dp}} = \frac{(\beta_U-\beta_0)(u-c)\frac{d\alpha_U}{dv}+[u-\beta_0(u-c)]\frac{d(\alpha_U-\alpha_0)}{dv}}{(1-\alpha_U)(u-c)\frac{d(\beta_U-\beta_0)}{dp}+(\alpha_U-\alpha_0)(u-c)\frac{d\beta_0}{dp}}$$

where all the following terms are positive:

- $\frac{d\alpha_U}{dv} = \frac{\gamma[\gamma(v-w)^2+(2v-w)w]}{[\gamma v+(1-\gamma)w]^2} > 0$ because both the numerator and denominator are positive. Recall that $v-w>0$ and $2v-w>0$, by definition.

- $\frac{d(\alpha_U-\alpha_0)}{dv} = \frac{\gamma(1-\gamma)(v-w)[\gamma(v-w)+2w]}{[\gamma v+(1-\gamma)w]^2} > 0$ because all the terms in the numerator and denominator are positive

- $\frac{d(\beta_U-\beta_0)}{dp} = \frac{\delta(1-\delta)(p-q)[\delta(p-q)+2q]}{[\delta p+(1-\delta)q]^2} > 0$ because both the numerator and denominator are positive. Recall that $p-q>0$.

- $\frac{d\beta_0}{dp} = \delta > 0$.

- $\frac{d(\beta_U-\beta_D)}{dp} = \frac{\delta(1-\delta)(p-q)[2(1-q)q+\delta(p-q)(1-2q)]}{[\delta p+(1-\delta)q]^2[\delta(1-p)+(1-\delta)(1-q)]^2}$. Here all the terms except $2(1-q)q+\delta(p-q)(1-2q)$ are always positive. Even this term is clearly positive if $q \leq 0.5$. So consider the case where $q > 0.5$. Then, $2(1-q)q+\delta(p-q)(1-2q) > 2(1-q)q+(1-q)(1-2q) = 1-q > 0$. Hence, $\frac{d(\beta_U-\beta_D)}{dp} > 0$ for all parameter values.

- $\frac{d\alpha_D}{dv} = \frac{\gamma\{1+\gamma v^2-(1-\gamma)w^2-2v[1-(1-\gamma)w]\}}{[\gamma(1-v)+(1-\gamma)(1-w)]^2}$. $\frac{d\alpha_D}{dv}$ is negative when the numerator is negative. Solving for the roots of the numerator, we have: $v_1 = \frac{1-w+\gamma w-\sqrt{1-\gamma-2w+2\gamma w+w^2-\gamma w^2}}{\gamma}$, $v_2 = \frac{1-w+\gamma w+\sqrt{1-\gamma-2w+2\gamma w+w^2-\gamma w^2}}{\gamma}$. The first root lies between 0 and 1. The second root is also always positive but lies above 1, and hence is excluded from the analysis. Hence, $\frac{d\alpha_D}{dv}$ is positive for $v < v_1$ and is negative after that.

Using the above results, we have:

1. $\frac{dp}{dv}\Big|_6$ and $\frac{dp}{dv}\Big|_9$ are always positive. Hence $6^=$ and $9^=$ are increasing in v.
2. In case of $\frac{dp}{dv}\Big|_7$ and $\frac{dp}{dv}\Big|_8$, the denominators are always positive, but the numerator consists of a term with the multiplier $\frac{d\alpha_D}{dv}$, which could be negative for large values of v. The other term in the numerator is $[u-\beta_0(u-c)]\frac{d(\alpha_U-\alpha_0)}{dv}$, which is always positive. Neither of these terms dominates the other. So, $\frac{dp}{dv}\Big|_7$ and $\frac{dp}{dv}\Big|_8$ are always increasing for small values of v (till v_1) and may decrease after that.

We can show that the left-hand-sides of $6^=$, $7^=$, $8^=$, and $9^=$ are increasing in p. Thus, if $f(p,v) > g(p,v)$, then $g(p,v) = 0$ lies above $f(p,v) = 0$ in the $v-p$ plane ($f(p,v) = g(p',v) = 0$ where $p' > p$). We can also show that the LHS of $6^=$ is always greater than the LHS of $9^=$, so $9^=$ lies above $6^=$. Similarly, LHS of $6^= <$ LHS of $7^=$, which implies that $6^=$ lies above $7^=$. Finally, LHS of $7^= > 8^=$, so $8^=$ always lies above $7^=$. In summary,

1. (L, L) exists above $6^=$ and (DL, L) exists between $9^=$ and $7^=$. Since, $9^=$ lies above $6^=$, both (DL, L) and (L, L) exist between $9^=$ and $6^=$, whereas (L, L) uniquely exists above $9^=$. Since $9^=$ goes through the origin ($v = w, p = q$) and has a positive slope, the region where (L, L) is unique is non-empty.

2. (DL, DL) exists below $8^=$ and (DL, L) exists between $9^=$ and $7^=$. Since $7^=$ lies below $8^=$, both (DL, L) and (L, L) exist between $8^=$ and $7^=$, whereas (DL, DL) uniquely exists below $7^=$. Since $7^=$ goes through the origin ($v = w, p = q$) and has a positive slope for small v, the region where (DL,DL) is unique is non-empty.

3. (DL, L) exists between unique regions of (L, L) and (DL, DL), *i.e.*, in between $9^=$ and $7^=$.

Q.E.D.

4 Extensions

In this section, we consider three extensions of the main model.

4.1 Correlated Abilities

First, we present the detailed analysis of the case where bloggers' abilities are positively correlated, as discussed in Section 5 in the main text.

Consider a blogger's decision to link following breaking (or not breaking) the news. We denote by δ_T the belief that the blogger is h-type on ability to find news in other blogs. For simplicity, we assume that the prior on ability to break the news is symmetric: $\gamma_0 = 0.5$. This implies that $\delta_0 = 0.5\rho + 0.5(1-\rho) = 0.5$. Hence, $\alpha_0 = 0.5v + 0.5w$,$\beta_0 = 0.5p + 0.5q$. Next, we denote the posterior following breaking the news with a subscript I and the posterior following no news with a subscript NI. That is: $\gamma_I = \frac{v}{v+w}$, $\delta_I = \rho\gamma_I + (1-\rho)(1-\gamma_I)$, $\gamma = \frac{(1-v)}{(1-v)+(1-w)}$, and $\delta_{NI} = \rho\gamma_{NI} + (1-\rho)(1-\gamma_{NI})$. This gives us $\alpha_J = \gamma_J v + (1-\gamma_J)w$ and $\beta_J = \delta_J p + (1-\delta_J)q$, where $J \in \{I, NI\}$.

Next, the blog may link or not link to another blog. We denote the posterior following a link with a subscript L and the posterior following no link with a subscript NL. The posterior beliefs now depend on the equilibrium that is being played: a superscript denoting the equilibrium is added.

(1) Posterior beliefs in (L, L):

$$\gamma_{L,J}^{L,L} = \frac{p\gamma_J\rho + q\gamma_J(1-\rho)}{p\gamma_J\rho + q\gamma_J(1-\rho) + p(1-\gamma_J)(1-\rho) + q(1-\gamma_J)\rho}$$

$$\gamma_{NL,J}^{L,L} = \frac{(1-p)\gamma_J\rho + (1-q)\gamma_J(1-\rho)}{(1-p)\gamma_J\rho + (1-q)\gamma_J(1-\rho) + (1-p)(1-\gamma_J)(1-\rho) + (1-q)(1-\gamma_J)\rho}$$

$$\delta_{L,J}^{L,L} = \frac{p\gamma_J\rho + p(1-\gamma_J)(1-\rho)}{p\gamma_J\rho + q\gamma_J(1-\rho) + p(1-\gamma_J)(1-\rho) + q(1-\gamma_J)\rho}$$

$$\delta_{NL,J}^{L,L} = \frac{(1-p)\gamma_J\rho + (1-p)(1-\gamma_J)(1-\rho)}{(1-p)\gamma_J\rho + (1-q)\gamma_J(1-\rho) + (1-p)(1-\gamma_J)(1-\rho) + (1-q)(1-\gamma_J)\rho}$$

(2) Posterior beliefs in (DL, DL):

$$\gamma_{L,J}^{DL,DL} = \frac{p\gamma_J\rho + q\gamma_J(1-\rho)}{p\gamma_J\rho + q\gamma_J(1-\rho) + p(1-\gamma_J)(1-\rho) + q(1-\gamma_J)\rho}$$

$$\gamma_{NL,J}^{DL,DL} = \gamma_J$$

$$\delta_{L,J}^{DL,DL} = \frac{p\gamma_J\rho + p(1-\gamma_J)(1-\rho)}{p\gamma_J\rho + q\gamma_J(1-\rho) + p(1-\gamma_J)(1-\rho) + q(1-\gamma_J)\rho}$$

$$\delta_{NL,J}^{DL,DL} = \delta_J$$

(3) Posterior beliefs in (L, DL)

$$\gamma_{L,J}^{L,DL} = \frac{p\gamma_J\rho + q\gamma_J(1-\rho)}{p\gamma_J\rho + q\gamma_J(1-\rho) + p(1-\gamma_J)(1-\rho) + q(1-\gamma_J)\rho}$$

$$\gamma_{NL,I}^{L,DL} = \frac{(1-p)\gamma_I\rho + (1-q)\gamma_I(1-\rho)}{(1-p)\gamma_I\rho + (1-q)\gamma_I(1-\rho) + (1-p)(1-\gamma_I)(1-\rho) + (1-q)(1-\gamma_I)\rho}$$

$$\gamma_{NL,NI}^{L,DL} = \gamma_{NI}$$

$$\delta_{L,J}^{L,DL} = \frac{p\gamma_J\rho + p(1-\gamma_J)(1-\rho)}{p\gamma_J\rho + q\gamma_J(1-\rho) + p(1-\gamma_J)(1-\rho) + q(1-\gamma_J)\rho}$$

$$\delta_{NL,I}^{L,DL} = \frac{(1-p)\gamma_I\rho + (1-p)(1-\gamma_I)(1-\rho)}{(1-p)\gamma_I\rho + (1-q)\gamma_I(1-\rho) + (1-p)(1-\gamma_I)(1-\rho) + (1-q)(1-\gamma_I)\rho}$$

$$\delta_{NL,NI}^{L,DL} = \delta_{NI}$$

(4) Posterior beliefs in (DL, L):

$$\gamma_{L,J}^{DL,L} = \frac{p\gamma_J\rho + q\gamma_J(1-\rho)}{p\gamma_J\rho + q\gamma_J(1-\rho) + p(1-\gamma_J)(1-\rho) + q(1-\gamma_J)\rho}$$

$$\gamma_{NL,I}^{DL,L} = \gamma_I$$

$$\gamma_{NL,NI}^{DL,L} = \frac{(1-p)\gamma_{NI}\rho + (1-q)\gamma_{NI}(1-\rho)}{(1-p)\gamma_{NI}\rho + (1-q)\gamma_{NI}(1-\rho) + (1-p)(1-\gamma_{NI})(1-\rho) + (1-q)(1-\gamma_{NI})\rho}$$

$$\delta_{L,J}^{DL,L} = \frac{p\gamma_J\rho + p(1-\gamma_J)(1-\rho)}{p\gamma_J\rho + q\gamma_J(1-\rho) + p(1-\gamma_J)(1-\rho) + q(1-\gamma_J)\rho}$$

$$\delta_{NL,I}^{DL,L} = \delta_I$$

$$\delta_{NL,NI}^{DL,L} = \frac{(1-p)\gamma_{NI}\rho + (1-p)(1-\gamma_{NI})(1-\rho)}{(1-p)\gamma_{NI}\rho + (1-q)\gamma_{NI}(1-\rho) + (1-p)(1-\gamma_{NI})(1-\rho) + (1-q)(1-\gamma_{NI})\rho}$$

After observing both the news-breaking and linking behavior, the final probabilities of breaking news and linking can be written as $\alpha_{S,J}^Q = \gamma_{S,J}^Q v + (1-\gamma_{S,J}^Q)w, \quad \beta_{S,J}^Q = \delta_{S,J}^Q p + (1-\delta_{S,J}^Q)q$, where $J \in \{I, NI\}$,$S \in \{L, NL\}$,$Q \in \{(L,L),(DL,DL),(L,DL),(DL,L)\}$. Note that

$$\alpha_{L,J}^{L,L} = \alpha_{L,J}^{DL,DL} = \alpha_{L,J}^{L,DL} = \alpha_{L,J}^{DL,L} \equiv \alpha_{L,J}, \alpha_{NL,I}^{L,L} = \alpha_{NL,I}^{L,DL} \equiv \alpha_{NL,I}$$

$$\alpha^{L,L}_{NL,NI} = \alpha^{DL,L}_{NL,NI} \equiv \alpha_{NL,NI}, \alpha^{DL,DL}_{NL,I} = \alpha^{DL,L}_{NL,I} \equiv \alpha_I$$

$$\alpha^{DL,DL}_{NL,NI} = \alpha^{L,DL}_{NL,NI} = \alpha_{NI}, \beta^{L,L}_{L,J} = \beta^{DL,DL}_{L,J} = \beta^{L,DL}_{L,J} = \beta^{DL,L}_{L,J} \equiv \beta_{L,J}$$

$$\beta^{L,L}_{NL,I} = \beta^{L,DL}_{NL,I} \equiv \beta_{NL,I}, \beta^{L,L}_{NL,NI} = \beta^{DL,L}_{NL,NI} \equiv \beta_{NL,NI}$$

$$\beta^{DL,DL}_{NL,I} = \beta^{DL,L}_{NL,I} = \beta_I, \beta^{DL,DL}_{NL,NI} = \beta^{L,DL}_{NL,NI} = \beta_{NI}$$

Let $K = (\alpha_I - \alpha_0)u + [(1 - \alpha_I)\beta_I - (1 - \alpha_0)\beta_0](u - c)$. The existence conditions for the four equilibria are: $(Link, Link)$

$$(\alpha_{L,I} - \alpha_{NL,I})u + [(1 - \alpha_{L,I})\beta_{L,I} - (1 - \alpha_{NL,I})\beta_{NL,I}](u - c) > K \tag{7}$$

$$(\alpha_{L,NI} - \alpha_{NL,NI})u + [(1 - \alpha_{L,NI})\beta_{L,NI} - (1 - \alpha_{NL,NI})\beta_{NL,NI}](u - c) > K \tag{8}$$

$(Link, Don'tLink)$

$$(\alpha_{L,I} - \alpha_{NL,I})u + [(1 - \alpha_{L,I})\beta_{L,I} - (1 - \alpha_{NL,I})\beta_{NL,I}](u - c) > K \tag{7}$$

$$(\alpha_{L,NI} - \alpha_{NI})u + [(1 - \alpha_{L,NI})\beta_{L,NI} - (1 - \alpha_{NI})\beta_{NI}](u - c) \leq K \tag{9}$$

$(Don'tLink, Link)$

$$(\alpha_{L,I} - \alpha_I)u + [(1 - \alpha_{L,I})\beta_{L,I} - (1 - \alpha_I)\beta_I](u - c) \leq K \tag{10}$$

$$(\alpha_{L,NI} - \alpha_{NL,NI})u + [(1 - \alpha_{L,NI})\beta_{L,NI} - (1 - \alpha_{NL,NI})\beta_{NL,NI}](u - c) > K \tag{8}$$

$(Don'tLink, Don'tLink)$

$$(\alpha_{L,I} - \alpha_I)u + [(1 - \alpha_{L,I})\beta_{L,I} - (1 - \alpha_I)\beta_I](u - c) \leq K \tag{10}$$

$$(\alpha_{L,NI} - \alpha_{NI})u + [(1 - \alpha_{L,NI})\beta_{L,NI} - (1 - \alpha_{NI})\beta_{NI}](u - c) \leq K \tag{9}$$

Proof of Proposition 4[1]

We can show that if (9) does not hold, then (8) does, and if (7) does not hold, then (10) does. Hence, we have nine possible regions that partition the space. Specifically, we have the following nine regions (each with a different set of equilibria):

Region I – (7) holds, (9) does not hold, (10) holds – **(DL, L)** and **(L, L)** exist

Region II – (7) holds, (9) does not hold, (10) does not hold – **(L, L)** is unique

[1]Proposition is available from the main article.

Region III – (7) does not hold, (8) and (9) hold – **(DL, DL)** and **(DL, L)** exist
Region IV – (7) does not hold, (8) does not hold, (9) holds – **(DL, DL)** is unique
Region V – (7) does not hold, (9) does not hold – **(DL, L)** is unique
Region VI – (7), (8), (9), and (10) hold – All the equilibria exist
Region VII – (7) holds, (8) does not hold, (9) holds, (10) does not hold – **(L,DL)** is unique.
Region VIII – (7) holds, (8) holds, (9) holds, (10) does not hold – **(L,L)** and **(L,DL)** exist.
Region IX – (7) holds, (8) does not hold, (9) holds, (10) holds – **(L,DL)** and **(DL,DL)** exist.

First we show that if (8) doesn't hold, then (9) holds. Note that in order to prove this, it is sufficient to show that:

$$\begin{aligned}(\alpha_{L,NI} - \alpha_{NL,NI})u + [(1-\alpha_{L,NI})\beta_{L,NI} - (1-\alpha_{NL,NI})\beta_{NL,NI}](u-c) \\ > (\alpha_{L,NI} - \alpha_{NI})u + [(1-\alpha_{L,NI})\beta_{L,NI} - (1-\alpha_{NI})\beta_{NI}](u-c)\end{aligned} \tag{11}$$

This simplifies to:

$$(\alpha_{NI} - \alpha_{NL,NI})u + [(1-\alpha_{NI})\beta_{NI} - (1-\alpha_{NL,NI})\beta_{NL,NI}](u-c) > 0 \tag{12}$$

Due to positive correlation between abilities, we know that $\alpha_{NI} > \alpha_{NL,NI}$. Hence, the first term in (12) is positive. If the second term is positive as well, then we have a positive expression and are done. If, on the other hand, the second term is negative, then LHS is minimized at $c = 0$. Thus, if we can show that it holds at $c = 0$, it will hold for a positive c. Thus, it is sufficient to show that $(\alpha_{NI} - \alpha_{NL,NI}) + [(1-\alpha_{NI})\beta_{NI} - (1-\alpha_{NL,NI})\beta_{NL,NI}] > 0$. Substituting in the appropriate expressions, we have:

$$(\gamma_{NI} - \gamma_{NL,NI})(v-w)(1-q) + (\delta_{NI} - \delta_{NL,NI})(1-w)(p-q) > (\gamma_{NI}\delta_{NI} - \gamma_{NL,NI}\delta_{NL,NI})(v-w)(p-q)$$

Dividing through by $(p-q)(v-w)$, we get $(\gamma_{NI} - \gamma^{k}_{NL,NI})\frac{1-q}{p-q} + (\delta_{NI} - \delta_{NL,NI})\frac{1-w}{v-w} > \gamma_{NI}\delta_{NI} - \gamma_{NL,NI}\delta_{NL,NI}$. Since the left hand side terms are positive and $\frac{1-w}{v-w} > 1$, and $\frac{1-q}{p-q} > 1$, it is sufficient to show that $\gamma_{NI} - \gamma_{NL,NI} + \delta_{NI} - \delta_{NL,NI} - \gamma_{NI}\delta_{NI} + \gamma_{NL,NI}\delta_{NL,NI} > 0$. The left hand side is equal to $(1-\gamma_{NL,NI})(1-\delta_{NL,NI}) - (1-\gamma_{NI})(1-\delta_{NI})$ which is always positive. Second, we show that if (7) doesn't hold, then (10) holds. In order to show this, it is sufficient to show that

$$\begin{aligned}(\alpha_{L,I} - \alpha_{NL,I})u + [(1-\alpha_{L,I})\beta_{L,I} - (1-\alpha_{NL,I})\beta_{NL,I}](u-c) \quad > \\ (\alpha_{L,I} - \alpha_{I})u + [(1-\alpha_{L,I})\beta_{L,I} - (1-\alpha_{NI})\beta_{I}](u-c)\end{aligned} \tag{13}$$

Note that (13) is equivalent to (11), with all the NI subscripts substituted by I. Since the first part of the proof did not depend on the fact that the site had or had not broken the news, we can re-create the proof above with the appropriate substitutions.

Existence: Now we demonstrate values at which regions I-IX exist.

Region I – $\rho = 1,\ u = 10,\ c = 1,\ w = 0.1,\ q = 0.2,\ p = 0.22,\ v = 0.11$

Region II – $\rho = 1,\ u = 10,\ c = 1,\ w = 0.1,\ q = 0.2,\ p = 0.201,\ v = 0.1$

Region III – $\rho = 1,\ u = 10,\ c = 1,\ w = 0.1,\ q = 0.2,\ p = 0.231,\ v = 0.12$

Region IV –$\rho = 1,\ u = 10,\ c = 1,\ w = 0.1,\ q = 0.2,\ p = 0.2,\ v = 0.11$

Region V – $\rho = 1,\ u = 10,\ c = 1,\ w = 0.1,\ q = 0.2,\ p = 0.577,\ v = 0.32$

Region VI – $\rho = 1,\ u = 10,\ c = 1,\ w = 0.1,\ q = 0.2,\ p = 0.216,\ v = 0.11$

Region VII – $\rho = 0.75,\ u = 10,\ c = 1,\ w = 0.8,\ q = 0.1,\ p = 0.14,\ v = 0.94$

Region VIII – $\rho = 0.75,\ u = 10,\ c = 1,\ w = 0.8,\ q = 0.1,\ p = 0.12,\ v = 0.88$

Region IX – $\rho = 1,\ u = 10,\ c = 1,\ w = 0.1,\ q = 0.2,\ p = 0.908,\ v = 0.96$

First, we show that for $0.5 < \rho \leq \bar{\rho}$, (7) implies (8) and (9) implies (10). Let g be the LHS of (7): $g \equiv (\alpha_{L,I} - \alpha_{NL,I})u + [(1-\alpha_{L,I})\beta_{L,I} - (1-\alpha_{NL,I})\beta_{NL,I}](u-c)$, and let f be the LHS of (8): $f \equiv (\alpha_{L,NI} - \alpha_{NL,NI})u + [(1-\alpha_{L,NI})\beta_{L,NI} - (1-\alpha_{NL,NI})\beta_{NL,NI}](u-c)$. At $\rho = 0.5$, there is no correlation between the two abilities, and hence we have the same equations as in the main model:

$$g(\rho = 0.5) = (1-\alpha_I)(\beta_L - \beta_{NL})(u-c)$$

$$f(\rho = 0.5) = (1-\alpha_{NI})(\beta_L - \beta_{NL})(u-c)$$

Since $\alpha_I > \alpha_{NI}$ and $\beta_L > \beta_{NL}$, $f > g$ at $\rho = 0.5$. This, along with the fact that f and g are continuous, implies that $f > g$ for ρ close to 0.5 and there exists $\tilde{\rho} > 0.5$ s.t. for all $0.5 < \rho \leq \tilde{\rho}$, $f > g$ or (7) implies (8). We can similarly show that for ρ close to 0.5 there exists $\hat{\rho} > 0.5$ s.t. for all $0.5 < \rho \leq \hat{\rho}$, (9) implies (10). Hence, if we let $\bar{\rho} = \min(\tilde{\rho}, \hat{\rho})$, we have shown that $0.5 < \rho \leq \bar{\rho}$, (7) implies (8) and (9) implies (10).

Hence, for small positive correlation, if (8) doesn't hold then (9) holds, if (7) doesn't hold, then (10) holds (as we showed earlier) and (7) implies (8) and (9) implies (10). We now refer to the Proof for Proposition 2 in the Appendix, to show that these conditions yield Regions I – IV. **Q.E.D.**

4.2 Endogenizing the Reader's Choice Set

In the main model we assume that if a reader sees a link from A to B at stage 1b, her set of alternatives is (A, B) in period 2. That is, with a link, R's outside option is B, a blog that is more likely to be H-type on the news-breaking ability. On the other hand, without a link, the set of alternatives in period 2 is (A, C), where C is a random blog that in expectation is average on both the relevant abilities. One concern about this specification is that in either case blog C is a viable outside option since R may choose to search for another alternative even if A had linked at stage 1b. Therefore, we now consider a more general specification, where R chooses to search for another blog after observing whether A broke the news and/or posted a link to another blog in Stage 1b. The cost of additional search is k.

Proposition A2. *There exists $\underline{k} < \overline{k}$ such that for $\underline{k} \leq k \leq \overline{k}$, R chooses to search for another alternative (C) after Stage 1b if and only if A did not post a link in stage 1b.*

Proof: Intuitively, the expected benefit from search is greater when the initial set of alternatives is smaller – the marginal benefit of search for an additional alternative is greater when the initial set is (A) versus (A, B). This of course would lead to the assumed variation in the set of alternatives: the set (A, B) following a link and the set (A, C) following no link. Hence, the simpler specification of the main model can be considered an equilibrium result of a more complicated game.

In order to prove our results, we need to use some mathematical identities. In particular, we use the following results:

$$E\left[\mathrm{x}|\mathrm{x} \geq w\ \right] P\left(x \geq w\right) = \frac{\int_{w}^{\infty} xf\left(x\right)dx}{1-F\left(w\right)}\left(1-F\left(w\right)\right) = \int_{w}^{\infty}\left[1-F(x)\right]dx \tag{14}$$

In addition, if x stochastically dominates y, $F(x) \leq G(y)$ $1 - F\left(x\right) \geq 1 - G(y)$. So:

$$E\left[\mathrm{x}|\mathrm{x} \geq 0\ \right] P\left(x \geq 0\right) \geq E\left[\mathrm{y}|\mathrm{y} \geq 0\ \right] P\left(y \geq 0\right) \tag{15}$$

Finally, consider a variable z that is distributed on the same support as x and y (independent of x and y). Suppose that $F(x) \leq G(y)$. Consider the c.d.f. function $H(Z)$, where $Z = \max(z, x)$. We can show that

$$P\left(Z \leq z_0\right) = P\left(x \leq z_0\right)P\left(z \leq z_0\right) < P\left(x \leq z_0\right) \rightarrow H\left(\max\left(x, z\right)\ \right) < G(y) \tag{16}$$

Suppose that after Stage 1b and prior to Stage 2a, R may choose to search for an additional blog at a cost K. R chooses to search if this will raise her expected utility (net of search of cost) in Stage 4. That is, she chooses to add to her set of alternatives, S, if $E\left[\max_{i\in S\cup j} EU_i + \varepsilon_{i,R}\ \right] - k \geq E\left[\max_{i\in S} EU_i + \varepsilon_{i,R}\ \right]$. Let $X = EU_C + \varepsilon_{C,R} - \left(EU_A\left(1,0\right) + \varepsilon_{A,R}\right)$, $Y = EU_C + \varepsilon_{C,R} - \max\left(EU_A\left(0,1\right) + \varepsilon_{A,R}, EU_B\left(1,\cdot\right) + \varepsilon_{B,R}\right)$, where $EU_A\left(0,1\right)$ is the consumer's expected utility from A when A posts no news but links, and $EU_B\left(1,\cdot\right)$ is the expected utility from B, where by definition B broke the news. The sufficient conditions for R to search iff A does not link (after some simplifications), are the following:

$$E\left[\mathrm{X}|\mathrm{X} \geq 0\,\right] P\left(X \geq 0\right) \geq k \tag{17}$$

$$E\left[\mathrm{Y}|\mathrm{Y} \geq 0\,\right] P\left(Y \geq 0\right) < k \tag{18}$$

R searches when the marginal benefit of search (the expected increase in utility) is greater than the expected cost of search (K). Intuitively, the expected benefit from search is greater when the initial size of the set of alternatives is lower – the marginal benefit of search for an additional alternative is greater when the initial set set is (A) v (A, B). Note that since $EU_B\left(1,\cdot\right) \geq EU_A\left(1,0\right)$, and the distributions of $\varepsilon_{B,R}$ and $\varepsilon_{A,R}$ are identical, $EU_B\left(1,\cdot\right) + \varepsilon_{B,R}$ stochastically dominates $EU_A\left(1,0\right) + \varepsilon_{A,R}$. Using (16), this implies that $H_1(\max\left(EU_A\left(0,1\right) + \varepsilon_{A,R}, EU_B\left(1,\cdot\right) + \varepsilon_{B,R}\right)) < H_2\left(EU_A\left(1,0\right) + \varepsilon_{A,R}\right)$, where H_1 and H_2 are the respective cumulative distribution functions. Finally, this implies that:

$$\begin{aligned} H_3(EU_C + \varepsilon_{C,R} - EU_A(1,0) - \varepsilon_{A,R}) \quad &< \\ & H_4(EU_C + \varepsilon_{C,R} - \max(EU_A(0,1) + \varepsilon_{A,R}, EU_B(1,\cdot) + \varepsilon_{B,R}) \end{aligned}$$

Using (15), we have: $E\left[\mathrm{X}|\mathrm{X} \geq 0\,\right] P\left(X \geq 0\right) > E\left[\mathrm{Y}|\mathrm{Y} \geq 0\,\right] P\left(Y \geq 0\right)$. That is, as long as $E\left[\mathrm{Y}|\mathrm{Y} \geq 0\,\right] P\left(Y \geq 0\right) \leq k \leq E\left[\mathrm{X}|\mathrm{X} \geq 0\,\right] P\left(X \geq 0\right)$, (17) and (18) are satisfied: the consumer only searches for additional blog following no link, but doesn't search for an additional blog if a link is provided by A. **Q.E.D.**

4.3 Consumer Learning and Bloggers' Incentives

In the main model, we show that linking speeds up consumer learning compared to a scenario where linking is not in equilibrium. However, there are some limits to this finding. Notice that linking benefits the reader in two ways. First, there is a direct benefit due to information obtained through a link. Second, there is an indirect benefit due to the learning on the linked blog's type, which hurts the linking blog. Hence, a blogger may link in a way that is not welfare maximizing for the reader. We now examine this by allowing for heterogeneity in the prior on the news-breaking ability (perhaps due to differences in past actions). Specifically, we assume that A can either link to a type D blog that has a prior of γ_D on being high type on news-breaking ability or a type E blog that has a prior γ_E. Assuming that A wants to link, which blog would it choose?

To simplify this analysis, we assume that the shock to the deterministic part of the utility ($\epsilon_{j,R}$) is i.i.d., according to the double exponential distribution, with $\mu = 1$. At the end of stage 1b, R chooses among three blogs: blog A (focal blog), a type D blog, and a type E blog. Given the assumption on the distribution of the error term, the resulting choice probabilities are $P(i) = \frac{exp(EU_i)}{\sum exp(EU_j)}$, where $j \in \{A, D, E\}$. The expected surplus for R from a choice set that contains these three alternatives is $S = ln\left[exp(EU_A) + exp(EU_D) + exp(EU_E)\right]$.

Proposition A3. *Let $F(\gamma_j) = exp\left(EU_j\left(\alpha_U, \beta_O\right)|\gamma_j\right) - exp\left(EU_j\left(\alpha_O, \beta_O\right)\right)|\gamma_j)$, where blog j's prior is γ_j.*

1. F is initially increasing and then decreasing in γ_j.

2. A blogger prefers to link to a rival k such that $F(\gamma_k)$ is minimized, whereas a reader benefits most from a link to a blog m s.t. $F(\gamma_m)$ is maximized: the bloggers' and the readers' incentives are perfectly misaligned.

Proof: A link is a signal on the ability of the rival blog, and the signal is most informative when it points to a blog in the medium range of the prior distribution, where the readers' uncertainty is maximized. However, though resolving this uncertainty in favor of the rival blog is costly, linking to a blog at the ends of the distribution does not greatly change R's prior and is thus less costly. Hence, bloggers prefer to link to blogs that are obviously high or low on news-breaking ability.

Now we present the detailed proof of Proposition below.

(1) $F(\gamma_j) = \exp\left[EU\left(\alpha_U, \beta_O\right)|\gamma_j\right] - \exp\left[EU\left(\alpha_O, \beta_O\right)|\gamma_j\right] = \exp\left[\alpha_U u + (1-\alpha_U)\beta_0(u-c)\right] - \exp\left[\alpha_0 u + (1-\alpha_0)\beta_0(u-c)\right] = \exp\left(\beta_0(u-c)\right)\left[\exp(\alpha_U B) - \exp(\alpha_0 B)\right]$, where $\alpha_0 = \gamma_j v + (1-\gamma_j)w$, $\alpha_U = \frac{\gamma_j v^2 + (1-\gamma_j)w^2}{\gamma_j v + (1-\gamma_j)w}$, and $B = u - \beta_0\left(u-c\right) > 0$. This implies that $\frac{\partial \alpha_U}{\partial \gamma_j} = \frac{vw(v-w)}{(\gamma_j v + (1-\gamma_j)w)^2}$ and $\frac{\partial \alpha_0}{\partial \gamma_j} = v - w$. So:

$$F'(\gamma) = B(v-w)\exp\left(\beta_0(u-c)\right)\left\{\exp\left(\alpha_U B\right)\frac{vw}{(\gamma_j v + (1-\gamma_j)w)} - \exp(\alpha_0 B)\right\}$$

Let γ^* be such that $\frac{vw(v-w)}{(\gamma_j v + (1-\gamma_j)w)^2} = 1$. We can see that for $\gamma < \gamma^*$, $\frac{vw(v-w)}{(\gamma_j v + (1-\gamma_j)w)^2} > 1$ and $\frac{\partial \alpha_U}{\partial \gamma_j} > \frac{\partial \alpha_0}{\partial \gamma_j}$, and for $\gamma \geq \gamma^*$, $\frac{vw(v-w)}{(\gamma_j v + (1-\gamma_j)w)^2} \leq 1$ and $\frac{\partial \alpha_U}{\partial \gamma_j} \leq \frac{\partial \alpha_0}{\partial \gamma_j}$. Due to this and due to the fact that $\exp(\alpha_U B) > \exp(\alpha_0 B)$, for all $\gamma < \gamma^*$, $F'(\gamma_j) > 0$. Since $F'(\gamma = 1) < 0$, we can use Mean Value Theorem to show that $F'(\gamma) = 0$ for some $\gamma^* < \gamma < 1$.

We can also show that $\gamma > \gamma^*$, $F'(\gamma) = 0$ only once. Suppose that $\gamma^* < \gamma_1 < \gamma_2$ and $F'(\gamma_1) = F'(\gamma_2) = 0$. That is, $\exp\left(\alpha_U\left(\gamma_1\right)B\right)\frac{vw}{(\gamma_1 v + (1-\gamma_1)w)^2} = \exp\left(\alpha_0\left(\gamma_1\right)B\right)$ and
$\exp\left(\alpha_U\left(\gamma_2\right)B\right)\frac{vw}{(\gamma_2 v + (1-\gamma_2)w)^2} = \exp\left(\alpha_0\left(\gamma_2\right)B\right)$. The ratio of these two expressions yields:
$\exp\left(\alpha_U\left(\gamma_2\right)B - \alpha_U\left(\gamma_1\right)B\right)\frac{(\gamma_1 v + (1-\gamma_1)w)^2}{(\gamma_2 v + (1-\gamma_2)w)^2} = \exp\left(\alpha_0\left(\gamma_2\right)B - \alpha_0\left(\gamma_1\right)B\right)$.

Since $\gamma_1 < \gamma_2$, we can see that $\frac{(\gamma_1 v + (1-\gamma_1)w)^2}{(\gamma_2 v + (1-\gamma_2)w)^2} < 1$ and $\alpha_U\left(\gamma_2\right) - \alpha_U\left(\gamma_1\right) > \alpha_0\left(\gamma_2\right) - \alpha_0\left(\gamma_1\right)$. However, this is a contradiction since in this domain $\frac{\partial \alpha_U}{\partial \gamma_j} \leq \frac{\partial \alpha_0}{\partial \gamma_j}$. Hence, $F'(\gamma) > 0$ for $\hat{\gamma} > \gamma^*$ (where $\gamma^* < \hat{\gamma} < 1$) and $F'(\gamma) \leq 0$ for $\hat{\gamma} \geq \gamma^*$: F is strictly increasing and then strictly decreasing in γ_j.

(2) A prefers to link to D-type versus E-type if $P(A|L^{A\to D}=1) \geq P(A|L^{A\to E}=1)$ or if

$$\frac{\exp\left(EU_A\left(.\right)\right)}{\exp\left(EU_A\left(.\right)\right)+\exp\left(EU_D\left(\alpha_U,\beta_O\right)\right)+\exp\left(EU_E\left(\alpha_O,\beta_O\right)\right)} \geq \frac{\exp\left(EU_A\left(.\right)\right)}{\exp\left(EU_A\left(.\right)\right)+\exp\left(EU_D\left(\alpha_O,\beta_O\right)\right)+\exp\left(EU_E\left(\alpha_U,\beta_O\right)\right)}$$

A prefers to link to i for which $\exp\left[EU_i\left(\alpha_U,\beta_O\right)\right]-\exp\left[EU_i\left(\alpha_O,\beta_O\right)\right]$ is minimized. From the reader's perspective, the difference in value across the two regimes (A links to a type D vs. type E blog) is proportional to $\ln\left[\exp\left(EU_A\left(.\right)\right)+\exp\left(EU_D\left(\alpha_U,\beta_O\right)\right)+\exp\left(EU_E\left(\alpha_O,\beta_O\right)\right)\right]-\ln\left[\exp\left(EU_A\left(.\right)\right)+\exp\left(EU_D\left(\alpha_O,\beta_O\right)\right)+\exp\left(EU_E\left(\alpha_U,\beta_O\right)\right)\right]$. The reader prefers a link for which $\exp\left[EU_i\left(\alpha_U,\beta_O\right)\right]-\exp\left[EU_i\left(\alpha_O,\beta_O\right)\right]$ is maximized, so the reader's and blogger's incentives are perfectly misaligned. **Q.E.D.**