MANAGEMENT SCIENCE
Articles in Advance, pp. 1–18
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.1110.1510

Link to Success: How Blogs Build an Audience by Promoting Rivals

Dina Mayzlin
Marshall School of Business, University of Southern California, Los Angeles, California 40089,
mayzlin@marshall.usc.edu

Hema Yoganarasimhan
Graduate School of Management, University of California, Davis, Davis, California 95616,
hyoganarasimhan@ucdavis.edu

Empirically, we find that Web logs (or "blogs") often link to other blogs in the same category. We present an analytical model that explains why a rational blogger may choose to link to another blog. We allow bloggers to differ along two dimensions: (1) the ability to post news-breaking content, and (2) the ability to find news in other blogs. By linking, a blog signals to the reader that it will be able to direct her to news in other blogs in the future. The downside of a link is that it is a positive signal on the rival's news-breaking ability. We show that linking will be in equilibrium when the heterogeneity on the ability to break news is low relative to the heterogeneity on the ability to find news in other blogs. One implication of the linking mechanism is that blogs that are high on the news-breaking ability are more likely to gain readers. Hence, despite the fact that bloggers link for purely selfish reasons, the macro effects of this activity is that readers' learning is enhanced.

Key words: game theory; social media; linking; signaling; blogs
History: Received September 30, 2010; accepted November 8, 2011, by J. Miguel Villas-Boas, marketing.
Published online in *Articles in Advance*.

1. Introduction

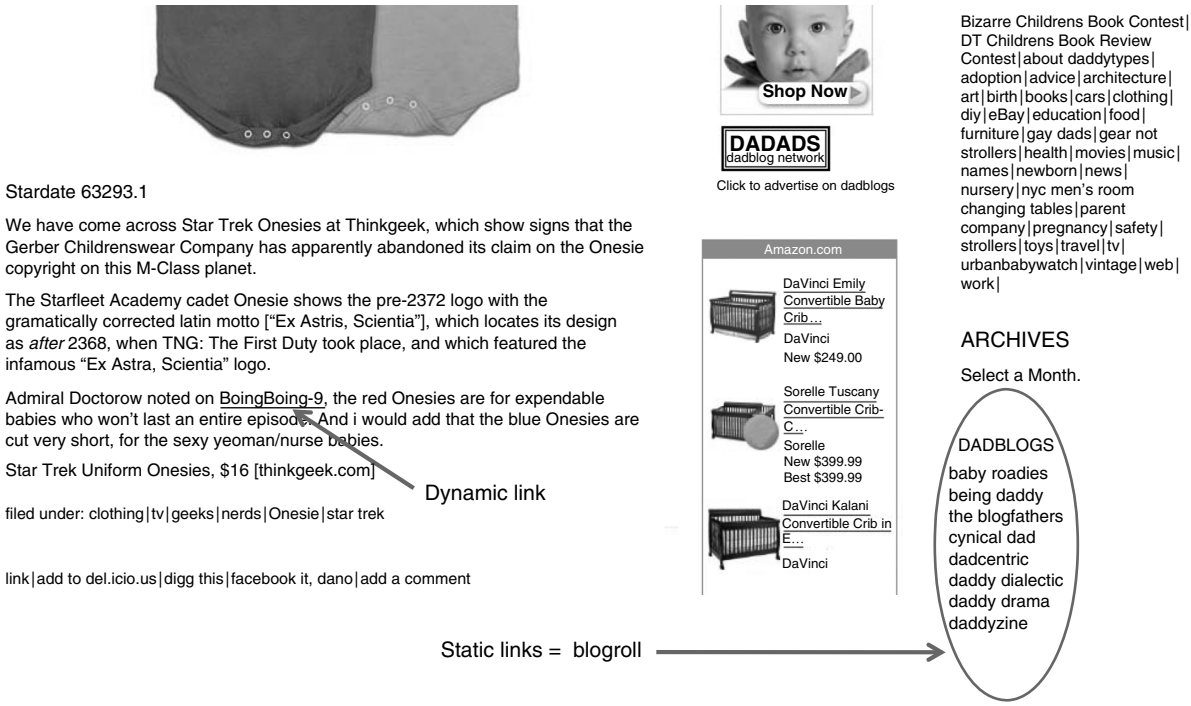
In 1994, a Swarthmore College student, Justin Hall, created an online personal journal called Links.net, now recognized as the first Web log or "blog" (Rosen 2004). Since then, creating (or "blogging") and reading blogs have become mainstream online activities. According to the Pew Internet Project, 12% of Internet users (9% of all adults) say that they blog, and 33% of Internet users (24% of all adults) say that they read blogs (Smith 2008). The growth of blogs as well as their perceived influence on purchases has motivated firms to engage with bloggers as well. For example, in a Society of Digital Agencies (2010) survey of executives from major global brands, agencies, and other major players in the digital space, 18% of respondents considered blogger outreach "top priority" and 44% considered it "important" in 2010.

Blogs are part of the larger set of online social media, which include online forums, bulletin boards, social networking sites, and video sharing sites. Although both blogs and other social media involve user-generated content, blogs also share some characteristics of newspapers. For example, blogs provide information to readers, and the mode of transmission is often one-to-many. David Winer, a blogging pioneer, gives the following definition: "A blog is like a personal newspaper . . . . It's sort of publishing on a small scale" (Potier 2003).[1]

For example, consider the blog AVC, at http://www.avc.blogs.com ("Musings of a VC [venture capitalist] in NYC"), by Fred Wilson, a partner in Union Square Ventures. The blog's posts range from the personal—"I've been in a funk for the past three days and I don't know why" (April 22, 2008)—to the general—"So why is Facebook worth $15 bn and Wordpress is worth $200 mm?" (April 18, 2008). Some posts break news, such as, "Disqus [which Union Square Ventures financed] announced a new feature release and an investment today" (March 18, 2008). Others contain information originally reported on another blog, for example, "Microsoft has apparently agreed to acquire Xobni [included a link to a post on TechCrunch]" (April 20, 2008). TechCrunch broke the story, "Two independent sources tell us that

[1] The following alternative definition of a blog is provided by Wikipedia (http://en.wikipedia.org/wiki/Blog; last modified June 27, 2012): "A blog . . . is a discussion or information site published on the World Wide Web consisting of discrete entries ('posts') typically displayed in reverse chronological order so the most recent post appears first. Until 2009 blogs were usually the work of a single individual, occasionally of a small group, and often were themed on a single subject . . . ."

Figure 1 A Snapshot of Daddytypes.com



Source. Used with permission.

the Microsoft/Xobni deal is moving along and that Microsoft signed an acquisition LOI in the last week" (April 20, 2008).

Although the nature of news-breaking events differs across domains, links to other blogs are common. (Here by "links" we mean dynamic links (or "permalinks") between blogs, which are links to specific posts in other blogs, as opposed to static links (or the "blog roll") that often appear on the right-hand side of a site (see Figure 1).) For example, on May 11, 2006, Daddytypes.com ("The Weblog for New Dads" authored by Greg Allen) posted an announcement about a two-day sale at Netto Collection, an upscale children's furniture store in Manhattan: "Looks like Netto Collection's having a sample sale. I have no idea what is there, but I do know that it's already been going for four hours… ." The post then provided a link to Daddydrama.com, which originally had posted the information on May 9, 2006, two days before Daddytypes.

In a small random sample of blogs, we found that 61% of blogs[2] contained at least one link to another site in the last 10 posts, with approximately 72% of links going to other blogs, 13% to newspaper sites, and the rest to other sites[3] (see §A.1 in the appendix for a description of the data collection method). Hence, we find that bloggers often choose to link to another blog. This is surprising on several levels. First, a reader who follows the outgoing link may not return to the original site in the short run. Second, a link implies that the linked blog has interesting content, which can improve the reader's perception of a competing site. For example, after seeing a link to the furniture sale post, the readers of Daddytypes now realize that Daddydrama can bring them useful information on sales. This of course may imply that readers will defect to Daddydrama in the future. Note that this is a concern only if sites do not already have established reputations, which is the case for most blogs to a much bigger extent than for newspaper sites. Hence, whereas all links may result in the short-term loss of an "eyeball," a link to another blog creates a stronger competitor, which may be detrimental in the long run.

One possible explanation for these links is that bloggers are irrational (or perhaps are not solely concerned about the size of their readership). However, linking may not necessarily be an irrational strategy even from an economic perspective. For example, in the same blog survey, we find that the blogs in the top quartile of subscribers are more likely to link than blogs in the bottom quartile (see Table 1). Although this anecdotal evidence does not establish causality, it does suggest that linking may not necessarily be associated with a decrease in readers.

Another possibility is that bloggers link to complementary blogs as opposed to direct competitors,

[2] Based on a sample of 258 blogs.

[3] Based on a random subsample of 438 outgoing links.

Table 1 Linking in the "Worst" vs. the "Best" Blogs

	Bottom quartile (no. of subscribers)			Top quartile (no. of subscribers)		
Category	*N*	% Blogs with outgoing links	Avg. no. of subscribers	*N*	% Blogs with outgoing links	Avg. no. of subscribers
Food	10	10	0	7	86	43
Health	16	50	0	4	75	52
Sports	23	30	0	9	78	10
Movies	13	46	0	4	100	77
Business	21	43	0	9	89	96
Music	8	50	0	4	100	331
Fashion	18	33	0	9	67	25
Politics	14	71	0	11	100	103

Notes. "No. of subscribers" is the number of people who subscribed to RSS feed of the blog through Bloglines. Note that this represents a small subset of the blog's total readership, because not all readers subscribe to RSS feeds, and Bloglines is one of many platforms that provides access to RSS feeds. "Outgoing link" is a link to another site embedded in 10 most recent posts sampled.

i.e., they link to blogs that provide information in a different category. For example, a political blog that links to a food blog faces a smaller danger of losing its readers than one that links to another political blog. However, in our sample, we find that 73% of the outgoing links to blogs are made to blogs in the same category. That is, more often than not, bloggers link to direct rivals. In this paper, we provide a theoretical explanation for this phenomenon. We explain how linking to rivals may increase a blogger's readership and explore the implications of linking on the evolution of the blogosphere.

Why would a link lead to an audience increase? As our examples demonstrate, one of the primary functions of a blog is to provide information to its readers. We focus on a particular aspect of information, namely, the ability to deliver timely news.[4] To capture the heterogeneity between blogs, we allow bloggers to differ along two dimensions: (1) the ability to post news-breaking content and (2) the ability to find news in other blogs. A blog that is higher on the ability to find news in other blogs is more likely to generate a link. Hence, a link signals to the consumer that the blogger is more likely to deliver timely news by directing her to other blogs in cases when the blogger is unable to break news on his own site. For example, Fred Wilson's (author of AVC) link to the Xobni post on TechCrunch allows him to signal that he can direct readers to interesting information posted on other blogs because of his extensive knowledge and interest in the category. Of course, a link signals the blogger's own ability to find news in other blogs, but it also signals a potential rival's news-breaking ability. For example, TechCrunch's post about the Xobni deal also demonstrates its ability to break news because of its well-placed sources. The relative benefit (positive signal about self) versus the relative cost of linking (positive signal about the other blog) determines whether a link increases a blog's audience and, hence, whether the blogger chooses to link in equilibrium. We show that linking will be in equilibrium when the heterogeneity on the ability to break news is low relative to the heterogeneity on the ability to find news in other blogs. We also show that as information "decays" at a more rapid rate (information obtained later becomes less valuable), the incentive to link decreases.

As a byproduct of the incentive to link, consumers can learn more efficiently which blogs deliver news-breaking content. Hence, despite the fact that bloggers link for purely selfish reasons on the micro level, the macro effect of this activity is that readers' learning is enhanced. Thus, through linking, blogs that are better at breaking news grow their readership more quickly than they would in the absence of linking. This of course also implies that the over-all quality of the blogosphere improves as well. This effect is further accentuated by search engines that commonly offer higher placement to sites with more incoming links.

Although the idea that incoming links contain information on the quality of a website is not a new one (the most prominent example of a model that assumes this is the Google search engine algorithm), this is the first paper that shows that linking can be incentive-compatible even in the absence of extrinsic incentives such as advertising links, for instance. That is, in our model, bloggers link because doing so improves the reader's inference about the blog's quality and ultimately increases the readership to their site. Hence, we provide an explanation for why better sites have more incoming links in equilibrium. In

[4] A survey of journalists and editors by Brodeur, a unit of Omnicom, confirms that blogs are an important source of news, even to the professionals in the media industry: 46% of respondents indicate that they find blogs helpful in getting information about breaking news, and 57% read blogs at least two or three times a week (see Brodeur 2008).

other words, this paper provides a micro foundation for models that assume that there is information in links.

We organize the remainder of this paper as follows: In §2, we discuss previous literature. We present the model setup in §3, the main results in §4, followed by extensions in §5. We conclude in §6, and we discuss some limitations and future work in §7.

2. Previous Work

We first turn to the question of why blogs may link to rivals. Katona and Sarvary (2008) investigate strategic linking online and propose a market for advertising, such that a website may sell advertising space or buy an incoming link from another site. Some similarities mark their article and the current work, in that they also find that a site with better content enjoys more incoming links. However, we differ with regard to the proposed mechanism driving this result; Katona and Sarvary (2008) focus on an explicit pricing scheme, whereas we address the role of inferences made by readers when they observe a link.

Because a link is a type of referral, the literature on referral services is a natural setting to explore the reasons behind linking. Garicano and Santos (2004) show that when an expert diagnoses a problem and decides to address it or refer it to another expert, different revenue-sharing schemes have unique implications for efficiency. Chen et al. (2002) consider infomediaries, Internet services that direct visitors to retailers that are members of their network. Both of these papers deal with an explicit contractual arrangement between the infomediary and its clients, without any inferences by clients about the infomediary's ability to refer to others. That is, here experts and sites refer to others in exchange for payment. In contrast, in our setting, there is no explicit payment structure between sites. Finally, Park (2005) examines the referral behavior of experts in repeated relationships with consumers; an expert may refer the client to another expert who is more qualified to address the client's current problem to maintain a long-term relationship with her. This broadly relates to the intuition in our model because in our model linking is also motivated by the desire to enhance a long-term relationship with a reader. However, the mechanisms in the two papers are very different. In Park (2005), an expert refers honestly because he is afraid to be punished by his customers in the future for dishonesty. In our model, on the other hand, a blogger links to signal his quality.

The literature on network formation also seeks to explain why people or firms form links. For example, Bala and Goyal (2000) and subsequent papers (see Demange and Wooders 2005 and Jackson 2008) study network formation as an equilibrium in a noncooperative game. In these papers, links are formed strategically, and the benefit of the link is to typically enhance the flow of information. That is, a link yields a direct benefit (for example, a customer may learn about a job opportunity). Despite the extensive literature in this area, to our knowledge, no research studies the issue of a third party (i.e., the reader) who makes inferences on the basis of the observed pattern of links.

Finally, link formation may be partially attributed to reciprocal giving between blogs. Resnick and Zeckhauser (2002) and Cabral and Hortaçsu (2010) find some evidence of reciprocity in buyer-seller feedback on *eBay*. Narayan and Yang (2007) find evidence of reciprocity in link formation between reviewers on Epinions, and Stephen and Toubia (2010) find evidence of reciprocal linking in an online social commerce marketplace. While reciprocity may explain some of the linking behavior, we observe linking even in situations where a reciprocal link is not expected (as would be the case for a relatively unknown blog linking to a well-known blog). Here we offer an explanation for linking that is above and beyond reciprocal giving between blogs.

Second, we show that an implication of linking is that a better "quality" site receives more incoming links in equilibrium. The idea that hyperlinks on the Internet contain information on site quality has been very influential in search engine design. Kleinberg (1999) proposed that hyperlinks offer valuable information because they reflect the subjective judgments of the author who created them. He further offered an algorithm, based on incoming links, to uncover the most authoritative webpages for a given query. Brin and Page (1998) expanded this idea to develop PageRank, a more flexible algorithm that calculates the authority rank of sites as a function of their incoming links, which continues to be the basic framework behind Google's search engine.

The assumption about the informativeness of the link structure is also analogous to the "wisdom of the crowd" hypothesis proposed by Surowiecki (2004). Even in the absence of search engines that amplify the effect of links, incoming links can increase traffic by directing people to the focal site. For example, Stephen and Toubia (2010) show that additional incoming links result in a better performance for a retailer in an online social marketplace.

In summary, the idea that sites may link to signal to a third party that they are high quality is novel to the literature. The result that this in turn leads to better sites having more incoming links, which implies that there is valuable information in links, is commonly assumed in the literature. Hence, the primary contribution of this paper is in providing a micro foundation for why we expect linking to occur in the absence of an explicit payment scheme.

3. Model

3.1. Setup

We use a finite-period game with an infinite number of risk-neutral consumers (we refer to a consumer as R, denoting reader) and an infinite number of blogs.[5] To clarify the exposition, we henceforth refer to the blogger as "he" and to the reader as "she." Bloggers obtain utility from the size of the readership. That is, the per-period utility of blogger A at time t is

$$V^A(N^t), \tag{1}$$

where N^t is the number of A's visitors during time t. Here we assume that $dV/dN > 0$: the blog's utility is increasing in the number of visitors. The blogger benefits from an increase in traffic in several ways. First, from a financial perspective, an increase in traffic results in an increase in advertising revenue. More importantly, the blogger's social utility is also increasing in site traffic because the blogger's social influence is increasing in the number of readers.[6] We also assume that all bloggers act in a way that maximizes their expected utility.[7]

Furthermore, we model bloggers as producers, and readers as consumers of information. We distinguish bloggers' abilities along the following two dimensions: (1) the ability to break news on their own site and (2) the ability to find news in other blogs. Although we initially assume that these abilities are independent, we relax this assumption subsequently in an extension. A blog can be either a high (h) type or a low (l) type with regard to breaking news[8]: h types receive it with probability v and l types with probability w, where $v > w$. The prior probability that the blogger is h type on ability to break news is γ. Thus, the prior probability that a blog breaks news, is $\alpha_0 = \gamma v + (1-\gamma)w$. The high-type's superior ability to break news derives either from its insider sources or being "in the know" through other means, such as one's social network. For example, Fred Wilson of the AVC blog, whose company has a stake in a number of start-ups, is more likely to break news compared to a blogger who engages in pure commentary.

Similarly, a blog can be either h type or l type with regard to finding news in other blogs: h types find it with probability p and l types find it with probability q (where $p > q > 0$[9] and the prior on h type is δ).[10] The prior probability that a random blog finds news in other blogs is $\beta_0 = \delta p + (1-\delta)q$. Note that being an h type here requires the knowledge of the other sites in the category, as opposed to access to specialized sources. In other words, all bloggers have access to the information in other sites, but the high-type blogger has either the ability or the desire to process the large amount of information scattered across different blogs. For example, Greg Allen of Daddytypes appears to be an avid reader of other parenting blogs, which enhances his ability to link to interesting posts elsewhere.

Hence, a blogger can be one of four types, $\varphi \in \{hh, hl, lh, ll\}$, where the first letter refers to the ability to break news on his own blog and the second letter refers to the ability to find news in other blogs. We also consider three benchmark cases where the bloggers are homogeneous along certain dimensions: (1) the case where there is only heterogeneity on the ability to find news, (2) the case where there is only heterogeneity on the ability to break news, and (3) the case where there is no heterogeneity on either dimension.

3.2. The Timeline of the Game

The game consists of two periods of two stages each (see Figure 2). Each period represents one news cycle, where the utility that the reader derives from the information depends on the speed with which it reaches her. At the beginning of the game, the bloggers know their own type, but the readers do not. Moreover, for simplicity, we also assume that bloggers do not observe their rivals' types: There is no informational asymmetry between readers and bloggers on other bloggers' quality. After observing the

[5] This technical assumption simplifies the model. Qualitatively, the results do not change as long as we assume a finite but very large number of blogs. The model does not depend on the assumption that the number of consumers is infinite.

[6] According to Lenhart and Fox (2006), 61% of bloggers listed the desire to "motivate others to action" as a reason for blogging, and 51% listed the desire to "influence the way others think" as a reason.

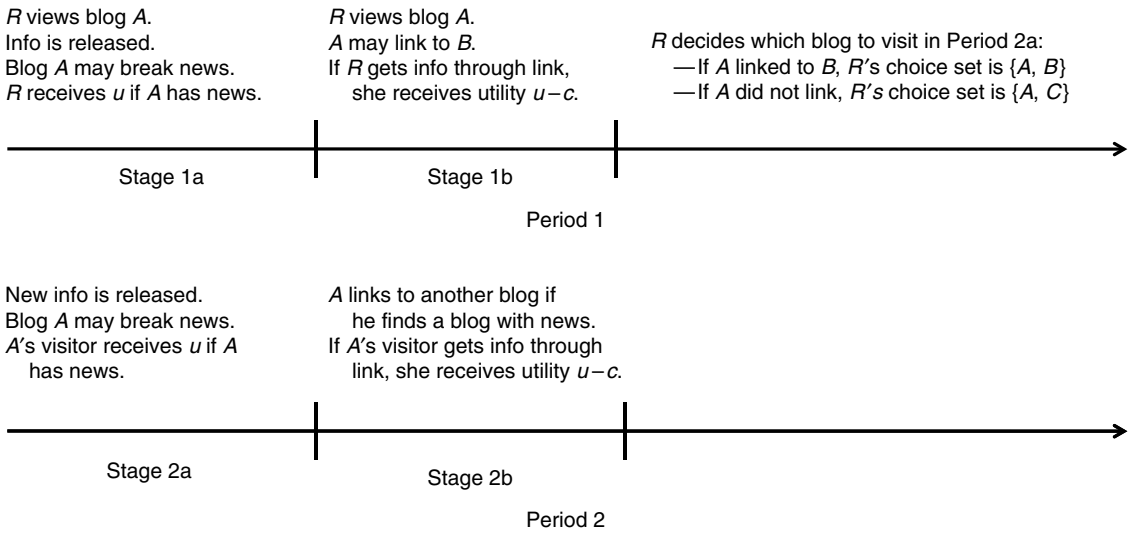
[7] Of course in reality bloggers may be partially motivated by behavioral phenomena such as altruism. Here we show that linking can occur even when bloggers are motivated solely by self-interest.

[8] Alternatively, we could differentiate bloggers according to the costs of cultivating insider sources or searching. For example, we could add an initial stage where the blogger invests a costly effort which determines the probability with which he will find breaking news in another blog, where the cost would differ across blogger types. Hence, a blogger with lower costs of finding insider sources could break news with a higher probability and a blogger with lower search costs would be more likely to find news-breaking content. Assuming differentiation across costs as opposed to probabilities of obtaining the information does not change the results qualitatively as long as $q > 0$ in equilibrium. We thank an anonymous reviewer for pointing this out.

[9] Here we assume that $q > 0$: The l type can find news in another blog with nonzero probability. This rules out a trivial separating equilibrium in linking (see the discussion following Proposition 1).

[10] The probability that a blog can find news on another blog is conditional on the event that at least one other blog breaks news. However, the probability of such an event is 1 because the number of blogs is infinite. Hence, p and q can be treated as independent of v and w.

Figure 2 Timeline of the Game



posting and linking behavior in the first period, the reader makes inferences on the blogs' types, which in turn will influence her blog choice in the second period. All readers and bloggers face the same game. To simplify the exposition, we outline the timeline of the game from the perspective of a random reader R and the blogs to which she may be exposed.

3.2.1. Period 1. At stage 1a, readers choose to visit a random blog and consume its content throughout Period 1.[11] That is, a reader (R) visits a random blog (A). Also during this stage a unique piece of verifiable information is released.[12] For example, Microsoft signs a letter of intent to purchase another company or Netto announces a furniture sale. Bloggers may gain access to the information depending on their news-breaking ability. A blogger j who obtains information may go on to post it on his blog: $a_j \in \{0, 1\}$, where $a_j = 1$ indicates the action of posting news on j's blog. We assume that, conditional on having access to breaking news, the act of posting this content is costless. We also assume that because the news is verifiable, bloggers cannot fabricate news stories. If A has posted news ($a_A = 1$), R derives utility u from the post. Otherwise, she derives zero utility from the post.

At stage 1b, bloggers search other blogs for information. Bloggers may gain access to news-breaking information on other blogs depending on their ability to find news. A blogger j who finds news in another blog may go on to post a link to that blog: $b_j \in \{0, 1\}$, where $b_j = 1$ indicates the action of posting a link. We assume that, conditional on having access to breaking news in another blog, the act of linking is costless. Reader R derives utility $u - c$ from the information if she sees a link to a news-breaking blog (say B), and the information is novel (which is the case if A had not posted news in stage 1a: $a_A = 0$). However, a reader who has seen the news at stage 1a ($a_A = 1$) does not derive any direct utility from the link, though she may learn about B's ability to break news. If the blog does not post news or link to news in another blog, we assume that R receives utility $\underline{u}$, which we normalize to 0. (See the second column of Table 2 for the summary of R's utility following A's actions in Period 1.)

Note that the value of the information declines over time; here c is the cost of delay. For example, because Daddytypes linked to the original post on the Netto Collection sales after a time delay, his readers may have already missed some of the better bargains. In short, our timeline captures the idea that original posts are more useful to consumers because they contain fresh information, unlike links, which contain relatively stale news.

Here we abstract away from the possibility that a blogger can plagiarize another blog's content without attribution. Instead, we assume that news-breaking blogs are credited by blogs who link to them. This

Table 2 Reader's Utility and Choice-Set at the End of Period 1

Blog A's action in Period 1	R's utility in Period 1	R's choice set at the end of Period 1
$a_A = 1, b_A = 1$	u	(A, B)
$a_A = 1, b_A = 0$	u	(A, C)
$a_A = 0, b_A = 1$	$u - c$	(A, B)
$a_A = 0, b_A = 0$	0	(A, C)

[11] We do not need to specify the number of readers that arrive at each blog as long as that number is finite. For example, we could assume that the number of arrivals is Poisson-distributed.

[12] All the results are unchanged if the information is released with probability $\theta < 1$.

is realistic for two reasons. First, we observe attribution.[13] Second, reputational concerns (as shown by Park 2005) may induce truth-telling. Finally, we also assume that bloggers cannot fabricate links because readers can easily verify the link's authenticity by clicking on it.

After stage 1b, R decides which blog to visit in Period 2a. If A had linked to B at stage 1b ($b_A = 1$), then R chooses between A and B. If A hadn't linked to any blog at stage 1b ($b_A = 0$), then R chooses between A and a random blog (which we denote by C). Furthermore, when making this choice, the reader also experiences a reader-blog–specific random shock to her utility. This is a technical assumption that simplifies the analysis.[14] Hence, R chooses the blog that delivers the greatest total utility, which is the sum of the expected utility and the random shock.

Note that the blogger has control in determining his competition. We model the consumer's choice as one between the focal blog (A) and a primary competitor (which may be B or C). If the blogger does not link at stage 1b, his primary competitor in Period 2 is a random blog (C), which is average in his abilities. If the blogger links, however, he makes his reader aware of a news-breaking blog (B), which becomes his competitor in the future (see the third column in Table 2). This highlights the downside of linking.[15]

To summarize, R's choice at the end of Period 1 is affected by her observations of the blogs' actions in Period 1. (See also Table 3 for summary of the information structure in different stages of Period 1.) That is, R updates her priors on A's abilities based on his posting and linking behavior. In addition, she updates her prior on B's news-breaking ability if she observes a link from A to B. To simplify the analysis, we assume that R does not update her priors about B's ability to find news in other blogs, either because she does not observe B's links (i.e., information from B may be consumed from A's post) or because she visits blog B and observes its links only after the news has become stale and the links have no signaling value.[16] We further assume that the reader does not learn about the abilities of any other blog during this time period, due to time constraints or because information quickly becomes stale in this environment.

Table 3 Information Structure During Period 1

Stage	A's information set at the *beginning* of the stage	R's information set at the *end* of the stage
0	$\Omega_0^A = \{\varphi \in \{hh, hl, lh, ll\}\}$	$\Omega_0^R = \{\alpha_0, \beta_0\}$
1a	$\Omega_{1a}^A = \{\Omega_0^A$, access to breaking news?$\}$	$\Omega_{1a}^R = \{\Omega_0^R, a_A\}$
1b	$\Omega_{1b}^A \equiv \Omega^A = \{\Omega_{1a}^A, a_A$, access to news in other blogs?$\}$	$\Omega_{1b}^R \equiv \Omega^R = \{\Omega_{1a}^R, b_A,$ $(a_B = 1$ if $b_A = 1)\}$

3.2.2. Period 2. At stage 2a, a *new* unique piece of verifiable information is released and bloggers may gain access to it depending on their types (h types with probability v and l types with probability w). Bloggers who receive the information go on to post it because there is no strategic reason to do otherwise. Reader R obtains utility u from consuming the information at this stage.

At stage 2b, blogs link to news-breaking blogs if they can find them. Here all blogs link if they find news because again, there are no strategic reasons to do otherwise. If A's visitor had not seen the news in stage 2a, she obtains utility $u - c$ from the news. Signaling in the first period is motivated by readers' desire to learn about bloggers' ability to link in the future, which in this case is the second period. The two-period model represents a simplification of an infinite-period overlapping generations model, without the added complexities of an infinite-period model.

4. Perfect Bayesian Nash Equilibrium

In our analysis we focus on the decision faced by a random reader R and a random blog A of type $\varphi \in \{hh, hl, lh, ll\}$. Given the symmetry in the readers' decisions and the bloggers' incentives, we can then generalize the findings to all blogs and all readers. The pure strategy perfect Bayesian Nash equilibrium in linking consists of the bloggers' optimal linking strategy at stage 1b as well as the readers' beliefs on the bloggers' abilities following the information received in Period 1.

We first turn to R's problem after stage 1b. We signify by Ω^R the information set of R at this

[13] For example, the Smoking Gun website received almost universal credit in the blogosphere for exposing James Frey's memoir *A Million Little Pieces* as largely fictional (The Smoking Gun 2006).

[14] There are two reasons to introduce the error term in the model. First, it explains why a blog with a negative outcome for either breaking news or linking still may attract readers in the next period. Second, the error term allows us to consider how linking affects the difference in the expected utilities between blog A and its primary rival, $EU_A - EU_j$, a continuous incentive, rather than a discrete incentive, as would be the case in a model without noise. Furthermore, the results are independent of the exact distribution of the error term. Finally, note that we could have added an error term to the utilities in the first period, too. However, it would be inconsequential because readers pick blogs randomly in the first period.

[15] Why is C not part of the choice set in the case when A links to B? We can think of this as an outcome of a more complicated game where R can invest in a (costly) search for another blog following her observation of A's linking behavior. In the online technical appendix (available at http://faculty.gsm.ucdavis.edu/~hema/blogs_tech_appendix.pdf), we show that under certain conditions R only chooses to search for another blog in the case when A does not link.

[16] The results of our analysis remain qualitatively the same if we assume that R can resolve uncertainty about B's ability to find news, but the analysis becomes much more cumbersome.

point, which consists of whether A broke the news (a_A) and whether A linked to another blog with news-breaking content (b_A): $\Omega^R = \{a_A, b_A, (a_B = 1 \text{ if } b_A = 1)\}$ (see the second column in Table 3). We denote by $\mu(\Omega^R) \equiv \{\mu^A(\Omega^R), \mu^j(\Omega^R)\} = \{[\mu^A_{hh}, \mu^A_{hl}, \mu^A_{lh}, \mu^A_{ll}], [\mu^j_{hh}, \mu^j_{hl}, \mu^j_{lh}, \mu^j_{ll}] \mid \Omega^R\}$ the vector of R's beliefs on blog A's and blog j's type, where $j = B$ if $b_A = 1$ and $j = C$ otherwise. Hence, the posterior probabilities that A will break and find the news are $\alpha_A = (\mu^A_{hh} + \mu^A_{hl})v + (\mu^A_{lh} + \mu^A_{ll})w$ and $\beta_A = (\mu^A_{hh} + \mu^A_{lh})p + (\mu^A_{hl} + \mu^A_{ll})q$, respectively. We can similarly define α_j and β_j.

Reader R's utility from blog i (where $i \in \{A, j\}$) after stage 1b is the sum of the expected utility based on R's updated beliefs about the blog's abilities and a random shock $\epsilon_{i,R}$, which we assume to be independent and identically distributed across readers and across blogs and distributed on the real line with the cumulative distribution function (CDF) F, where density is nonzero everywhere,

$$\begin{aligned} U_i^R &= EU_i(\alpha_i, \beta_i \mid \mu^A(\Omega^R)) + \epsilon_{i,R} \\ &= \alpha_i u + (1-\alpha_i)\beta_i(u-c) + \epsilon_{i,R}. \end{aligned} \tag{2}$$

Therefore, the probability that R visits blog A at stage 2a is

$$\begin{aligned} &\Pr[U_A^R > U_j^R \mid \mu(\Omega^R)] \\ &= \Pr[\epsilon_{j,R} - \epsilon_{A,R} < EU_A(\alpha_A, \beta_A \mid \mu^A(\Omega^R)) \\ &\qquad - EU_j(\alpha_j, \beta_j \mid \mu^j(\Omega^R))] \\ &= G[EU_A(\alpha_A, \beta_A \mid \mu^A(\Omega^R)) - EU_j(\alpha_j, \beta_j \mid \mu^j(\Omega^R))], \end{aligned} \tag{3}$$

where G is the CDF of the random variable $\epsilon_{j,R} - \epsilon_{A,R}$.

Second, we turn to A's optimal strategy at stage 1b. Because by assumption the blogger cannot link if he does not find news in another blogs, we focus on the scenario in which A finds news in another blog B. The blogger can condition his linking decision on his information set at this point, which contains his type (φ) and whether he broke news earlier (a_A) (see Table 3). Blogger A chooses an action (link or no link) that maximizes his expected utility in stage 2a:

$$\begin{aligned} b_A^* &= \arg\max_{b_A \in \{0,1\}} E[V \mid \Omega^A = \{\varphi, a_A\}] \\ &= \arg\max_{b_A \in \{0,1\}} E[u(N^R + N^I + N^A(\mu(\Omega^R))) \mid \Omega^A = \{\varphi, a_A\}], \end{aligned}$$

where N^R are readers who choose the blog randomly, N^I are the readers who visit A because of previous incoming links (if A had broken news at stage 1a and other blogs had linked to it at stage 1b), and N^A are A's returning readers from Period 1 (each one of whom returns with probability given in Equation (3)). Because b_A only affects the last term in A's utility function by affecting the beliefs of returning readers, henceforth, we focus on this term. Also, because we assumed that $dV/dN > 0$, we can show that A links if doing so increases the probability that it will be chosen over the primary rival in stage 2a. Furthermore, we assume that if the blogger is indifferent between linking and not linking, he chooses to link. In other words, A links if

$$\begin{aligned} &G[EU_A(\alpha_A, \beta_A \mid \mu^A(a_A, b_A = 1)) \\ &\quad - EU_B(\alpha_B, \beta_B \mid \mu^B(a_A, b_A = 1))] \\ &\quad \geq G[EU_A(\alpha_A, \beta_A \mid \mu^A(a_A, b_A = 0)) \\ &\qquad - EU_C(\alpha_C, \beta_C \mid \mu^C(a_A, b_A = 0))]. \end{aligned} \tag{4}$$

Because ϵ's density function is assumed to be nonzero on the real line, the density of $\epsilon_{j,R} - \epsilon_{A,R}$ is also nonzero on the real line. This, along with the fact that G is a CDF, implies that G is a strictly increasing function. Hence, A links if

$$\begin{aligned} &EU_A(\alpha_A, \beta_A \mid \mu^A(a_A, b_A = 1)) \\ &\quad - EU_B(\alpha_B, \beta_B \mid \mu^B(a_A, b_A = 1)) \\ &\geq EU_A(\alpha_A, \beta_A \mid \mu^A(a_A, b_A = 0)) \\ &\quad - EU_C(\alpha_C, \beta_C \mid \mu^C(a_A, b_A = 0)). \end{aligned} \tag{5}$$

Note that the linking condition in Equation (5) does not depend on a specific distribution of ϵ: The only necessary assumption is that ϵ's density function is nonzero on the real line. Intuitively, the blogger will link if this action makes him look on average more attractive than his primary rival.

Because bloggers observe their own type at the beginning of the game, in principle there could be separating equilibria where the blogger's decision to link depends on his type. However, we can show that no separating equilibria exist here.

Proposition 1. *No fully separating or semiseparating equilibria in linking exist.*

Proof. See the appendix.

The intuition behind the proof of Proposition 1 is the following. Note that given $q > 0$, even the low type can find news in another blog with nonzero probability. A blogger's utility depends solely on the size of his readership, and linking is costless. Therefore, if a link convinces the reader that the blogger is likely to be of higher ability, and hence increases the probability that the reader will come back in the future, all types of bloggers prefer to link if they find information in other blogs. Thus, separation is impossible. In contrast, if we were to assume that $q = 0$, the presence of a link could separate $\{hh, lh\}$ from $\{hl, ll\}$ because the blogger who is l type on ability to find news is never able to link.

The discussion above suggests that there are two ways to generate separation: (1) to ensure that the low type is unable to find breaking news in another blog, or (2) to introduce a cost of linking that differs by type to enable the single-crossing property. The second assumption is clearly not realistic; conditional on finding news in another blog, posting a link does not require any specialized skills. Similarly, it is difficult to believe that in reality the l type would never find news in other blogs ($q = 0$), because even minimal consumption of others' content by the blogger would result in a nonzero probability of finding news.

Proposition 1 implies that we can focus on equilibria where the decision to link depends only on whether the blogger broke news in stage 1a.[17] There are four possible pure strategy equilibria: (L, L), (L, DL), (DL, L), and (DL, DL), where L stands for link and DL indicates do not link, and the cells refer to the actions when the blog has or has not broken news, respectively. For example, in (DL, L), the blogger does not link if he broke news, but does link if he did not break news.

In equilibrium, how are R's beliefs at the end of stage 1b affected by the information received in Period 1? First, we turn to the inference on the blog's news-breaking ability. Suppose that the blogger always breaks news on his blog if he can (which we will show to be the case). Because bloggers who are h type on the ability to break news are more likely to do so than bloggers who are l type on this ability ($v > w$), a news-breaking post is a positive signal on that blogger's news-breaking ability. This implies that if R observes a news-breaking action by blog i ($a_i = 1$), she updates α_i *u*pward: $\alpha_i \mid (\mu(\Omega^R = (a_A = 1, b_A))) \equiv \alpha_U > \alpha_0$.[18] On the other hand, if $a_i = 0$, R updates α_i *d*ownward: $\alpha_i \mid (\mu(\Omega^R = (a = 0, b_A))) \equiv \alpha_D < \alpha_0$. Finally, if R has no information on news-breaking action, she holds a prior belief on the news-breaking ability, $\alpha_i \equiv \alpha_0$.

Second, we examine the inference on the blogger's ability to find news in other blogs. Because bloggers who are h type on the ability to find news are more likely to do so than bloggers who are l type on that ability ($p > q$), the presence or absence of a link could be a signal on the blogger's ability to find news in other blogs. However, as we show below, a blogger who finds news in other blogs may not always choose to link to that information. Hence, each equilibrium generates a different set of posterior beliefs following the observation of links. For example, the absence of a link in (L, L) is a negative signal about the blog's ability to find news because linking is expected in equilibrium, whereas in (DL, DL) it has no negative effect because linking is not expected. That is, if linking is not expected in equilibrium, $\beta_i \mid \mu(\Omega^R = (a_A, b_A = 0)) \equiv \beta_0$. On the other hand, if linking is expected in equilibrium, $\beta_i \mid \mu(\Omega^R = (a_A, b_A = 1)) \equiv \beta_U > \beta_0$ and $\beta_i \mid \mu(\Omega^R = (a_A, b_A = 0)) \equiv \beta_D < \beta_0$.

In addition, as is the case for most signaling models, some actions are not in equilibrium, and in this case we must specify off-equilibrium beliefs. Specifically, if linking is observed but is not played in equilibrium, Bayes' rule does not apply. There are several approaches that have been used to deal with off-equilibrium beliefs. One approach is not to make any assumptions on the set of out-of-equilibrium beliefs and to narrow it using a refinement such as the intuitive criterion (Cho and Kreps 1987) or the D1 criterion (Fudenberg and Tirole 1991). Unfortunately, here the refinements do not constrain the set of beliefs because the blogger's utility and best response do not differ by type. Hence, the approach we use here is to assume a certain set of out-of-equilibrium beliefs following a deviation (McAfee and Schwartz 1994). Because linking is only possible in the case when the blogger found news in another blog, a link perfectly signals that the blogger found news in another blog. For this reason we assume that a link is a positive signal on the blogger's ability to find news in other blogs even if linking is *not* expected in equilibrium, $\beta_i \mid \mu(\Omega^R = (a_A, b_A = 1)) \equiv \beta_U > \beta_0$. Note that this is a modified form of passive beliefs. That is, upon seeing a deviation, R assumes the prior distribution of types, but also takes into account that a link is a credible positive signal on the ability to find news in other blogs.

Finally, A's linking behavior impacts the beliefs about the rival's abilities. That is, if A links to B, R believes that $\alpha_B = \alpha_U > \alpha_0$ since the act of linking to B implies that $a_B = 1$. That is, by linking in stage 1b, A sends a positive signal on the rival's news-breaking ability. Note that because R does not have information on B's linking behavior in Period 1, her belief on B's ability to find news in other blogs in Period 2 is the prior: $\beta_B = \beta_0$. On the other hand, if A does not link to another blog in stage 1b ($b_A = 0$), his primary rival in Period 2 is a random blog (C): $\alpha_C = \alpha_0$ and $\beta_C = \beta_0$ because R does not have information on the actions of C in Period 1. Table 4 summarizes the effect of blogger's actions on R's beliefs on the two abilities for the focal blog A and the primary rival (which is

[17] Of course, it is still the case that linking could send a positive signal on the blogger's type because the h type on ability to find news is more likely to be able to link than l type. However, this signal is noisy because both types can find news with nonzero probability. Hence, here the signal is about the information that the blogger obtained in stage 1b, which in turn has implications on the blogger's type. Unlike classical signaling models, however, it is not a direct signal on the blogger's type.

[18] In the proof of Proposition 1 in the appendix, we derive α_U, α_0, α_D, β_U, β_0, and β_D as functions of the primitive parameters.

Table 4 The Effect of the Blog's Action on *R*'s Beliefs

Cases	α_A	β_A	α_{other}	β_{other}
A breaks news; *A* links to *B*	α_U	β_U	$\alpha_B = \alpha_U$	$\beta_B = \beta_0$
A breaks news; *A* does not link	α_U	β_D if (L, L) or (L, DL), else β_0	$\alpha_C = \alpha_0$	$\beta_C = \beta_0$
A does not break news; *A* links to *B*	α_D	β_U	$\alpha_B = \alpha_U$	$\beta_B = \beta_0$
A does not break news; *A* does not link	α_D	β_D if (L, L) or (DL, L), else β_0	$\alpha_C = \alpha_0$	$\beta_C = \beta_0$

B or *C* depending on *A*'s linking action in stage 1c). For example, the first row in Table 4 states that if A breaks news and links, *R* updates upward her belief on *A*'s ability to break news and to find news in other blogs, as well as *B*'s ability to break news. As we stated before, *R*'s belief on *B*'s ability to finds news in other blogs does not change.

Let us now consider the following inequalities:

$$EU_A(\alpha_U, \beta_U) - EU_B(\alpha_U, \beta_0) \geq EU_A(\alpha_U, \beta_D) - EU_C(\alpha_0, \beta_0); \quad (6)$$

$$EU_A(\alpha_D, \beta_U) - EU_B(\alpha_U, \beta_0) \geq EU_A(\alpha_D, \beta_D) - EU_C(\alpha_0, \beta_0); \quad (7)$$

$$EU_A(\alpha_D, \beta_U) - EU_B(\alpha_U, \beta_0) < EU_A(\alpha_D, \beta_0) - EU_C(\alpha_0, \beta_0); \quad (8)$$

$$EU_A(\alpha_U, \beta_U) - EU_B(\alpha_U, \beta_0) < EU_A(\alpha_U, \beta_0) - EU_C(\alpha_0, \beta_0). \quad (9)$$

Combining *R*'s equilibrium beliefs and *A*'s incentive to link (see (5)) generates the following four sets of equilibrium conditions:

- In (L, L), *A* chooses to link if it breaks news (6) and if it does not break news (7).
- In (L, DL), *A* links only if it breaks news: (6) and (8).
- In (DL, L), *A* links only if it does not break news: (9) and (7).
- In (DL, DL), *A* never chooses to link: (9) and (8).

Inequalities (6)–(9) can also be rewritten as the difference between the marginal benefit (increase in own utility) and marginal cost (increase in rival's utility) from linking; for example, (6) becomes

$$[EU_A(\alpha_U, \beta_U) - EU_A(\alpha_U, \beta_D)] - [EU_B(\alpha_U, \beta_0) - EU_C(\alpha_0, \beta_0)] \geq 0,$$

where the left-hand side can be interpreted as a measure of the incentive to link.

Last, we turn to *A*'s optimal strategy in stage 1a. We can see that in this game the blogger always chooses to post information since there is no strategic reason to do otherwise. Hence, the only strategic consideration that the blogger faces is whether to link to another blog at stage 1b. Inequalities (6)–(9) are written from the perspective of blogger *A*, but it is trivial to show that all bloggers face the same problem. Hence, these equations define linking behavior for all blogs. Similarly, the inference made by *R* describes the inference made by all readers. Hence, even though we have approached the problem from the perspective of a single blog and a single reader, we can obtain the general equilibrium of the game. We will show that the blogger links if the benefit of linking outweighs the cost. To highlight the benefit as well as the cost of linking in the main model, and their effect on the equilibrium outcome, we first present three benchmark cases where we assume that bloggers are homogeneous along certain dimensions.

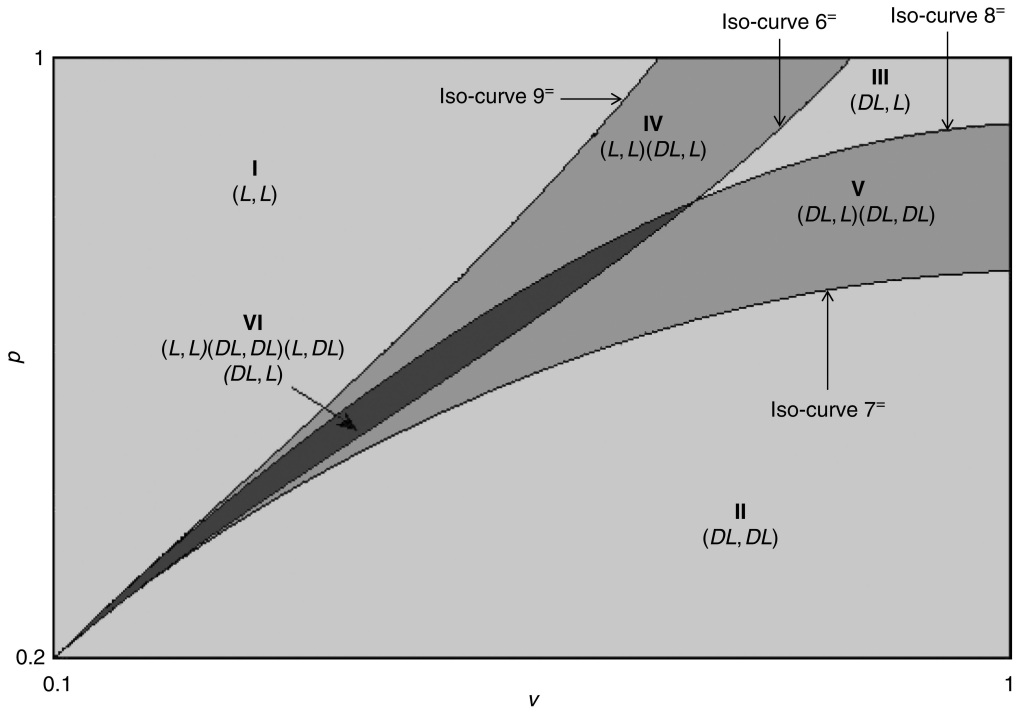
4.1. Benchmark Cases

Below we present the intuition behind the results. Please refer to the online technical appendix for a formal proof.

4.1.1. No Heterogeneity on the Ability to Break News. First, consider a modified model where there is no heterogeneity between bloggers on the ability to break news, i.e., $v = w > 0$. Because all bloggers are equally likely to break news, a news-breaking post does not provide any new information on the blog's type. Hence, there is no downside to linking: When *A* links to *B*, *R* does not draw any positive inference about *B*'s news-breaking ability. However, the benefit of linking remains: *R* updates upward her belief on *A*'s ability to find news in other blogs. Because linking in this case does not send a positive signal on the rival's ability, but still sends a positive signal about own ability to find news in other blogs, (L, L) is the unique equilibrium. This case highlights the benefit of linking for blogs: In a setting where links do not benefit the potential rival, a blogger always prefers to link.

4.1.2. No Heterogeneity on Ability to Find News in Other Blogs. Next, consider a modified model where there is no heterogeneity among bloggers on the ability to find news, i.e., $p = q > 0$. Here, *R*'s posterior beliefs on blog *A*'s ability to find news remain at β_0 regardless of whether *A* linked. That is, $\beta_U = \beta_D = \beta_0$. Because this eliminates the benefit of linking, but the cost of linking remains due to the heterogeneity on the blogs' abilities to break news, linking is not an equilibrium strategy. Hence, (DL, DL) is the unique equilibrium for all parameter values. This case highlights the cost of linking: In a setting where links do not function as signals on blogger's own abilities, there is no strategic reason for the blogger to link to his rival.

Figure 3 **Placement of Equilibria in the (*v*–*w*)–(*p*–*q*) Space** ($q = 0.2, w = 0.1, \gamma = \delta = 0.5, u = 10, c = 1$)



4.1.3. No Heterogeneity on Any Dimension. Finally, consider a scenario where there is no heterogeneity among bloggers on both the ability to break news and the ability to find news in other blogs, i.e., $v = w > 0$ and $p = q > 0$. Because all blogs are homogeneous along both dimensions, there is no updating of beliefs, i.e., $\alpha_U = \alpha_D = \alpha_0$ and $\beta_U = \beta_D = \beta_0$. This is a knife-edge case where the blogger is exactly indifferent between linking and not linking, and, by assumption, chooses to link. However, even though (L, L) is the unique equilibrium here, R does not gain any information on blog quality through the link. Moreover, any nonzero cost of linking would result in (DL, DL), which is not the case for the first benchmark case.

In sum, these three benchmark cases highlight the fact that linking is motivated by the bloggers' desire to signal own ability to find news, while it is tempered by the realization that links also act as positive signals on the linked blog's ability to break news.

4.2. Existence and Uniqueness Results of the Full Model

Next we solve for the equilibria of the full model. In Proposition (2) we summarize our findings on equilibrium existence and uniqueness. We find that (L, L), (DL, L), and (DL, DL) can all be unique, but there are also regions that contain multiple equilibria. Note that, for brevity in the proposition, we provide the conditions for uniqueness only, whereas the remaining details are provided in the appendix (§A.3).

We also graphically illustrate the results of the model in Figure 3, which depicts the relative placement of the different equilibria in the $(v$–$w)$–$(p$–$q)$ space. That is, we fix w, q and vary v, p. The boundaries of the regions in Figure 3 are defined by the iso-curves derived from Equations (6)–(9). For example, we define $6^=$ as $EU_A(\alpha_U, \beta_U) - EU_B(\alpha_U, \beta_0) = EU_A(\alpha_U, \beta_D) - EU_C(\alpha_O, \beta_0)$. In the online technical appendix, we show that there are regularities in the relative placement of the regions that span different parameter values.[19] The numbering of the regions (Region I, etc.) in Proposition (2) refers to the regions depicted in Figure 3.

Proposition 2. (Existence and Uniqueness)
Conditions for Uniqueness:

1. *Region* I: *If* $(1-\alpha_U)(\beta_U-\beta_0)(u-c) \geq (\alpha_U-\alpha_0)\cdot[(1-\beta_0)u+\beta_0 c]$, (L, L) *is unique.*
2. *Region* II: *If* $(1-\alpha_D)(\beta_U-\beta_D)(u-c) < (\alpha_U-\alpha_0)\cdot[(1-\beta_0)u+\beta_0 c]$, (DL, DL) *is unique.*
3. *Region* III: *If* $(1-\alpha_U)(\beta_U-\beta_D)(u-c) < (\alpha_U-\alpha_0)\cdot[(1-\beta_0)u+\beta_0 c] \leq (1-\alpha_D)(\beta_U-\beta_0)(u-c)$, (DL, L) *is unique.*

Regions with Multiple Equilibria:

4. *Region* IV: *The equilibria of the model are* (DL, L) *and* (L, L).
5. *Region* V: *The equilibria are* (DL, DL) *and* (DL, L).

[19] In particular, we show that $6^=$ 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 $9^=$.

6. *Region* VI: *The equilibria are* (L, L), (DL, L), (L, DL), *and* (DL, DL).

Proof. See the appendix. □

Proposition 2 and Figure 3 demonstrate the importance of the role of heterogeneity between h and l types on the existence of linking as an equilibrium strategy. By varying the p and v parameters in Figure 3, we alter the heterogeneity between types. In this section we illustrate the intuition behind the results by focusing on extreme cases—minimal versus maximal heterogeneity. For example, if v and w are very close, the heterogeneity across types on the ability to break news is minimal. The opposite is the case if v is much greater than w. Similarly, if p and q are close, there is little distinction across types on the ability to find news in other blogs, whereas the opposite is true is p is much greater than q.

From Figure 3 we can see that (L, L) is unique when p is much higher than q and v is close to w. Note that here the heterogeneity on ability to find news is low, which implies that a link to another blog would not significantly change the posterior about the rival blog's ability to break news: $\alpha_U - \alpha_0$ is small. This of course implies that the marginal cost of linking is minimal. In the limit, when $v = w$ or $\alpha_U = \alpha_0$, we obtain the benchmark case of no heterogeneity on the ability to break news, which we showed results in the unique (L, L) equilibrium. In contrast, if p is much greater than q, $\beta_U - \beta_0$ is large, and the marginal benefit of linking is high. Hence, (L, L) is unique when the marginal benefit of linking is relatively large, whereas the marginal cost is relatively low. Algebraically, we can also see this result in the condition on the uniqueness of (L, L) in Proposition 2: Linking is unique when the benefit of linking (which is a function of $\beta_U - \beta_0$) is significantly greater than the cost of linking (which is a function of $\alpha_U - \alpha_0$).

In contrast, we can see from Figure 3 that (DL, DL) is unique when p is close to q (the marginal benefit of linking is low) and v is much higher than w (the marginal cost of linking is high). Again, when p and q are close, there is little heterogeneity on the ability to find news and little difference in the posterior beliefs following link or no link: $\beta_U - \beta_D$ is small. In the limit, as β_U approaches β_D we replicate the result of the benchmark model: (DL, DL) is unique. Similarly, from the second point in Proposition 2 we can see that (DL, DL) is unique when the cost of linking (which is a function of $\alpha_U - \alpha_0$) is significantly greater than the benefit of linking (which is a function of $\beta_U - \beta_D$).

The other two equilibria, (DL, L) and (L, DL), exist when the marginal benefit and the marginal cost are close, in which case the outcome depends on the equilibrium beliefs and whether the blog breaks news (see Figure 3 and Proposition 2). Finally, as we decrease the heterogeneity along both dimensions, in the limit, only the condition for (L, L) is (trivially) satisfied at equality in Proposition 2. This is the benchmark case of no heterogeneity along both dimensions where the blogger is indifferent between linking and no linking, and R learns nothing about the blogs' types from a link. Again, because this is a knife-edge case, the result would be (DL, DL) if we introduce a nonzero cost of linking.

Because content on blogs is substitutable, a blogger who is more likely to break news earns a smaller marginal benefit from being able to direct readers to news in other blogs. Hence, a blogger should have less incentive to link when he has broken news. However, this reasoning does not take into account changes in beliefs across information states: In (L, DL), the reader's inference following "no link" is more punishing if the blog had broken the news. Because (L, DL) is the only equilibrium that defies the intuition that the incentive to link is greater if the blog had not broken the news, it is comforting to see that it is never unique.

In summary, the amount of heterogeneity on the two abilities determines whether linking is in equilibrium. That is, as we decrease the heterogeneity on the ability to break news, the marginal cost of linking decreases. On the other hand, a decrease in the heterogeneity on the ability to find news decreases the marginal benefit of linking. When linking occurs in equilibrium, the blogs that are h type on the ability to break news are more likely to attract incoming links. Thus, our model provides a micro-foundation for why incoming links may serve as an enduring measure of quality.[20] From a consumer's perspective, linking increases the attractiveness of the blogosphere, because links enable readers to locate information more efficiently. From bloggers' perspective, outgoing links enable them to signal their ability to locate information. It also enables the h type blogs to grow at a faster rate. The desire to signal generates an incentive for blogs to promote high-quality rivals, which is a byproduct of selfish behavior and not the result of altruism.

4.3. Comparative Statics

Here our objective is to clarify and reinforce the main intuition of the basic model, i.e., that the decision to link involves a trade-off between the benefit of linking (a positive signal about own ability to find news in other blogs) and the cost of linking (a positive signal about the primary rival's ability to break news).

[20] Note that the equilibrium outcome here (i.e., readers learn about the quality of the blog from the number of incoming links) is similar to that of observational learning where customers learn about the quality of the product from the number of people who purchased it (or chose not to purchase it) in the past (see Zhang 2010).

We demonstrate how a change in parameters affects the incentive to link, where the incentive to link was defined as the difference between the marginal benefit and the marginal cost of linking.

In Proposition 3, we focus on two kinds of comparative statics. First, we focus on how the changes in learning associated with posting news-breaking content and linking to another blog impact the incentive to link. From the previous section, we obtain the intuition associated with extreme cases: A large amount of heterogeneity on ability to find news compared to minimal heterogeneity on ability to break news leads to linking and vice versa. Here we ask whether a small increase in informativeness of the signal on ability to find news always increases the incentive to link and whether a small increase in informativeness of the signal on ability to break news always decreases the incentive to link. We use the following notion of informativeness: We consider the signal s' on ability to break news to be more informative than s if s' is the mean-preserving spread of s: $\alpha'_0 = \alpha_0$ and $\alpha'_U > \alpha_U > \alpha_D > \alpha'_D$ (see Kim 1995). We can also think of this increase in informativeness as *spreading the posteriors* on the signal. We can similarly define a more informative signal on ability to find news.[21] Second, we examine the effect of a change of c (the cost of delay) on the incentive to link.

Proposition 3. (Comparative Statics)

1. *A more informative signal on ability to find news always increases the incentive to link.*

2. *A more informative signal on ability to break news decreases the incentive to link if* (a) *the blog was able to break news, or* (b) *the blog was unable to break news and* $\beta_U - \beta_D$ *and* $u - c$ *are relatively low.*

3. *The incentive to link is decreasing in* c.

Proof. See the appendix.

Spreading the posteriors on the ability to find news in other blogs makes both the positive signal (a link) and the negative signal (no link) more informative. Because this increases the marginal benefit of linking and has no effect on marginal cost, the overall incentive to link increases. In contrast, the effect of spreading the posteriors on news-breaking ability has a more ambiguous effect on the incentive to link. A more informative signal improves the reader's assessment of the rival's ability and hence increases the marginal cost of linking. The effect on the marginal benefit depends on whether the blog broke news; If it did, the marginal benefit of linking decreases, because the reader infers that she is less likely to rely on a link to obtain news in the next period since she can obtain it directly from the blog. However, if the blog did not break news, the marginal benefit of the link increases, because the reader infers that she is less likely to obtain news directly from the blog. Thus, the overall effect depends on the relative importance of these two factors. In particular, when $\beta_U - \beta_D$ and $u - c$ are low, the marginal benefit of a link is relatively low, and hence the increase in cost outweighs the increase in the benefit of linking: The blog's incentive to link decreases.

Finally, we turn to the effect of a change in delay cost on the incentive to link. Increasing c decreases the utility that the reader expects to receive from information obtained through a link, which decreases the marginal benefit of linking. In addition, increasing c increases the benefit of receiving information directly versus through a link, which makes the linked blog look particularly attractive because of its news-breaking ability, which increases the marginal cost of linking. Hence, an increase in c decreases the incentive to link. Therefore, our model would predict that in categories where the timeliness of information is more important, we would observe less linking.

5. Correlated Abilities

Here we relax the simplifying assumption that the ability to break news and the ability to link are independent.[22] In particular, if there were an underlying expertise that drives both theses abilities, we would observe a positive correlation. We model this correlation across types as follows:

$$\begin{aligned} &P(h \text{ type on ability to find news} \mid h \text{ type on ability to break news}) = \rho, \\ &P(h \text{ type on ability to find news} \mid l \text{ type on ability to break news}) = 1 - \rho, \end{aligned} \tag{10}$$

where $0.5 \le \rho \le 1$.[23]

[21] Note that when we vary v and p in Figure 3, we change both the prior and posterior probabilities.

[22] In the online technical appendix, we consider two additional extensions to the main model. In the first extension we endogenize the reader's choice set by allowing her to choose to search for another random blog irrespective of A's linking decision. We find that this does not qualitatively affect the results of the main model. In the second extension, we examine the extent to which the blogger's incentive to link and the reader's incentive to learn about the bloggers' abilities are aligned. We find that the incentives on learning about the rival blog are perfectly misaligned between the blogger and the reader. That is, the blogger links in a way that minimizes the learning about the rival site, whereas the reader's utility is strictly increasing in the precision of learning.

[23] That is, since $\rho \ge 0.5$, we assume that a blog that is high type on the ability to break news is more likely to be high type on the ability to find news in other blogs. Note that at $\rho = 0.5$, the two abilities are independent; an increase in ρ implies an increase in positive correlation; and at $\rho = 1$, the two abilities are perfectly correlated. We can show that the reverse inference holds as well: A blog that is high type on the ability to find news in other blogs is more likely to be high type on the ability to break news.

Proposition 4. *There exists a* $\bar{\rho} > 0.5$, *such that*

1. *for all* $0.5 < \rho \leq \bar{\rho}$, *the equilibrium existence and uniqueness results are qualitatively the same as under independence*;

2. *for all* $\rho > \bar{\rho}$, *three additional regions* (*each with a different set of equilibria*) *become possible*, *including the region in which* (L, DL) *is unique*.

Proof. See the online technical appendix.

In the case of correlated abilities, the learning about the two abilities is coupled: A positive signal on ability to break news also results in the inference that the blogger is more likely to find news in other blogs and vice versa. This significantly complicates the trade-off that the blogger faces. For example, the blogger who posts a link also sends a positive signal on own ability to break news, an increase in the marginal benefit of linking. On the other hand, a link to another blog not only signals that a potential rival is more likely to break news in the future, but also that he is more likely to find news in another blog, an increase in the marginal cost of linking.

Interestingly, when the amount of correlation is large $(\rho > \bar{\rho})$, (L, DL) may be unique. This contradicts the earlier intuition that the incentive to link is lower when the blog can break news. The intuition behind this new result is the following: If the two abilities are highly correlated, and the blog can break news, the posterior on his ability to find news in other blogs may change in a way that increases the reader's uncertainty on this ability (which would be the case, for example, if the initial prior were low and then increased to the intermediate region where uncertainty is maximized). This in turn implies that another positive signal (a link) would result in a significant jump in the reader's belief that the blog is h type on ability to find news in other blogs. This effect may dominate the previously discussed effect that the ability to find news in other blogs is less important for a blog that is perceived as more likely to break news in the future. Hence, the incentive to link may be higher following a news-breaking story.

6. Conclusion

Empirically, linking between blogs is common. Here we provide a rational explanation for why a blogger may link to a potential rival. In the blogosphere, bloggers are providers, and readers are consumers of information. A blogger links in order to credibly signal to the reader that he is high type on the ability to find news in other blogs. By doing so, his blog can become a destination site: A reader can gain access to news-breaking content even in cases when the blogger is unable to break the news on his own. Although links are referrals to competing sites, recommending rivals may actually be profitable in certain circumstances. These individual incentives result in a system in which high-quality blogs gain prominence, which enhances consumer learning.

Several interesting marketing implications result from our findings. First, we find that both the number of incoming and outgoing links may serve as a metric of blog quality. Second, we point out the trade-offs inherent to recommending a rival. Moreover, although we focus in this paper on blogs, we believe that our findings apply to many sites that provide information as content. For example, Amazon features links to other book sellers who often have a new version of the book at a discounted price. Although Amazon takes a commission from sales generated through such links, it nonetheless chooses to facilitate a relationship between its customers and another bookstore. Hence, Amazon Marketplace allows it to act as a destination site. Another example involves the decision by WashingtonPost.com (among others) to feature links to related articles and blog posts in other publications next to the Post article (Tedeschi 2006).

What can a blogger learn from this paper? Here we suggest a framework for weighing the relative benefits and costs of linking in a context where an important function of blogs is to provide information to the readers. That is, a link demonstrates that a blog can find news in other sites, and this ability is especially valued if there is a large amount of heterogeneity along this dimension among the blogs within the category. The costs of linking arise because of the positive signal sent on the other site's ability to break news. The comparative statics presented in Proposition 3 also suggest that in a category with a great need for up-to-date information (such as sports, for example), the benefit from linking is lower.

7. Limitations and Future Research

We note several limitations to this study, which offer opportunities for further research. First, our model makes a number of simplifications. For example, we assume that blogs do not observe each other's type. Including this type of informational asymmetry (where blogs observe their own type as well as the type of their rivals) would result in a richer set of equilibria where bloggers may not post a link in cases where the linked blog is in fact of low quality, and not linking would be less of a negative signal. However, the basic trade-off between linking and not linking would remain even in this richer model. In addition, here we take a rather narrow view of information as "breaking news," which implies that information contained in two blogs (during the same period) is substitutable. We could alternatively model "information" as any content potentially useful to readers,

which would decrease the substitutability between blog content. Note that although our model requires the heterogeneity in abilities between blogs, we do not explicitly model the mechanism that gives rise to this heterogeneity because in our model the abilities of bloggers are exogenously determined. Instead we could allow bloggers to invest in both abilities and thus determine whether bloggers prefer to invest in both or to specialize.

Another limitation in our model is that we only consider links to news-breaking blogs and do not consider links to other types of content such as another blogger's set of relevant links. Allowing these other types of links would allow a link to convey information on the linked blog's ability to find news in other blogs. Hence, in the future it may be interesting to examine the incentives to link to these different types of content. Note that even in this richer model the basic trade-offs would remain the same: By linking a blogger increases the reader's perception of its own quality as well as the quality of a potential rival.

In addition to modeling the phenomena above, we see several other promising directions for further research. First, it would be interesting to model how search engines (and, in particular, news search) affect the bloggers' incentives to link. One concern may be that as the news search engine becomes very efficient at directing readers to breaking news sites, the individual blogger's ability to find news in other sites becomes less valuable to the readers, which in turn may decrease the blogger's incentive to link. Paradoxically, this may also adversely affect the performance of search engines because their algorithms rely on links as inputs. Hence, it may be interesting to investigate the equilibrium of the game between the news search engine, blogs, and the consumer. However, we do not believe that search engines can completely replace blogs in breaking news discovery because the quality of the results that a search engine produces depends on the user's ability to enter news-related search terms.[24]

Finally, there are many opportunities for empirical work in this area. Our model makes several testable predictions on when the incentives to link are greatest. In general, very little is known about the determinants of links between blogs, and between blogs and other sites, and the implications of these links. For example, some newspapers have been wary of allowing blogs to link to them without permission.[25] Richard Posner, a judge on the United States Court of Appeals, even suggested that such links be banned (*Becker-Posner Blog* 2009). Hence, empirical work in this area may have important policy implications.

[24] For example, consider a recent college graduate who has an idea for a high-tech start-up. If she is a regular reader of the AVC blog, she would learn about a 2011 New York summer accelerator program called *SeedStart Media* that nurtures start-ups. In contrast, Google would return *SeedStart* as a first page result only if she enters a specific search term such as "New York tech start-up program," but would not return it if she enters a more generic term like "tech start-up program."

Acknowledgments

The authors thank the department editor, the associate editor, and two anonymous reviewers for very constructive feedback during the review process. The authors also thank Anthony Dukes, David Godes, Yogesh Joshi, Sachin Sancheti, Jose Silva, Jiwoong Shin, K. Sudhir, and Birger Wernerfelt for comments that have improved the paper. Finally, the authors thank participants at the Yale University School of Management (SOM) Ph.D. seminar, the Marketing Science 2006 Conference, the Summer Institute in Competitive Strategy 2006 Conference, the Yale SOM Wednesday lunch seminar, the Carnegie Mellon University Heinz School seminar, the University of Minnesota marketing seminar, the University of Maryland marketing camp, the University of Massachusetts at Amherst marketing seminar, the University of Alberta marketing seminar, the Georgetown University marketing seminar, the University of Wisconsin marketing seminar, the Boston University marketing seminar, the Fourth Symposium on Statistical Challenges in Electronic Commerce Research, the University of California, San Diego marketing seminar, and the University of Southern California marketing seminar for helpful comments. Both authors contributed equally, and their names are listed in alphabetical order.

Appendix

A.1. Empirical Example

To motivate the research questions posed in the paper, we constructed a data set of blogs in eight categories: politics, health, fashion, food, business, music, sports, and movies. We used the popular blog search engine Technorati to collect a list of blogs for sampling. In particular, we used Technorati's blog directory feature. For each category, we sorted all blogs according to the recency of their updates and selected the most recently updated 150 blogs at the time of the data collection (April 19, 2007). We then collected the last 10 posts (as of June 25, 2007) for each selected blog using blog feeds available through the Bloglines aggregator. Note that this data also includes outlinks. We also used the Bloglines aggregator to collect the RSS feed subscriber data. To perform further data cleaning and classification, we used three independent raters to visit each site and decide whether it was a blog and whether it belonged in the category to which it had been tagged by Technorati. In case of conflicts, we chose the majority rating. This data cleaning and classification method was used to all the outlinks, too.

[25] For a discussion of various newspaper policies on linking, see *Online Journalism Blog* (2009).

A.2. Proof of Proposition 1

Here we present a proof by contradiction. Let us partition the set of all possible blogger types ($S = \{hh, hl, lh, ll\}$) into two nonempty sets: the set S_L and the set S_{DL}. Suppose that there exists a (semi) separating equilibrium where all the types in set S_L link if they find news and the types in the set do not link if they find news in another blog. In such an equilibrium, let $\{\alpha_{\text{Link}}, \beta_{\text{Link}}\}$ be the posteriors following a link and $\{\alpha_{\text{NoLink}}, \beta_{\text{NoLink}}\}$ be the posteriors following no link. Because the types in set S_L choose to link in equilibrium, it must be the case that for $i \in S_L$, $V_i = EU_i(\alpha_{\text{Link}}, \beta_{\text{Link}}) - EU_B(\cdot) \geq EU_i(\alpha_{\text{NoLink}}, \beta_{\text{NoLink}}) - EU_C(\cdot)$, where B is a linked blog (that was able to break news), and C is a random blog. Since V_i is independent of i and is a function of posteriors only, the trade-offs remain the same for all types. This implies that all types in the set S_{DL} would (weakly) prefer to link if they find news in another blog, which happens with nonzero probability. This contradicts the assumption that all types in the set S_{DL} do not link in equilibrium. Note that allowing the blogger to strategically post or withhold breaking news also would not enable a separating equilibrium because the same logic would hold, albeit with more complicated posteriors, such as $\{\alpha_{\text{News, Link}}, \beta_{\text{News, Link}}\}$, etc. Q.E.D.

A.3. Proof of Proposition 2

We show in the main body of the paper that for a certain set of parameters an equilibrium exists if its conditions (a subset of Equations (6)–(9) in the main body of the paper) are satisfied for that set of parameters. By substituting particular expressions for expected utilities and simplifying, these conditions can be rewritten as follows:

$$(1-\alpha_U)(\beta_U-\beta_D)(u-c) \geq (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c], \quad (6')$$

$$(1-\alpha_D)(\beta_U-\beta_D)(u-c) \geq (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c], \quad (7')$$

$$(1-\alpha_D)(\beta_U-\beta_0)(u-c) < (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c], \quad (8')$$

$$(1-\alpha_U)(\beta_U-\beta_0)(u-c) < (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c]. \quad (9')$$

Then we have
(1) (L, L): (6′) and (7′),
(2) (L, DL): (6′) and (8′),
(3) (DL, L): (9′) and (7′),
(4) (DL, DL): (9′) and (8′),
where $\alpha_U = (\gamma v^2 + (1-\gamma)w^2)/(\gamma v + (1-\gamma)w) > \alpha_0 = \gamma v + (1-\gamma)w > \alpha_D = (\gamma v(1-v) + (1-\gamma)w(1-w))/(\gamma(1-v) + (1-\gamma)(1-w))$, and $\beta_U = (\delta p^2 + (1-\delta)q^2)/(\delta p + (1-\delta)q) > \beta_0 = \delta p + (1-\delta)q > \beta_D = (\delta p(1-p) + (1-\delta)q(1-q))/(\delta(1-p) + (1-\delta)(1-q))$.

For example, (L, L) exists under a given set of parameters if both (6) and (7) are satisfied (or hold) for that set of parameters. More generally, to obtain the existence and uniqueness results, we use the four different conditions to divide the parameter space into $2^4 = 16$ nonintersecting regions, some of which may be empty. (For example, one region is defined where (6′), (7′), (8′), and (9′) hold, another where (6′), (7′), and (8′) hold but (9′) does not hold, etc.) Note that each region contains a different set of equilibria. (For example, all equilibria exist in the former region, whereas only (L, L) and (L, DL) exist in the latter.)

Even though theoretically 16 distinct regions are possible, some regions may be empty because conditions (6′)–(9′) are not independent. Note that the right-hand side (RHS) of all the equations is the same. In addition,

- The left-hand side (LHS) of (7′) > the LHS of (6′) because $(1-\alpha_D) > (1-\alpha_U)$. Hence, if (6′) holds, then (7′) holds.
- The LHS of (6′) > the LHS of (9′) because $(\beta_U - \beta_D) > (\beta_U - \beta_0)$. Hence, if (6′) does not hold, then (9′) holds.
- The LHS of (8′) > the LHS of (9′) because $(1-\alpha_D) > (1-\alpha_U)$. Hence, if (8′) holds, then (9′) holds.
- The LHS of (7′) > the LHS of (8′) because $(\beta_U - \beta_D) > (\beta_U - \beta_0)$. Hence, if (8′) does not hold, then (7′) holds.

We can show that 10 of the 16 potential regions are empty because they violate at least one of the relationshops above. For example, the region where (6′) holds and (7′) does not hold violates the first relationship. The surviving six regions partition the parameter space as follows. Figure (3) demonstrates that all of these regions are nonempty.

1. *Region* I: (6′) holds, (8′) does not hold, (9′) does not hold. As before, if (8′) does not hold, then (DL, DL) and (L, DL) do not exist. Furthermore, because (8′) does not hold, (7′) holds. Because (6′) and (7′) hold, (L, L) exists. However, because (9′) does not hold, (DL, L) does not exist. Note that if (9′) does not hold, then (6′) holds and (8′) does not hold. Thus, we can describe this region by stating that (9′) does not hold. To summarize, if $(1-\alpha_U)(\beta_U-\beta_0) \cdot (u-c) \geq (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c]$, (L, L) is unique.

2. *Region* II: (6′) does not hold, (8′) holds, (7′) does not hold. If (6′) does not hold, then (L, L) and (L, DL) do not exist. (DL, L) also does not exist because (7′) does not hold. Moreover, when (6′) does not hold, then (9′) holds. Because (8′) and (9′) hold, (DL, DL) exists. Note that if (7′) does not hold, then (6′) does not hold and (8′) holds. Thus, we can describe this region by stating that (7′) does not hold. To summarize, if $(1-\alpha_D)(\beta_U-\beta_D)(u-c) < (\alpha_U-\alpha_0) \cdot [(1-\beta_0)u+\beta_0 c]$, (DL, DL) is unique.

3. *Region* III: (6′) does not hold, (8′) does not hold. If (6′) does not hold, then (L, L) and (L, DL) do not exist. If (8′) does not hold, (DL, DL) does not exist. If (6′) does not hold, then (9′) holds. Also, if (8′) does not hold, then (7′) holds. Hence, only (DL, L) exists. To summarize, if $(1-\alpha_U) \cdot (\beta_U-\beta_D)(u-c) < (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c] \leq (1-\alpha_D) \cdot (\beta_U-\beta_0)(u-c)$, (DL, L) is unique.

4. *Region* IV: (6′) holds, (8′) does not hold, (9′) holds. If (8′) does not hold, then (DL, DL) and (L, DL) do not exist. Moreover, if (8′) does not hold, then (7′) holds. Because (6′) and (7′) hold, (L, L) exists. Similarly, because (9′) and (7′) hold, (DL, L) exists. To summarize, if $(1-\alpha_U)(\beta_U-\beta_0)(u-c) < (\alpha_U-\alpha_0) \cdot [(1-\beta_0)u+\beta_0 c] \leq \min[(1-\alpha_U)(\beta_U-\beta_D)(u-c), (1-\alpha_D) \cdot (\beta_U-\beta_0)(u-c)]$, the equilibria are (DL, L) and (L, L).

5. *Region* V: (6′) does not hold, (8′) holds, (7′) holds. If (6′) does not hold, then (L, L) and (L, DL) do not exist. Also, when (6′) does not hold, then (9′) holds. Because (8′), (7′), and (9′) hold, both (DL, DL) and (DL, L) equilibria exist. To summarize, if $\max[(1-\alpha_U)(\beta_U-\beta_D)(u-c), (1-\alpha_D)(\beta_U-\beta_0)(u-c)] < (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c] \leq (1-\alpha_D)(\beta_U-\beta_D)(u-c)$, the equilibria are (DL, DL) and (DL, L).

6. *Region* VI: (6′) holds, (8′) holds. If (6′) holds, then (7′) holds. If (8′) holds, then (9′) holds. Hence, all four equilibria exist. To summarize, if $(1-\alpha_D)(\beta_U-\beta_0)(u-c) < (\alpha_U-\alpha_0)\cdot[(1-\beta_0)u+\beta_0 c] \le (1-\alpha_U)(\beta_U-\beta_D)(u-c)$, all equilibria exist. Q.E.D.

A.4. Proof of Proposition 3

In the main body of the paper, we show that Equations (6)–(9) can also be rewritten as the incentive to link. Let f_6 denote the incentive to link when (L, L) is being played and the blog was able to break news, f_7 denote the incentive to link when (L, L) is being played and the blog was unable to break news, etc. Using the results derived in §A.3, we can show that

$$f_6 = (1-\alpha_U)(\beta_U-\beta_D)(u-c) - (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c],$$
$$f_7 = (1-\alpha_D)(\beta_U-\beta_D)(u-c) - (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c],$$
$$f_8 = (1-\alpha_D)(\beta_U-\beta_0)(u-c) - (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c],$$
$$f_9 = (1-\alpha_U)(\beta_U-\beta_0)(u-c) - (\alpha_U-\alpha_0)[(1-\beta_0)u+\beta_0 c].$$

1. Next, we demonstrate how f_6, f_7, f_8, and f_9 change as we change the parameters of the model. We first turn to the comparative statics dealing with the change in learning (point 1 and 2 in Proposition (3)). That is, we compare the incentive to link at (p, q) and $(p' = p + \varpi, q' = q - \gamma\varpi/(1-\gamma))$ as well as at (v, w) and $(v' = v + \varpi, w' = w - \delta\varpi/(1-\delta))$, where $\varpi > 0$ is a small change. This approach constitutes a *mean-preserving spread* because the priors are kept constant ($\alpha_0' = \alpha_0$, $\beta_0' = \beta_0$), and one set of posteriors brackets the other set ($\alpha_U' > \alpha_U > \alpha_D > \alpha_D'$ and $\beta_U' > \beta_U > \beta_D > \beta_D'$). Hence, according to Blackwell's theorem (Blackwell 1951, Kim 1995), the signal on ability to find news in other blogs associated with (p', q') is more informative than the signal associated with (p, q), and the signal on ability to break news associated with (v', w') is more *informative* than the signal associated with (v, w).

We use Bayes' rule to derive the posterior probabilities under (v', w') and (p', q') (the posterior probabilities under (v, w) and (p, q) were derived in A.3) as

$$\alpha_U' = \frac{\gamma(v+\varpi)^2 + \gamma(w-\gamma\varpi/(1-\gamma))^2}{\gamma v + (1-\gamma)w},$$

$$\alpha_D' = \left(\gamma(v+\varpi)(1-v-\varpi) + (1-\gamma)\left(w - \frac{\gamma\varpi}{1-\gamma}\right)\cdot\left(1 - w + \frac{\gamma\varpi}{1-\gamma}\right)\right)\cdot(1-\gamma v-(1-\gamma)w)^{-1},$$

$$\beta_U' = \frac{\delta(p+\varpi)^2 + (1-\delta)(q-\delta\varpi/(1-\delta))^2}{\delta p + (1-\delta)q},$$

$$\beta_D' = \left(\delta(p+\varpi)\left(1 - p - \frac{\delta\varpi}{1-\delta}\right) + (1-\delta)(q-\varpi)\cdot\left(1 - q + \frac{\delta\varpi}{1-\delta}\right)\right)\cdot(1-\delta p-(1-\delta)q)^{-1}.$$

Because we have a mean-preserving spread, we also know that $\alpha_U' > \alpha_U$, $\alpha_D' < \alpha_D$, $\beta_U' > \beta_U$, $\beta_D' < \beta_D$, $(\beta_U' - \beta_0') - (\beta_U - \beta_0) = (\beta_U' - \beta_U) > 0$ and $(\beta_U' - \beta_D') - (\beta_U - \beta_D) > 0$.

(a) We can easily show that the incentive to link is always higher under (p', q') than under (p, q) because

$$f_6(p', q') - f_6(p, q) = [(\beta_U' - \beta_D') - (\beta_U - \beta_D)](1-\alpha_U)(u-c) > 0,$$
$$f_7(p', q') - f_7(p, q) = [(\beta_U' - \beta_D') - (\beta_U - \beta_D)](1-\alpha_D)(u-c) > 0,$$
$$f_8(p', q') - f_8(p, q) = [(\beta_U' - \beta_0') - (\beta_U - \beta_0)](1-\alpha_D)(u-c) > 0,$$
$$f_9(p', q') - f_9(p, q) = [(\beta_U' - \beta_0') - (\beta_U - \beta_0)](1-\alpha_U)(u-c) > 0.$$

(b) Next, we compare the incentive to link under (v', w') and under (v, w):

$$\begin{aligned} f_6(v', w') - f_6(v, w) &= (\alpha_U - \alpha_U')[(\beta_U - \beta_D)(u-c) + (1-\beta_0)u + \beta_0 c] < 0, \\ f_7(v', w') - f_7(v, w) &= (\alpha_D - \alpha_D')(\beta_U - \beta_D)(u-c) + (\alpha_U - \alpha_U')((1-\beta_0)u + \beta_0 c) \\ &= K(\alpha_0(\beta_U - \beta_D)(u-c) - (1-\alpha_0)((1-\beta_0)u + \beta_0 c)), \\ f_8(v', w') - f_8(v, w) &= (\alpha_D - \alpha_D')(\beta_U - \beta_0)(u-c) + (\alpha_U - \alpha_U')((1-\beta_0)u + \beta_0 c) \\ &= K(\alpha_0(\beta_U - \beta_0)(u-c) - (1-\alpha_0)((1-\beta_0)u + \beta_0 c)), \\ f_9(v', w') - f_9(v, w) &= (\alpha_U - \alpha_U')[(\beta_U - \beta_0)(u-c) + (1-\beta_0)u + \beta_0 c] < 0, \end{aligned}$$

where $K = 2\gamma\varpi(v-w) + \gamma\varpi^2/(1-\gamma) > 0$. Note that here the sign of $f_7(v', w') - f_7(v, w)$ is the same as the sign of $\alpha_0(\beta_U - \beta_D)(u-c) - (1-\alpha_0)((1-\beta_0)u + \beta_0 c)$, and the sign of $f_8(v', w') - f_8(v, w)$ is the same as the sign of $\alpha_0(\beta_U - \beta_0)\cdot(u-c) - (1-\alpha_0)((1-\beta_0)u + \beta_0 c)$. Hence, both are negative as long as $\alpha_0(\beta_U - \beta_D)(u-c) - (1-\alpha_0)((1-\beta_0)u + \beta_0 c) < 0$.

2. Finally, $df_6/dc = -(1-\alpha_U)(\beta_U - \beta_D) - (\alpha_U - \alpha_0)\beta_0 < 0$. The same holds for f_7, f_8, and f_9. Q.E.D.

References

Bala, V., S. Goyal. 2000. A noncooperative model of network formation. *Econometrica* **68**(5) 1181–1229.

Becker-Posner Blog. 2009. The future of newspapers—Posner. (June 23), http://www.becker-posner-blog.com/2009/06/.

Blackwell. 1951. Comparison of experiments. *Proc. Second Berkeley Sympos. Math. Statist. Probab.*, University of California Press, Berkeley.

Brin, S., L. Page. 1998. The anatomy of a large-scale hypertextual Web search engine. *Comput. Networks and ISDN Systems* **30**(1–7) 107–117.

Brodeur. 2008. Journalism, social media, and the blogosphere. Accessed May 1, 2012, http://www.brodeur.com/.

Cabral, L., A. Hortaçsu. 2010. The dynamics of seller reputation: Evidence from ebay. *J. Indust. Econom.* **58**(1) 54–78.

Chen, Y., G. Iyer, V. Padmanabhan. 2002. Referral infomediaries. *Marketing Sci.* **21**(4) 412–434.

Cho, I. K., D. M. Kreps. 1987. Signaling games and stable equilibria. *Quart. J. Econom.* **102**(2) 179–221.

Demange, G., M. Wooders. 2005. *Group Formation in Economics*. Cambridge University Press, Cambridge, UK.

Fudenberg, D., J. Tirole. 1991. *Game Theory*. MIT Press, Cambridge, MA.

Garicano, L., T. Santos. 2004. Referrals. *Amer. Econom. Rev.* **94**(3) 499–525.

Jackson, M. 2008. *Social and Economic Networks*. Princeton University Press, Princeton, NJ.

Katona, Z., M. Sarvary. 2008. Network formation and the structure of the commercial World Wide Web. *Marketing Sci.* **27**(5) 764–778.

Kim, S. K. 1995. Efficiency of an information system in an agency model. *Econometrica* **63**(1) 89–102.

Kleinberg, J. 1999. Authoritative sources in a hyperlinked environment. *J. ACM* **46**(5) 604–632.

Lenhart, A., S. Fox. 2006. Bloggers: A portrait of the Internet's new storytellers. Pew Internet and American life project. *Pew Internet* (July 19), http://www.pewinternet.org/~/media/Files/Reports/2006/PIP%20Bloggers%20Report%20July%2019%202006.pdf.pdf.

McAfee, R. P., M. Schwartz. 1994. Opportunism in multilateral vertical contracting: Nondiscrimination, exclusivity, and uniformity. *Amer. Econom. Rev.* **84**(1) 210–230.

Narayan, V., S. Yang. 2007. Trust between consumers in online communities: Modeling the formation of dyadic relationships. Working paper, New York University, New York.

Online Journalism Blog. 2009. Newspaper sites: Do not link to us. (March 11), http://onlinejournalismblog.com/2009/03/11/newspapers-dont-link/.

Park, I. 2005. Cheap-talk referrals of differentiated experts in repeated relationships. *RAND J. Econom.* **36**(2) 391–411.

Potier, B. 2003. Berkman Center fellow Dave Winer wants to get Harvard blogging: Weblog pioneer preaches the gospel of blog. *Harvard University Gazette* (April 17), http://news.harvard.edu/gazette/2003/04.17/13-blogging.html.

Resnick, P., R. Zeckhauser. 2002. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. M. R. Baye, ed. *Economic Internet and E-Commerce*. Advances in Applied Microeconomics, Vol. 11. Emerald Group, Bingley, UK, 127–157.

Rosen, J. 2004. Your blog or mine? *New York Times Magazine* (December 19), http://www.nytimes.com/2004/12/19/magazine/19PHENOM.html?pagewanted=1.

Smith, A. 2008. New numbers for blogging and blog readership. *Pew Internet* (July 22), http://www.pewinternet.org/Commentary/2008/July/New-numbers-for-blogging-and-blog-readership.aspx.

Society of Digital Agencies. 2010. 2010 SoDA digital marketing outlook. Accessed June 15, 2012, http://dmroundtable.com/page/2010-soda-digital-marketing.

Stephen, A. T., O. Toubia. 2010. Deriving value from social commerce networks. *J. Marketing Res.* **47**(2) 215–228.

Surowiecki, J. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. Doubleday, New York.

Tedeschi, B. 2006. Newspapers to use links to rivals on Web sites. *New York Times* (July 31), http://www.nytimes.com./2006/07/31/technology/3lecom.html.

The Smoking Gun. 2006. A million little lies: Exposing James Frey's fiction addiction. (January 4), http://www.thesmokinggun.com/documents/celebrity/million-little-lies.

Zhang, J. 2010. The sound of silence: Observational learning in the U.S. kidney market. *Marketing Sci.* **29**(2) 315–335.