

Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms

Chrysanthos Dellarocas

R. H. Smith School of Business, University of Maryland, 4341 Van Munching Hall,
College Park, Maryland 20742, cdell@rhsmith.umd.edu

There is growing evidence that consumers are influenced by Internet-based opinion forums before making a variety of purchase decisions. Firms whose products are being discussed in such forums are therefore tempted to manipulate consumer perceptions by posting costly anonymous messages that praise their products. This paper offers a theoretical analysis of the impact of such behavior on firm profits and consumer surplus. There are three main results. First, if every firm's manipulation strategy is a monotonically increasing (decreasing) function of that firm's true quality, strategic manipulation of online forums increases (decreases) the information value of a forum to consumers. This result implies the existence of settings where online forum manipulation benefits consumers. Second, equilibria where strategies are monotonically increasing (decreasing) functions of a firm's true quality exist in settings where the firm's net payoff function, inclusive of the cost of manipulation, is supermodular (submodular) in the firm's quality and manipulation action. Third, in a broad class of settings, if the precision of honest consumer opinions that firms manipulate is sufficiently high, firms of all types, as well as society, would be strictly better off if manipulation of online forums was not possible. Nonetheless, firms are locked into a "rat race" and forced to spend resources on such profit-reducing activities; if they don't, consumer perceptions will be biased against them. The social cost of online manipulation can be reduced by developing "filtering" technologies that make it costlier for firms to manipulate. Interestingly, as the amount of user-contributed online content increases, it is firms, and not consumers, that have most to gain from the development of such technologies.

Key words: online product reviews; electronic commerce; manipulation; word of mouth

History: Accepted by Ramayya Krishnan, information systems; received August 26, 2004. This paper was with the author 3 months for 2 revisions.

1. Introduction

The Internet has enabled individuals all over the world to make their personal experiences, thoughts, and opinions easily accessible to the global community "at the click of a mouse." This has led to the creation of a diverse mosaic of online word-of-mouth communities (online forums), where individuals exchange experiences and opinions on a variety of topics ranging from products and services, to politics and world events. Examples of such communities include online product review forums, Internet discussion groups, instant messaging chat rooms, mailing lists, and Web logs (see Schindler and Bickart 2003 for a nice overview).

There is growing evidence that consumers are influenced by opinions posted in online forums before making a variety of purchase decisions (Thompson 2003, Senecal and Nantel 2004, Chevalier and Mayzlin 2006). Similar evidence suggests that forums play an increasingly important role in public opinion formation. Internet forums are thus emerging as an alternative source of information to mainstream mass media, replacing our societies' traditional reliance on the

"wisdom of the specialist" by the "knowledge of the many."

Many have argued that the ease with which the Internet can aggregate information from large numbers of, otherwise unrelated, individuals can lead to better informed and, ultimately, more efficient markets and societies. Nevertheless, these same properties (large scale, relative anonymity) make it relatively easy for interested parties to manipulate the information propagated through online forums by anonymously adding their own, strategically biased, messages to the total mix of posted opinions.

Online forum manipulation strategies can take many forms, and firms (or depending on the context of interest, political parties and special interest groups) are getting more sophisticated by the day. The simplest firm strategy is to anonymously post online reviews praising its own products, or bad-mouthing those of its competitors. There is ample evidence that such manipulation occurs. For example, when in February 2004, because of a software error, Amazon.com's Canadian site mistakenly revealed the true identities of some of its book reviewers, it turned out that a

sizable proportion of these reviews were written by the books' own publishers, authors, and competitors (Harmon 2004). The music industry is known to hire professional marketers who surf various online chat rooms and fan sites to post positive opinions on behalf of new albums (White 1999, Mayzlin 2006).

Given the potential backlash of such activities, firms are experimenting with less overt methods. Some firms offer rewards to consumers who start favorable conversations about their products on popular online forums. For example, a recent marketing campaign promised prizes to fans who would start conversations in online forums praising singer Lucinda Williams's albums (see <http://slate.msn.com>, July 26, 2001). Other firms routinely monitor online forums to identify influential community members. They then target them directly and try to persuade them to write favorable reviews by sending them free samples, inviting them to special events, etc. There is at least one professional marketing firm that conducts such campaigns on behalf of its clients (see <http://www.electricartists.com>).¹

As more firms, political parties, and special interest groups realize the power of online forums, it is expected that more will engage in direct or indirect word-of-mouth manipulation practices.

It is, therefore, important and timely to understand what the impact of such activity is likely to be on the informativeness of Internet forums and on the payoffs to the various parties who are affected by them. The results of such analyses will be relevant to policy decisions (Is Internet forum manipulation socially harmful?), R&D decisions (Does it pay to invest in technologies that discourage online manipulation? Who should bear the cost of such investments?), and, of course, firm and consumer attitudes toward Internet forums (How much should consumers trust online forums? How much should firms invest in trying to manipulate them?).

This paper contributes to answering some of these questions by analyzing how the strategic manipulation of Internet opinion forums affects the payoffs of consumers and firms in markets of vertically differentiated experience goods. Specifically, I consider settings where one or more firms simultaneously launch products whose true quality is initially unknown to consumers and difficult to verify before purchase and use. I assume that the main source of quality information for consumers is an online product review

forum (such as Epinions.com or Amazon Reviews), where past consumers post opinions about their experiences with the goods. New consumers read these opinions and form perceptions about the qualities of the products. Based on these perceptions, they make purchase decisions. Firms can try to manipulate consumer perceptions by posting anonymous reviews that praise their own product, at a cost. All firms are assumed to be strategic; that is, they manipulate opinion forums to maximize their payoff, given their correct anticipation of other firms' strategies and consumer beliefs. Furthermore, consumers are smart; even though they cannot directly distinguish honest opinions from fake opinions, they are aware that manipulation occurs and adjust their interpretation of online opinions accordingly.

My analysis derives three principal results. First, if every firm's manipulation strategy (amount by which the firm inflates its true ratings) is monotonically increasing in that firm's true quality, manipulation *increases* the informativeness of online forums, in the sense of increasing the ex ante expected payoff of consumers who base their decisions on information published in these forums. In such settings, high-quality firms inflate their, already higher, true ratings more than low-quality firms. Manipulation activity then increases the separation of the probability distributions of ratings that correspond to adjacent product qualities. This allows consumers to make more accurate inferences about a firm's true quality from its published ratings. The inverse result holds in settings where every firm's manipulation strategy is monotonically decreasing in that firm's true quality. In such settings, low-quality firms manipulate more than high-quality firms; i.e., they are shrinking the gap between their respective ratings and making it more difficult for consumers to infer a firm's true quality (because everybody's ratings will be clustered together). Manipulation then *decreases* the informativeness of online forums, hurting consumers.

Second, informativeness-enhancing manipulation equilibria exist in settings where every firm's net payoff function, inclusive of the cost of manipulation, is supermodular in firm type and manipulation action. In a broad class of settings, net firm payoffs are supermodular if firm profits (before manipulation costs) are sufficiently steeply increasing convex functions of consumer perceptions of their quality. In such settings, the better a firm is perceived to be, the more it has to gain from being perceived to be even better. This provides high-quality firms with higher incentives to inflate their ratings. In contrast, in settings where firm profits are concave functions of consumer perceptions of their quality, the better a firm is perceived to be, the less it has to gain from being perceived to be even better. It is then low-quality firms

¹ Word-of-mouth manipulation practices are not confined to the online domain. The anonymity of modern urban societies is prompting several marketing firms to experiment with word-of-mouth marketing campaigns in which paid individuals circulate large cities, endorsing a firm's products to strangers as if they were expressing their personal opinions (Walker 2004). The models and results of this paper apply to these contexts as well.

that have higher incentives to inflate their ratings, leading to equilibria where manipulation decreases informativeness.

Third, in a broad class of settings, if the cumulative precision of honest ratings is sufficiently high (for example, because a sufficiently large number of consumers post honest opinions online), the cost of manipulation to firms always outweighs its benefits. A high precision baseline signal cannot be substantially affected by firm manipulation. Resources spent on manipulation are then wasted, because they do not change consumer beliefs. Nevertheless, firms have no choice. The fact that anonymous manipulation is possible induces rational consumers to anticipate that firms will engage in it (and thus to appropriately discount the nominal values of online ratings they observe). Firms are then trapped into performing the equilibrium level of manipulation that is expected of them, because as in a rat race, a lower level will bias consumers' perceptions against them.

The overall picture painted by these results has interesting, and somewhat counterintuitive, implications for practice. On the one hand, my analysis shows the existence of settings where forum manipulation is equivalent to a form of *quality signaling* that benefits consumers. On the other hand, it shows that if consumers come to expect that firms will manipulate, as the volume and quality of user-generated online content increases, there will be a threshold beyond which firms will be trapped into having to engage in profit-reducing online manipulation practices, simply because consumers expect them to. All firms would then be better off if consumers (rationally) expected them to manipulate less. My analysis shows that one way of accomplishing this is by developing filtering technologies that make it costlier for firms to manipulate, and thus lower the equilibrium levels of manipulation. Interestingly, it is firms, and not consumers, that have most to gain from the development of such technologies.

The rest of this paper is organized as follows. Section 2 introduces the main intuitions by analyzing a set of simple monopoly settings. Section 3 shows how the results generalize for a broad class of payoff functions and signal distributions. Section 4 discusses related work. Finally, §5 summarizes the strategic implications of our findings for consumers, firms, and forum operators and concludes. An online appendix, available on the *Management Science* website at <http://mansci.pubs.informs.org/ecompanion.html>, provides several modeling extensions. Table 1 summarizes the key notation used throughout this paper.

2. The Main Intuitions

This section introduces the main ideas underlying this work by analyzing a series of simple monopoly set-

Table 1 Key Notation Used Throughout the Paper

Symbol	Meaning
c	Manipulation cost
D	Demand
g	Constant component of a firm's manipulation strategy
h	Type-dependent component of a firm's manipulation strategy
p	Price
q	Firm type (quality)
u	Consumer utility
v	Firm net profit function (inclusive of manipulation costs)
w	Firm sales profit function (exclusive of manipulation costs)
x	Baseline word-of-mouth signal (average value of honest ratings)
y	Observable signal (average value of honest plus fake ratings)
η	Manipulation strategy
θ	Mean of consumer posterior beliefs (perceived quality)
λ	Unit cost of manipulation
ρ_x	Precision of baseline signal
ρ_z	Precision of adjusted observable signal (used in consumer inference)
τ	Precision of prior beliefs

tings. The emphasis is on conveying the fundamental intuitions using examples that admit closed-form solutions. Section 3 then states the general form of the results in multifirm settings.

Consider a monopolist firm that introduces a product to a new market. The appeal of the product to consumers is the sum of two independent components: (1) a horizontal component (*location*), representing product attributes whose valuation depends on each individual consumer's taste (e.g., color, shape, look and feel, etc.) and (2) a vertical component (*quality*), representing attributes whose valuation is identical among all consumers (e.g., ease of use, durability, etc.). I assume that a product's location can be reliably communicated to consumers, whereas a product's true quality q can only become known after the good is bought and consumed.²

Consumers are uniformly distributed in the unit interval $[0, 1]$ and have quadratic transportation costs. According to standard theory, the monopolist will locate his product at the interval's center. A consumer's utility from consuming a product of quality q is then given by

$$u^i = f(q) - \frac{16}{27} \left(i - \frac{1}{2} \right)^2 - p,$$

where $f(\cdot)$ is an arbitrary nonnegative function, p is the product's price, and $i \in [0, 1]$ is the consumer's location in the unit interval. The scaling factor $16/27$

² Throughout the paper, I will refer to parameter q as the product's quality. Keep in mind, however, that the model is independent of the specific interpretation of q ; the main ideas apply to any situation regarding some aspect of their products (e.g., the established user base of a software application in a setting with network externalities) that matters to consumers and cannot be easily verified before purchase and use.

simply serves to simplify the final expression of equilibrium firm profits. Expected utility maximization and price-taking behavior imply the following demand function:

$$D = \frac{3\sqrt{3}}{2} \sqrt{f(q) - p}.$$

Variable costs are assumed to be zero or, alternatively, marginal costs are constant and prices are defined net of marginal costs. Firm profits are then simply equal to sales revenues $w = Dp$. If q is common knowledge, then profit maximization would imply price $p = 2f(q)/3$, demand $D = 3\sqrt{f(q)}/2$, and sales revenues $w = [f(q)]^{3/2}$.

In our setting, q is known to the firm but not to consumers. Consumers share a common prior regarding q . The prior is normally distributed with mean m and precision τ . In addition, consumers have access to an exogenously generated, normally distributed signal x of the product’s true quality with mean q , precision ρ_x , and full support. In the context of this paper, this signal can be thought of as the arithmetic mean of online ratings posted by consumers who have already tried the product.³ The signal’s precision then is the sum of precisions of individual ratings. The firm knows the parameters of the consumers’ prior and signal. However, it does not know the exact realization of the signal until it has been published by the forum.

Throughout this paper, I am treating word of mouth as an exogenous phenomenon, and do not make an attempt to justify why consumers engage in this costly activity on the basis of economic grounds. This perspective is consistent with a large body of empirical evidence (Dichter 1966, Engel et al. 1993, Sundaram et al. 1998, Hennig-Thurau et al. 2004) that has identified a variety of extraeconomic motivations to explain why consumers engage in (offline and online) word of mouth (desire to achieve social status, utility from engaging in social interaction, altruism, concern for others, easing anger, dissonance reduction, vengeance, etc.). In the context of information systems research, see Dellarocas et al. (2003) and Gu and Jarvenpaa (2003) for empirical studies of extraeconomic drivers of consumer contributions in online forums.

All consumers (and the firm) have access to the same realization of signal x . Let $\theta = E[q | x]$ denote the mean of consumer posterior beliefs regarding the product’s quality after they observe x . I will refer to θ as the firm’s *perceived quality*. If the firm sets its price after consumers observe signal x , no signaling

³ Alternative interpretations of signal x are also possible. For example, in the context of an online forum that solicits text reviews, x can be thought of as the mean of the judgments contained in all text reviews read by consumers. In an offline context, x could be the average of all opinions “heard on the street.”

of quality through price is possible. Maximization of expected sales revenues then implies price $p = 2f(\theta)/3$ and sales revenues $w = [f(\theta)]^{3/2}$.

Taking advantage of the anonymity of the online medium, the monopolist can manipulate signal x by posting fake anonymous ratings that praise its product or by providing incentives to past consumers (who would otherwise not have posted online ratings) to do so. This way the monopolist can attempt to increase consumer perceptions of its quality by shifting the mean of the distribution of average ratings from q to $q + \eta$ at cost $c(\eta) = \lambda\eta^2$. The parameter λ captures the unit cost of manipulation.

Denote by y the signal that results from manipulation of the original signal x . The motivating question of this paper is to understand how manipulation affects (1) the information quality of signal y (relative to x) and (2) the profits of the firm.

2.1. Manipulation and Informativeness

We begin the analysis by assuming that $f(\theta) = \theta^{4/3}$. Such a function characterizes settings where consumers have increasing marginal returns from quality. In such settings, sales revenues $w = [f(\theta)]^{3/2} = \theta^2$ are convex with respect to the firm’s perceived quality. We assume that consumers cannot distinguish between honest and fake online ratings. Nevertheless, they are aware that the monopolist manipulates online ratings and adjust their posterior beliefs accordingly. The notion of equilibrium that applies to this setting is a perfect Bayesian equilibrium (PBE). In such an equilibrium, the firm maximizes its expected payoffs given the buyers’ beliefs about its manipulation strategy, and buyer beliefs are consistent (in the Bayesian sense) with the firm’s strategy. Buyer beliefs are obtained from equilibrium strategies and observations using Bayes’ rule.

I will show the existence of PBE in linear strategies. Suppose that consumers conjecture that the amount by which the firm inflates its online quality ratings is a linear function $\eta(q) = g + hq$ of the firm’s true quality, where g, h are real numbers that at equilibrium, correspond to correct conjectures. Consumers then understand that the ratings they observe are the sum of three (indistinguishable) components

$$y = q + \eta(q) + \varepsilon = q + (g + hq) + \varepsilon,$$

where q is the product’s true quality and ε is a normally distributed noise term with mean zero and precision ρ_x . From the above expression it follows that:

$$q = \frac{y - g}{h + 1} - \frac{\varepsilon}{h + 1}.$$

In other words, the publicly observable statistic $z = (y - g)/(h + 1)$ is a normally distributed unbiased estimator of q with precision $\rho_z = \rho_x(h + 1)^2$. If consumers

update their beliefs using Bayes rule, given the normality of prior beliefs and signal z , standard theory (DeGroot 1970) predicts that their posterior beliefs about q will be normally distributed with mean

$$\theta = \frac{\tau m + \rho_z(y - g)/(h + 1)}{\tau + \rho_z}.$$

The measure of information quality that I will use in this section is the signal's precision.⁴ The impact of manipulation on the precision of signal y (relative to x) depends on the value of the parameter h . If $h > 0$, then $\rho_z = \rho_x(h + 1)^2 > \rho_x$; in such cases, firm manipulation *increases* the precision of the original quality signal x . Positive h means that a firm's manipulation strategy is a monotonically increasing function of its true quality: higher quality firms inflate their (already higher) ratings more than lower quality firms. Manipulation activity then spreads out the means $q + \eta(q)$ of the signal distributions that correspond to adjacent qualities q . As a result, the distributions of signal y become less *crowded* than the distributions of the original quality signal x . This makes the probabilistic mapping between an observed signal and the underlying firm type more reliable and allows consumers to make more accurate inferences about the firm's true quality from published ratings. On the other hand, if $-1 < h < 0$, then $\rho_z < \rho_x$; manipulation *decreases* the precision of the original quality signal x . Negative h means that a firm's manipulation strategy is a monotonically decreasing function of its quality: lower quality firms manipulate more than higher quality firms, shrinking the gap between their respective ratings. The distributions of signal y then become more crowded than the distributions of the original signal x , making the probabilistic mapping between signals and types less reliable and reducing the ability of consumers to make inferences about firm quality from online ratings.⁵

The special case $h = 0$ corresponds to situations where the amount of manipulation is a constant g independently of the firm's quality. Consumers then can recover the original quality signal by subtracting g from the observed signal y . The precision of the signal remains unchanged. On the other hand, the special case $h = -1$ corresponds to situations where the

amount of manipulation $\eta(q) = g - q$ is inversely proportional to the firm's quality. The observable signal $y = q + (g - q) + \varepsilon = g + \varepsilon$ then carries no information about quality: manipulation activity completely cancels whatever quality information was contained in the original signal x , resulting in observable signals y whose distribution is identical for all firm qualities. In such settings $\rho_z = 0$: if consumers expect such firm behavior, they will ignore online ratings.

The equilibrium values of g and h can be computed by solving the firm's optimization problem. Assume that the firm decides its manipulation strategy before it sees the realization of signal x (e.g., before the forum publishes the average ratings posted by honest consumers), but sets its price after the forum publishes signal y . At the time of manipulation, the firm's expected revenues are

$$\begin{aligned} E[w(y | q, \eta)] &= E\left[\left(\frac{\tau m + \rho_z(y - g)/(h + 1)}{\tau + \rho_z}\right)^2 \middle| q, \eta\right] \\ &= \left(\frac{\tau m + \rho_z(q + \eta - g)/(h + 1)}{\tau + \rho_z}\right)^2 + \frac{\rho_z}{(\tau + \rho_z)^2}. \end{aligned}$$

The firm's objective is to select η to maximize

$$\begin{aligned} v &= -c(\eta) + E[w(\theta | q, \eta)] \\ &= -\lambda\eta^2 + \left(\frac{\tau m + \rho_z(q + \eta - g)/(h + 1)}{\tau + \rho_z}\right)^2 + \frac{\rho_z}{(\tau + \rho_z)^2}. \end{aligned} \quad (1)$$

The following proposition provides details of the solution.

PROPOSITION 1. *Let θ be the mean of the consumers' posterior beliefs about the firm's quality. If sales revenues are equal to $w = \theta^2$ and manipulation cost is equal to $c(\eta) = \lambda\eta^2$, there exists linear PBE where the firm's manipulation strategy is a linear function of its true quality $\eta = g + hq$, and where*

$$g = \frac{\tau h}{\rho_x(h + 1)^2} m$$

and h is a positive real solution of the fifth-degree polynomial equation

$$h = \frac{\rho_x^2(h + 1)^3}{(\tau + \rho_x(h + 1)^2)^2 \lambda}. \quad (2)$$

The most important part of Proposition 1 is that h is positive for all values of parameters τ , ρ_x and λ . In other words, in all linear PBE, higher quality firms inflate their (already higher) ratings more than lower quality firms; manipulation activity then increases the precision of the quality signal.

⁴ Section 3 reformulates all results using the more general notion of Blackwell informativeness.

⁵ The case $h < -1$ corresponds to situations where low-quality firms manipulate so much more than high-quality firms that the ordering of the resulting online ratings is reversed: higher average ratings y then correspond to lower quality firms. In our specification, rational consumers would anticipate such firm behavior, and would, thus, not be fooled. One, therefore, expects that firms would never find it optimal to engage in such behavior. Our results confirm that this case can never occur at equilibrium.

The main drivers of this result are the facts that (1) the firm’s sales revenues $w = \theta^2$ are a convex function of its perceived quality θ and (2) the effects of manipulation are additive to an exogenous signal (honest ratings) that is centered on the firm’s true quality q . Thanks to the presence of the exogenous quality signal x , in the absence of manipulation, a firm’s expected sales revenues would be proportional to $w = ((\tau m + \rho_x E[x]) / (\tau + \rho_x))^2$, where $E[x] = q$. Therefore, given a set of consumer beliefs, the marginal revenue that a firm obtains by inflating the expected value of its ratings by a small amount would be

$$\begin{aligned} \frac{\partial w}{\partial E[x]} &= \frac{2\tau\rho_x m}{(\tau + \rho_x)^2} + \frac{2\rho_x^2}{(\tau + \rho_x)^2} E[x] \\ &= \frac{2\tau\rho_x m}{(\tau + \rho_x)^2} + \frac{2\rho_x^2}{(\tau + \rho_x)^2} q. \end{aligned}$$

Observe that the marginal revenue is a monotonically increasing function of the firm’s true quality: the higher the true quality (and therefore the higher the expected value of honest ratings posted on behalf of the firm), the higher the marginal gains that the firm derives from convincing consumers that its quality is even higher. Higher quality firms then have a higher incentive to manipulate.

To appreciate the essence of the above result, let us examine how different assumptions about firm revenues affect the form of equilibrium manipulation strategies. First, consider a setting identical to the above except that $f(\theta) = \theta^{2/3}$. Such a function characterizes settings where consumers have decreasing marginal returns from quality. In such settings, sales revenues are equal to $w = [f(\theta)]^{3/2} = \theta$. Because revenues are now a linear function of perceived quality, the marginal revenue that a firm can obtain by convincing consumers that its quality is higher by one unit is constant and independent of its current perceived quality. One, therefore, expects that a firm’s equilibrium manipulation strategy will also be independent of its true quality. This intuition is confirmed by the following result.

PROPOSITION 2. *Let θ be the mean of the consumers’ posterior beliefs about the firm’s quality. If sales revenues are equal to $w = \theta$ and manipulation cost is equal to $c(\eta) = \lambda\eta^2$, there exists linear PBE where the firm’s manipulation strategy is a constant*

$$\eta = \frac{\rho_x}{2(\tau + \rho_x)\lambda}.$$

A constant manipulation strategy can be thought of as a special case of a linear strategy $\eta = g + hq$, where $g = \rho_x / (2(\tau + \rho_x)\lambda)$ and $h = 0$. Observe that in this setting, firms of all qualities inflate their ratings by a constant amount. As discussed above, such

firm behavior does not affect the informativeness of the quality signal: consumers can recover the information contained in the original signal by subtracting the constant $g = \rho_x / (2(\tau + \rho_x)\lambda)$ from the published ratings they observe.

Finally, consider a setting where $f(\theta) = [\theta - \theta^2]^{2/3}$, and thus $w = [f(\theta)]^{3/2} = \theta - \theta^2$. In such a setting, both consumer demand and firm revenues are concave functions of perceived quality, rising, then declining as perceived quality increases.⁶ The marginal revenue $\partial w / \partial \theta = 1 - 2\theta$ that a firm can obtain by increasing the consumers’ perceptions of its quality is a declining function of their current perceptions: the higher the firm’s current perceived quality θ , the less the firm has to gain from inflating that perception by a given amount. The details of the solution are provided below.

PROPOSITION 3. *Let θ be the mean of the consumers’ posterior beliefs about the firm’s quality. If sales revenues are equal to $w = \theta - \theta^2$ and manipulation cost is equal to $c(\eta) = \lambda\eta^2$, there exists linear PBE where the firm’s manipulation strategy is a linear function of its true quality $\eta = g + hq$ and where*

$$g = \left(\frac{\tau}{\rho_x(h+1)^2} \left(m - \frac{1}{2} \right) - \frac{1}{2} \right) h$$

and $-1 < h < 0$ is a negative real solution of the fifth-degree polynomial equation

$$h = - \frac{\rho_x^2(h+1)^3}{(\tau + \rho_x(h+1)^2)^2 \lambda}.$$

Interestingly, in this setting, a firm’s manipulation strategy is a monotonically decreasing function of its true quality. The end result is that manipulation reduces the precision of the quality signal. The driver of this result is the fact that the firm’s marginal revenue from inflating consumer perceptions of its quality is a decreasing function of the firm’s current perceived quality.

The preceding examples show that if consumers are rational, strategic manipulation of online forums may increase, decrease, or leave the information value of forums unchanged. In all three cases, the outcome relates to the properties of the marginal revenue that results from increases in the firm’s perceived quality.

⁶ Such functional forms are easier to justify in settings where the unknown parameter q stands for something other than quality. For example, q might be the privately known number of members of a subscription-based online community (e.g., a multiplayer game). Consumers derive utility from larger communities because they can interact with more players, which makes the game more interesting. However, once communities increase past a certain size, certain disutilities (server congestion, lots of novice players) dominate the equation and reduce the community’s attractiveness.

In settings where the firm’s revenues are a convex function of perceived quality, marginal revenue is an increasing function of consumer perceptions of its quality: the better a firm is perceived to be, the more it has to gain from being perceived to be even better. Thanks to the presence of honest consumer ratings, in the absence of manipulation, the firm knows that its expected perceived quality will be equal to its true quality. If marginal revenues are increasing functions of perceived quality, a firm whose *true* quality is higher will then have higher incentives to inflate its ratings. This induces equilibrium manipulation strategies that are monotonically increasing functions of a firm’s true quality. Interestingly, in such cases, the more intensely a firm manipulates (i.e., the higher the h), the more its perfectly informative manipulation signal component $\eta = g + hq$ will crowd out the noisy signal component $x = q + \varepsilon$ obtained through honest consumer opinions, and thus the higher the precision of the resulting observable signal $y = x + \eta$. By *manipulating*, firms are, in fact, *signaling* their true quality to consumers.

The situation is reversed if a firm’s marginal revenue is a decreasing function of consumer perceptions of its quality. In such environments, the better a firm is perceived to be, the *less* it has to gain from being perceived to be even better. Equilibrium firm manipulation strategies are then monotonically decreasing functions of the firm’s true quality. Lower quality firms manipulate more than higher quality firms, shrinking the gap between their respective ratings and making it more difficult for consumers to infer a firm’s true quality from its online ratings: firm manipulation jams the online signal, reducing its precision.

In both cases, the presence of honest (but noisy) consumer ratings x that are centered on a firm’s true quality is crucial to the result. If these ratings did not exist, a firm’s incentive to manipulate would be independent of its true quality; the resulting equilibria would then convey no information.

2.2. Manipulation and Firm Profits

In this section, I explore the impact of forum manipulation on the firm’s profits. Specifically, I compare the firm’s equilibrium profits in environments where manipulation is possible to its profits in a setting where the exogenous quality signal x cannot be manipulated.

The easiest case is when the firm’s sales revenues are a linear function of its perceived quality ($w = \theta$). At a linear PBE, the firm will manipulate by a constant amount $\eta = g = \rho_x / (2(\tau + \rho_x)\lambda)$ independently of its true quality. Because consumers correctly anticipate the firm’s strategy, they can simply subtract g from the observable signal y to recover the original

signal x . Manipulation then has positive cost, but does not change consumer perceptions or the sales revenues of the firm; it becomes a net cost of doing business. The fact that manipulation is possible induces consumers to anticipate that the firm will engage in it. The firm is then trapped into performing the equilibrium level of manipulation that is expected of it, because, as in a rat race, a lower level will bias the consumers’ perceptions against it.⁷

Consider now the case where the firm’s sales revenues are convex with respect to its perceived quality (e.g., $w = \theta^2$). At a linear PBE, the firm’s manipulation strategy will be a monotonically increasing function of its true quality $\eta = g + hq$, $h > 0$. Anticipating the firm’s strategy, at equilibrium consumers subtract g from y and divide the result by $h + 1$ to get an estimate of the firm’s true quality. As before, the g component of manipulation is wasted, because it does not in any way affect the posterior beliefs of consumers. The hq component, on the other hand, serves to increase the precision of the signal $z = (y - g)/(h + 1)$ that consumers use in their inference process from ρ_x to $\rho_z = \rho_x(h + 1)^2$. The firm’s expected sales revenues then change from

$$E[w(x | q)] = \left(\frac{\tau m + \rho_x q}{\tau + \rho_x} \right)^2 + \frac{\rho_x}{(\tau + \rho_x)^2}$$

to

$$E[w(y | q, \eta)] = \left(\frac{\tau m + \rho_z q}{\tau + \rho_z} \right)^2 + \frac{\rho_z}{(\tau + \rho_z)^2}.$$

Observe that if the second term $\rho_z / (\tau + \rho_z)^2$ is small relative to the first term (which is the case when ρ_z is large), $E[w(y | q, \eta)]$ is a monotonically increasing function of the expected mean of posterior consumer beliefs $(\tau m + \rho_z q) / (\tau + \rho_z)$. The latter can be equivalently expressed as a weighted average $(1 - \kappa)m + \kappa q$ of the prior and signal means. Factor $\kappa = \rho_z / (\tau + \rho_z) = \rho_x(h + 1)^2 / (\tau + \rho_x(h + 1)^2)$ is proportional to the precision of signal z . Therefore, firms whose true quality q is higher than the mean of prior beliefs m benefit from the higher precision of signal z , because factor κ (and, thus the weighted average $(1 - \kappa)m + \kappa q$) monotonically increases with h . For the same reason, firms whose quality is lower than the mean of prior beliefs suffer a decline in sales revenues as h increases. This result is intuitive: because manipulation increases signal precision, it helps consumers better distinguish a firm’s true quality. This benefits the sales revenues of firms whose quality is higher than the prior mean and hurts firms whose quality is below the prior mean.

Since manipulation has a cost, what matters to firms is the net impact of this activity on profits. The

⁷ The situation is reminiscent of Holmstrom’s (1999) well-known “career concerns” model. The analogy is discussed in detail in §4.

crucial observation here is that the impact of manipulation on the mean of consumer posterior beliefs is a diminishing function of the precision ρ_x of honest consumer ratings. Intuitively, if the baseline signal of quality that the firm manipulates is already precise enough, further precision improvements have a very small impact on consumer beliefs. Mathematically, this is expressed by the fact that $\partial^2 \kappa / \partial \rho_x \partial h < 0$: the higher the precision of the original signal x , the smaller the impact that a given amount of manipulation has on the posterior mean (and thus, on firm revenues). At the limit $\rho_x \rightarrow \infty$, sales revenues converge to q^2 , and thus become unaffected by manipulation. On the other hand, from Equation (2), it is straightforward to show that as $\rho_x \rightarrow \infty$, the firm's manipulation strategy converges to a nonzero value $\eta = (\sqrt{\lambda^2 + 4\lambda} - \lambda)q/2\lambda$. Therefore, as the precision of honest ratings increases, the situation converges to a Holmstrom-like "rat race" (1999), where the firm spends resources on manipulation simply because it is expected to, without affecting consumer perceptions of its quality. This intuition is verified by the following result.

PROPOSITION 4. *Consider the setting of Proposition 1. The following statements are true:*

(a) *There exists a threshold R_1 such that if the precision of honest ratings satisfies $\rho_x < R_1$, a firm whose true quality is sufficiently high realizes net profits from manipulating online ratings.*

(b) *There exists a threshold R_2 such that if the precision of honest ratings satisfies $\rho_x > R_2$, firms of all qualities would be strictly better off if manipulation of online ratings was not possible.*

Finally, consider the case where the firm's sales revenues $w = \theta - \theta^2$ are concave with respect to its perceived quality. In this case, revenues are maximized when $\theta = 0.5$ and decline as θ moves away from 0.5 in either direction. At equilibrium, the firm's manipulation strategy will be a monotonically decreasing function of its true quality $\eta = g + hq$, $h < 0$. As before, component g is wasted because it does not change consumer beliefs. However, component hq now serves to decrease the precision of the signal $z = (y - g)/(h + 1)$ that consumers use in their inference process. The firm's expected revenues then change from

$$E[w(x | q)] = \frac{\tau m + \rho_x q}{\tau + \rho_x} - \left(\frac{\tau m + \rho_x q}{\tau + \rho_x} \right)^2 - \frac{\rho_x}{(\tau + \rho_x)^2}$$

to

$$E[w(y | q, \eta)] = \frac{\tau m + \rho_z q}{\tau + \rho_z} - \left(\frac{\tau m + \rho_z q}{\tau + \rho_z} \right)^2 - \frac{\rho_z}{(\tau + \rho_z)^2}.$$

Assuming that the last term is small relative to the first two, we see that lower signal precision increases

the sales revenues of firms whose true quality is sufficiently far away from the optimum $q = 0.5$ (and decreases the revenues of those that are close to 0.5). As in the convex revenues case, however, the marginal impact of manipulation is a declining function of the original signal's precision ρ_x . Thus, one expects that for sufficiently high ρ_x , the cost of manipulation will outweigh its benefits for firms of all qualities. It is possible to show that a version of Proposition 4 (where, for small ρ_x , firms whose qualities are sufficiently away from 0.5 realize net profits from manipulation, whereas for large ρ_x , all firms lose) holds in this case as well.

The most important conclusion of the preceding analysis is that if the exogenous quality signal x has low precision, then, for certain types of manipulation cost functions and demand functions, manipulation can result in net gains for some types of firms. On the other hand, once the precision of signal x surpasses a threshold, manipulation becomes a profit-reducing rat race for all types of firms: consumers expect that because firms can manipulate, they will. Firms are then trapped into performing such activities even though it results in a net loss for them: if they don't, consumer perceptions will be biased against them.

As the number of people who contribute to online forums goes up, one expects that the aggregate precision of consumer ratings posted in those forums (equal to the sum of the precisions of individual ratings) will also increase. If consumers come to expect that firms will manipulate, as forums keep growing in popularity, my model predicts that there will be a threshold beyond which firms will be trapped into having to engage in profit-reducing online manipulation practices. All firms would then be better off if consumers (rationally) expected them to manipulate less. The following proposition shows that one way of accomplishing this is by increasing the unit cost of manipulation λ .

PROPOSITION 5. *If the baseline precision of honest consumer ratings ρ_x is sufficiently high, as the unit cost of manipulation λ grows, all linear PBE have the following properties:*

- (a) *h declines*
- (b) *net firm payoffs grow.*

The unit cost of manipulation can be increased by implementing mechanisms that make it more difficult to anonymously post large numbers of strategically biased messages. The development of such *antimanipulation* technologies (somewhat analogous to anti-spam technologies) is the topic of active research. Most forums currently attempt to provide a rudimentary level of quality control by asking their users to rate the usefulness of posted feedback. Other proposed

approaches include statistical filtering of posted ratings (Dellarocas 2000), side-payment mechanisms that provide incentives to submit honest feedback (Miller et al. 2005), and cryptographic schemes that discourage the creation of multiple online identities through which manipulators can flood a forum with anonymous feedback (Friedman and Resnick 2001). Interestingly, the preceding analysis shows that in healthy forums where honest consumer participation is sufficiently high, it is firms and not consumers that have most to gain from the further development of such technologies.

3. The General Case

The preceding analysis derived the following three main results:

- Manipulation of online ratings increases (decreases) signal precision if firm strategies are monotonically increasing (decreasing) functions of the firm’s true quality.
- Equilibria where strategies are monotonically increasing (decreasing) functions of the firm’s true quality exist in settings where firm revenues are convex (concave) functions of the firm’s perceived quality.
- If the precision of the baseline signal that firms manipulate is sufficiently high, firms of all types would be strictly better off if manipulation of online ratings was not possible.

Our objective thus far has been to emphasize the underlying intuitions. All results were thus derived in the context of simple monopoly settings, assuming specific functional forms for firm payoffs, manipulation costs, and signal distributions that allowed us to obtain closed-form solutions. The current section shows that, appropriately generalized, versions of all three results hold true in a wide range of multifirm settings and for a broad class of consumer utilities, firm payoff functions and signal distributions.

3.1. The Setting

Consider a setting where M firms are facing N consumers. Each firm j , $j = 1, \dots, M$ is characterized by an attribute ω_j (the firm’s “type”), whose true value is unknown to consumers, but common knowledge among firms. A firm’s type might refer to its quality, its customer base, or any other attribute that matters to consumers and is difficult to verify before purchase. I assume that firm types are independently drawn from probability distributions $h_j(\omega_j)$. Denote by $\omega = (\omega_1, \dots, \omega_M)$ the vector of all firm types (the state of nature) and by ω_{-j} the vector of every other firm’s type (except firm j).

Each consumer i , $i = 1, \dots, N$ is faced with a decision problem (e.g., which product to purchase, on which firm’s stock to invest, etc.) whose value (consumer utility) $u_i(a_i, \omega)$ depends on her action a_i and

on the true state of nature ω . Firms’ payoff functions $w_j(a)$ depend on the consumers’ collective action $a = (a_1, \dots, a_N)$.

Word of mouth (e.g., the average value of honest online consumer ratings) is modeled as a noisy exogenous signal $x = (x_1, \dots, x_M)$ of the state of nature, characterized by mutually independent information structures $f_j(x_j | \omega_j)$. Firms can manipulate the word-of-mouth component that relates to their type in an attempt to influence consumer beliefs (and, therefore, actions) to their favor. Specifically, I assume that firms can shift the probability distribution of their signal component, so that it mimics the distribution that corresponds to another type. Signal manipulation is costly to firms. Denote by $\eta_j(\omega_j, \omega_{-j})$ firm j ’s manipulation strategy and by $c(\eta_j, \omega_j, \omega_{-j})$ the associated manipulation cost. The result of manipulation is a new signal $y = (y_1, \dots, y_M)$ characterized by information structures $g_j(y_j | \omega_j) = f_j(y_j | \omega_j + \eta_j)$.

It is easy to see that all settings we considered in §2 are special cases of this more general setting for $M = 1$ and $\omega = q$.

3.2. Manipulation and Informativeness

The first question that motivates this work is to distill the conditions under which signal y will be more (less) informative than x with respect to the decision problem faced by consumers. I use the standard notion of informativeness introduced by Blackwell (1951). Assume that a family of decision makers must make decisions whose payoff depends on an unknown state of nature ω , drawn from a distribution $h(\omega)$. Let x and y be two noisy signals of ω with corresponding information structures $f(x | \omega)$ and $g(y | \omega)$, respectively.

DEFINITION. Signal y is more informative than signal x for a class of decision problems if and only if every decision maker’s (ex ante) average surplus is higher when she bases her decisions on observations of y than when she bases her decisions on observations of x . Stated formally, y is more informative than x if and only if

$$\int_{\omega} h(\omega) \left(\int_y u_i(a_i(y), \omega) g(y | \omega) dy \right) d\omega > \int_{\omega} h(\omega) \left(\int_x u_i(a_i(x), \omega) f(x | \omega) dx \right) d\omega \quad \text{for all } i. \quad (3)$$

In the above equation, $a_i(y)$, $a_i(x)$ represent a decision maker’s best response upon observing signal realizations y and x , respectively.

It is well known since Blackwell (1951) that the comparison of information structures in arbitrary decision problems is often not possible. More meaningful results can be obtained if one restricts the set of decision problems and information structures

of interest. Lehmann (1988) studied decision problems in which the decision maker’s payoff function has single-crossing incremental returns, and posterior beliefs have the monotone likelihood ratio property (MLRP).⁸

Building on Lehmann’s (1988) work, I derive the following result.

THEOREM 1. *Assume that consumer payoff functions have single-crossing incremental returns in (a_i, ω_j) for all $i = 1, \dots, N, j = 1, \dots, M$. Assume, further, that all $f_j(\cdot | \cdot)$ belong to location families that have the MLRP. The following statements are true:*

(a) *Firm manipulation increases signal informativeness if manipulation strategies $\eta_j(\omega_j, \omega_{-j})$ are monotonically increasing functions of each firm’s own type (i.e., if $\partial\eta_j(\omega_j, \omega_{-j})/\partial\omega_j \geq 0$ for all j and all ω_j).*

(b) *Firm manipulation decreases signal informativeness if manipulation strategies $\eta_j(\omega_j, \omega_{-j})$ are monotonically decreasing functions of each firm’s own type (i.e., if $\partial\eta_j(\omega_j, \omega_{-j})/\partial\omega_j \leq 0$ for all j and all ω_j).*

Intuitively, if $\partial\eta_j(\omega_j, \omega_{-j})/\partial\omega_j \geq 0$ for all j , manipulation activity spreads out the location parameters $\omega_j + \eta_j(\omega_j, \omega_{-j})$ of the signal distributions that correspond to adjacent types ω_j . The distributions of y_j then become less crowded relative to the distributions of x_j . This makes the probabilistic mapping between an observed signal and the type that generated it more reliable. Conversely, if $\partial\eta_j(\omega_j, \omega_{-j})/\partial\omega_j \leq 0$ for all j , manipulation activity condenses the locations of the signals that correspond to adjacent types. The distributions of y_j then become more crowded relative to the distributions of x_j . This makes the mapping between signals and types less reliable.

An interesting observation is that even though a firm’s manipulation strategy may also depend on every other firm’s type, it is only the dependence of the strategy on the firm’s own type that matters to the above result.

Theorem 1 is consistent with the results of the previous section. In all cases we studied, equilibrium firm manipulation strategies followed a linear form $\eta = g + hq$. Theorem 1 then predicts that manipulation increases informativeness if $\partial\eta/\partial q = h > 0$ and decreases informativeness if $h < 0$. These are precisely the results we found in all three special cases we considered.

⁸ A utility function $u(a, \omega)$ has single-crossing incremental returns in (a, ω) if, for any action $a' > a$, the function $R(\omega) = u(a', \omega) - u(a, \omega)$ satisfies the single-crossing property (crosses zero only once and from below as ω grows). Athey and Levin (2001) show that if utility functions have single-crossing incremental returns and signal distributions satisfy the MLRP, then optimal consumer responses are monotone in the observed signal; that is, higher signal realizations result in higher consumer actions.

The following result connects the informativeness criterion of Theorem 1 to properties of firms’ payoff functions. This connection allows us to understand in what settings we can expect the existence of equilibria where online forum manipulation benefits or hurts consumers. The result itself is an immediate corollary of the work of Athey (2001) and McAdams (2003).

THEOREM 2. *A necessary and sufficient condition for the existence of equilibria where firm j ’s manipulation strategy $\eta_j(\omega_j, \omega_{-j})$ is a monotonically increasing (decreasing) function of that firm’s type ω_j is that the firm’s expected net payoff function (inclusive of the cost of manipulation) satisfies the single-crossing property in (ω_j, η_j) (or $(\omega_j, -\eta_j)$, respectively).*

A sufficient condition for Theorem 2 to hold is that the firm’s payoff function be supermodular (submodular) in (ω_j, η_j) . If a firm’s payoff function is twice-continuously differentiable, supermodularity (submodularity) is equivalent to positivity (negativity) of the cross-partial derivative of the payoff function with respect to η_j and ω_j .

It is straightforward to show that the results of §2 can be derived as special cases of Theorems 1 and 2. Consider, first, the setting of Proposition 1. In that setting, sales revenues are equal to $w = \theta^2$, manipulation costs equal to $c(\eta) = \lambda\eta^2$, and the mean of consumer beliefs after observing signal $y = q + \eta + \varepsilon$ equal to

$$\theta = \frac{\tau m + \rho_z((y - g)/(h + 1))}{\tau + \rho_z}.$$

The firm’s net payoff function is, thus, equal to

$$v = \left(\frac{\tau m + \rho_z(q + \eta + \varepsilon - g)/(h + 1)}{\tau + \rho_z} \right)^2 - \lambda\eta^2.$$

The cross-partial derivative of v with respect to q and η is equal to

$$\frac{\partial^2 v}{\partial q \partial \eta} = \frac{\rho_z^2}{(h + 1)^2(\tau + \rho_z)^2} > 0,$$

which implies that v is supermodular in q and η . The combination of Theorems 1 and 2 then predicts the existence of informativeness-enhancing equilibria. If we apply the same analysis to the setting of Proposition 3 ($w = \theta - \theta^2$), we find

$$\frac{\partial^2 v}{\partial q \partial \eta} = -\frac{\rho_z^2}{(h + 1)^2(\tau + \rho_z)^2} < 0,$$

which, according to Theorems 1 and 2, implies the existence of informativeness-reducing equilibria. More generally, assume that a firm’s payoff function is twice-continuously differentiable and can be written as

$$v_j = w_j(\theta_j(y_j), \theta_{-j}(y_{-j})) - c(\eta_j, \omega_j, \omega_{-j}),$$

where $\theta_j(y_j)$ represents the mean of consumers' posterior beliefs regarding ω_j after observing signal $y_j = \omega_j + \eta_j + \varepsilon_j$, $w_j(\cdot)$ denotes firm j 's sales profits given consumer beliefs, and $c(\cdot)$ is the cost of manipulation. Differentiating we obtain

$$\begin{aligned} \partial^2 v_j / \partial \eta_j \partial \omega_j &= w_j''(\theta_j(\omega_j + \eta_j + \varepsilon_j), \theta_{-j}(y_{-j}))[\theta_j'(\omega_j + \eta_j + \varepsilon_j)]^2 \\ &\quad + w_j'(\theta_j(\omega_j + \eta_j + \varepsilon_j), \theta_{-j}(y_{-j}))\theta_j''(\omega_j + \eta_j + \varepsilon_j) \\ &\quad - c_{12}(\eta_j, \omega_j, \omega_{-j}). \end{aligned} \quad (4)$$

In all three examples we considered in §2, it was $\theta''(y) = 0$ and $c_{12}(\eta, q) = 0$. In such settings, positivity of (4) is equivalent to convexity of $w(\cdot)$, which is exactly the result we found. In the more general case represented by Equation (4), assuming that $w_j'(\cdot) \geq 0$, for the cross-partial derivative to be positive, it must be $w_j''(\cdot) > 0$ and sufficiently positive to compensate for any negative terms that might be present in (4). Negative terms might be present if $c_{12}(\eta_j, \omega_j, \cdot) > 0$; that is, if the marginal cost of manipulation increases with a firm's true quality and/or if $\theta_j''(\cdot) < 0$. In such circumstances, the existence of equilibria where manipulation increases the information value of a forum is more likely when sales profits $w(\cdot)$ are sufficiently, steeply increasing convex functions of consumer perceptions regarding the unknown firm attribute ω_j . On the other hand, in settings where $c_{12}(\eta_j, \omega_j, \cdot) < 0$ (that is, where the marginal cost of manipulation decreases with a firm's true quality) and/or $\theta_j''(\cdot) > 0$, Equation (4) may be positive even when $w_j''(\cdot)$ takes negative values. We see, therefore, that the dependence of the marginal cost of manipulation on a firm's true quality increases or decreases the bar with respect to the amount of convexity that function $w(\cdot)$ must possess for informativeness-enhancing equilibria to exist. Observe that even though a firm's manipulation cost function may also depend on every other firm's quality, it is only the dependence of the cost function on the firm's own quality that matters to the result.

Convex profit functions arise in settings where the consumers' willingness to pay is a convex function of the unknown firm attribute (i.e., when consumers experience nondecreasing marginal utility from higher values of the unknown attribute) and/or when a firm's unit costs are concave functions of the unknown firm attribute. More generally, if consumer demand is given by $D_j(\omega_j) = f(\omega_j) - p$ and unit costs are given by $c(\omega_j)$, in settings with endogenous prices profit maximization implies sales profits $w_j = (f(\omega_j) - c(\omega_j))^2/4$. The second derivative of profits with respect to ω_j is equal to

$$\begin{aligned} w_j'' &= (f(\omega_j) - c(\omega_j))(f''(\omega_j) - c''(\omega_j))/2 \\ &\quad + (f'(\omega_j) - c'(\omega_j))^2/2. \end{aligned}$$

The above is positive if $f''(\omega_j) - c''(\omega_j)$ is positive (or, at least, not too negative). This, in turn, happens when $f''(\omega_j)$ is greater than (or, at least, not much smaller than) $c''(\omega_j)$, i.e., in settings where the acceleration of demand that is associated with higher values of parameter ω_j is not much smaller than the corresponding rate of increase of marginal costs.

One area where such profit functions arise are settings with network effects. Consider, for example, competition among incompatible mobile phone networks. In such environments, the higher the number of a network's users, the higher the marginal utility that consumers derive from using it, and therefore the higher the marginal increase in their willingness to pay for subscribing to it. If the marginal cost of operating the network stays constant with size, firm profits would most likely be convex functions of a network's size. If competing networks were able to manipulate an exogenous noisy signal of their network's true size (e.g., the number of postings associated with each network in a popular online chat room), the results of the paper then predict the existence of equilibria where such manipulation would increase the signal's informativeness.

Another area where convex profit functions often arise are settings with economies of scale. In that case, a firm's marginal costs decrease with the volume of sales. The resulting firm profits would, again, most likely be convex functions of sales volume. The above results then predict the existence of equilibria where attempts to manipulate online information that relates to the firm's volume of sales would end up increasing the signal's informativeness.

3.3. Manipulation and Firm Profits

The second important question motivating this paper is the impact of manipulation practices on firm profits. Presumably, firms manipulate word of mouth in an attempt to shift consumer beliefs to their favor, and thus increase their revenues. Nevertheless, in §2, we found that if consumers anticipate such firm behavior at equilibrium, they will appropriately adjust their beliefs (i.e., deflate what they see and hear) to compensate for the anticipated signal manipulation. Consumers then cannot be misled by inflated signals; the revenue gains or losses to firms from engaging in manipulation are purely because of the resulting changes in signal precision. As the precision of the baseline signal that firms manipulate increases, the marginal changes in precision that result from a given amount of manipulation decline. At the limit where the baseline signal has infinite precision, manipulation produces no change in signal precision, and thus no change in firm revenues. On the other hand, consumer expectations may force firms to engage in nonzero manipulation even in such cases. Firms are

then trapped into performing the equilibrium level of manipulation that is expected of them, because as in a rat race, a lower level will bias the consumers’ perceptions against them.

Proposition 4 proved this result for specific firm payoff and manipulation cost functions. In this section, I establish more general conditions under which the same result holds. Consider a setting where each firm’s net payoff function is twice-continuously differentiable and can be written as

$$v_j = w_j(\theta_j(y_j), \theta_{-j}(y_{-j})) - c(\eta_j, \omega_j, \omega_{-j}), \quad (5)$$

where $\theta_j(y_j)$ represents the mean of consumers’ posterior beliefs regarding ω_j after observing signal $y_j = x_j + \eta_j$, $w_j(\cdot)$ denotes firm j ’s sales profits given consumer beliefs, and $c(\cdot)$ is the cost of manipulation. Sales profits satisfy $w_j'(\cdot) > 0$, whereas the manipulation cost function satisfies $c(0, \omega_j, \omega_{-j}) = 0$ and $c_1(\eta_j, \omega_j, \omega_{-j}) > 0$ for all η_j .

Assume that the baseline word-of-mouth signals x_j are distributed according to $f_j(x_j | \omega_j, \rho_j)$, where $\rho_j \geq 0$ is a scalar parameter that relates to the signal’s informativeness. For the purposes of this discussion, I do not need to precisely define what ρ_j means. My only requirement is to assume that

- A1: Higher values of ρ_j correspond to higher informativeness of the baseline signal x_j (in the sense of Blackwell).
- A2: Equilibrium strategies and payoffs are continuous in ρ_j .
- A3: $\lim_{\rho_j \rightarrow \infty} f_j(x_j | \omega_j, \rho_j) = \delta(x_j - \omega_j)$, where $\delta(\cdot)$ is Dirac’s delta function.

The following Theorem establishes conditions that make manipulation profit and welfare reducing when the informativeness of the baseline word-of-mouth signal rises above a threshold.

THEOREM 3. *Consider a setting where each firm’s payoff function is described by (5), and where the probability distribution of each baseline word-of-mouth signal x_j is parameterized by an informativeness parameter ρ_j that satisfies (A1)–(A3). If there is no type interval $[\underline{\omega}, \bar{\omega}]$, $\underline{\omega} < \bar{\omega}$ such that*

$$w_j'(\omega_j) = c_1(0, \omega_j, \omega_{-j}) \quad \text{and} \quad w_j''(\omega_j) < c_{11}(0, \omega_j, \omega_{-j})$$

for all $\omega_j \in [\underline{\omega}, \bar{\omega}], \omega_{-j}, \quad (6)$

then there exists a threshold R_j such that if $\rho_j > R_j$, in all separating equilibria, all types of firm j (with the possible exception of types that belong to intervals of Lebesgue measure zero), as well as society, would be strictly better off if manipulation was not possible.

The intuition behind Theorem 3 is that in separating equilibria, as $\rho_j \rightarrow \infty$ firm profits (before manipulation costs) and consumer surplus become independent of

manipulation because the baseline signal allows consumers to perfectly infer the firm’s true type. On the other hand, if no type of firm j satisfies conditions (6), manipulation strategies of almost all types of firm j converge to a nonzero rat race limit. Manipulation then incurs a positive cost to firm j but provides no benefit to the firm or to consumers. Continuity of payoffs with respect to ρ_j establishes that the property holds for all sufficiently high ρ_j . The exceptions to Theorem 3 are settings where there exist firm-type intervals for which the profit and manipulation cost functions satisfy the special relationship (6). In such settings, and for those type intervals, as $\rho_j \rightarrow \infty$, there exists separating equilibria where manipulation strategies tend to zero. In such equilibria, at the limit $\rho_j \rightarrow \infty$, the relevant firm types are indifferent between being able to manipulate and not being able to.⁹

4. Related Work

This work relates to a number of streams in the economics and marketing literature.

In the quality signaling literature (Kihlstrom and Riordan 1984, Milgrom and Roberts 1986), producers of experience goods use costly signals (such as prices and advertising) to communicate their quality to consumers. Specifically, firms distort introductory prices (relative to myopically optimal levels) or burn money on advertising that conveys no direct quality information. Signaling occurs when, at equilibrium, higher quality firms advertise or distort their prices more than lower quality firms.

In the above papers, signaling behavior is driven by repeat-purchase arguments: higher quality firms are more confident that new adopters will become repeat customers. Therefore, their long-term marginal benefit from attracting one additional consumer is higher than that of their lower quality competitors. They can thus afford higher upfront advertising costs or lost revenue because of short-term price distortions.

In my model, quality signaling also arises when higher quality firms manipulate more than lower quality firms. However, the mechanism that gives rise to such behavior is entirely different and does not rely on repeat purchasing. Instead, as discussed in detail in §2.1, if firms have stage-game payoffs that are sufficiently convex functions of consumer perceptions of their quality, the presence of honest consumer ratings (modeled as a noisy exogenous signal that is centered

⁹ An example function pair that satisfies (6) for all $\omega_j \geq 0$ is $w(\omega_j) = \sqrt{\omega_j}$ and $c(\eta_j, \omega_j) = e^{\eta_j w_j(\omega_j)} = e^{\eta_j / 2\sqrt{\omega_j}}$. For settings and types that satisfy (6), the author was not able to derive general results regarding the relationship of payoffs with versus without manipulation near the limit $\rho_j \rightarrow \infty$. Theorem 3 thus provides *sufficient* conditions for profit- and welfare-reducing manipulation equilibria as $\rho_j \rightarrow \infty$.

on each firm's true quality) induces higher quality firms to manipulate more than lower quality firms even in a one-stage game.

Another related stream is the signal-jamming literature (Riordan 1985, Fudenberg and Tirole 1986, Mirman et al. 1994). In that literature, competing firms have uncertainty about some parameter (e.g., cost, demand curve) of their rivals. Moreover, they cannot observe this parameter directly but can make indirect noisy inferences from observing some other variable (e.g., prices). Rival firms then find it optimal to distort their prices or their outputs (relative to myopically optimal levels) to influence the direction or degree to which their opponents update their beliefs.

In the signal-jamming literature, firms manipulate to mislead one another by distorting prices and output. In contrast, in this paper, firms manipulate to mislead consumers by manipulating online word of mouth. Moreover, the central question of this paper (conditions under which manipulation increases or decreases the degree of learning) has not been addressed by any of the papers in the signal-jamming literature.

In Holmstrom's (1999) "career concerns" model, a manager of initially unknown ability is rewarded by the market on the basis of his perceived ability. The market infers the manager's ability from his output. Output is the sum of the manager's ability, plus labor, plus noise. At equilibrium, given the same history of past outputs, managers of all ability levels supply a constant amount of costly labor, even though the resulting output increases do not change the market's perception of their ability: the market knows that a manager can exert labor to inflate his output, and thus expects him to do so. The manager is then trapped in a rat race and forced to provide labor simply to fulfill the market's expectations, and thus to avoid being considered as less capable than he truly is.

If we replace *worker* with *firm*, *ability* with *quality*, and *labor* with *manipulation*, my model becomes a generalization of a two-period version of Holmstrom's (1999) model. Whereas Holmstrom's emphasis was on examining how career incentives affect the amount of labor that a manager provides over the span of his career, my objective is to determine how the form of a firm's payoff and manipulation cost functions determine whether manipulation increases or decreases the consumers' degree of learning regarding the firm's unknown parameter. Holmstrom (1999) only considers payoff functions that exhibit constant marginal returns to the market's perception of a manager's ability. Thus, in his work, the manager's actions cannot affect the market's degree of learning. In contrast, I show that if payoff functions exhibit increasing

(decreasing) marginal returns with respect to consumer perceptions of the firm's unknown parameter, at equilibrium, manipulation increases (decreases) the consumers' degree of learning.

Finally, in the marketing literature, Mayzlin (2006) offers a theoretical model of promotional chat in Usenet groups where consumers discuss products and services. Mayzlin's basic result is that if the ratio of profits to manipulation cost is high enough, there exists an equilibrium in which both firms manipulate but the low-quality firm manipulates more. Promotional chat thus decreases the informativeness of online forums. My work generalizes Mayzlin's (2006) result, shows that there exist settings where manipulation can increase forum informativeness, and considers the impact of manipulation cost and degree of consumer participation on firm profits.

5. Concluding Remarks

Recent advances in information technologies have made it easier than ever for, otherwise unrelated, individuals to pool their opinions and experiences on any imaginable topic on Internet opinion forums. Such forums appear to be exerting increasing influence on consumer behavior. Because of their scale and relative anonymity, Internet forums are vulnerable to manipulation from interested parties. This paper offers a systematic exploration of the impact of such manipulation on both consumers and the firms that might be doing it. The principal results can be summarized as follows:

- If every firm's manipulation strategy is a monotonically increasing (decreasing) function of that firm's true quality, strategic manipulation of online forums increases (decreases) the information value of a forum to consumers. This result implies that there exist settings where forum manipulation ends up benefiting consumers.

- Equilibria where strategies are monotonically increasing (decreasing) functions of a firm's true quality exist in settings where that firm's net payoff function, inclusive of the cost of manipulation, is supermodular (submodular) in the firm's quality and manipulation action. In many settings, the net payoff function is supermodular if and only if the firm's profit function (before manipulation costs are taken into account) is sufficiently convex with respect to the firm's perceived quality.

- In a broad class of settings, if the precision of the baseline signal that firms manipulate is sufficiently high, firms of all types as well as society, would be strictly better off if manipulation of online forums was not possible.

The results of this paper have interesting implications for practice. First, they show that in most

cases, all firms will manipulate their online ratings. This implies that consumers should come to expect a certain amount of hype to be present in most online forums and must learn to compensate for it (by properly deflating what they see and read in such forums) when making inferences from such information. Second, contrary to intuition, my analysis shows that in a broad class of settings, there exist equilibria where manipulation of online forums by firms becomes a form of quality signaling that improves the information value of honest ratings provided by consumers. Such equilibria are, in fact, more likely to exist in settings where consumers have increasing marginal returns from higher quality (and thus have a higher appreciation for quality goods). Third, as online forums keep growing in popularity, one expects that the volume and quality of content that will be contributed by honest consumers will also increase. My model predicts that if consumers come to expect the presence of firm manipulation, once the quality of honest user content surpasses a threshold, manipulation of online forums will become a profit-reducing necessity for all firms.

Since the quality of honest content is high, firms will not be able to substantially change consumer beliefs through manipulation. However, because (anticipating that firms will manipulate) consumers will still deflate what they see, firms will be trapped into performing the equilibrium amount of manipulation that is expected of them. If they don't, consumer perceptions will be biased against them.

Online forum operators acquire a pivotal role in this new competitive environment. By investing in technologies that make it more difficult to manipulate (see §2.2), operators can help reduce the manipulation expenditures of all competing firms in their industry. Interestingly, as the volume and quality of user-contributed online content increases, it is firms, and not consumers, that have most to gain from the further development of such technologies.

The online appendix to this paper provides examples indicating that the results of this paper are qualitatively robust to several modeling extensions. Specifically,

- Appendix A.2 studies a setting where, in addition to praising themselves, firms can bad-mouth their competitors.
- Appendix A.3 studies a setting where there are multiple forums and each consumer only visits a subset of them.
- Appendix A.4 studies a setting where firms can also imperfectly signal their quality through prices.

There are several promising directions for future research. This paper analyzes settings where forum discussions are focused on topics for which the underlying state of the world has an exogenous and objective true value. There are several important forum

settings (ranging from discussions of fashion goods to political debates) where there is no such objective true state of the world. The impact of online manipulation in such forums is an interesting open question. Also, all results assume that consumers are rational and capable of correctly anticipating the firms' manipulation strategies. The degree to which this happens in actual practice, and the resulting implications for firms and consumers, is an empirical question of considerable practical interest. Finally, this paper looks at single-period settings. In real life, Internet forums are continuously updated with new information from consumers. It is well known that multiperiod social learning is vulnerable to several anomalies (such as herding and information cascades) that may slow down or completely inhibit learning. The study of forum manipulation in multiperiod settings is, thus, an intriguing next step of this line of research.

An online appendix to this paper is available on the *Management Science* website (<http://mansci.pubs.informs.org/eocompanion.html>).

Acknowledgments

This material is based on work supported by the National Science Foundation under CAREER grant 9984147. The author is grateful to Vidyanand Choudhary, Il-Horn Hann, Terry Hendershott, Flavio Toxvaerd, Dimitri Vayanos, seminar participants at the University of California, Berkeley, Carnegie Mellon University, the University of California, Irvine, University of Southern California, The Wharton School, and the anonymous referees for helpful comments.

Appendix. Proofs

PROOF OF PROPOSITION 1. Expected firm revenues are given by

$$\begin{aligned}
 E[w(y | q, \eta)] &= E\left[\left(\frac{\tau m + \rho_z(y - g)/(h + 1)}{\tau + \rho_z}\right)^2 \middle| q, \eta\right] \\
 &= \left(E\left[\frac{\tau m + \rho_z(y - g)/(h + 1)}{\tau + \rho_z}\right]\right)^2 \\
 &\quad + V\left[\frac{\tau m + \rho_z(y - g)/(h + 1)}{\tau + \rho_z}\right] \\
 &= \left(\frac{\tau m + \rho_z(q + \eta - g)/(h + 1)}{\tau + \rho_z}\right)^2 + \frac{\rho_z}{(\tau + \rho_z)^2}.
 \end{aligned}$$

The firm's objective is to select η to maximize

$$\begin{aligned}
 v &= -c(\eta) + E[w(\theta | q, \eta)] \\
 &= -\lambda\eta^2 + \left(\frac{\tau m + \rho_z(q + \eta - g)/(h + 1)}{\tau + \rho_z}\right)^2 + \frac{\rho_z}{(\tau + \rho_z)^2}.
 \end{aligned}$$

The first-order condition $\partial v / \partial \eta = 0$ results in a linear equation on η . Solving for η verifies that the firm's optimal manipulation strategy is a linear function $\eta = g + hq$, where

$$g = \frac{(\tau m(h + 1) - \rho_z g)\rho_z}{\lambda(\tau + \rho_z)^2(h + 1)^2 - \rho_z^2} \quad h = \frac{\rho_z^2}{\lambda(\tau + \rho_z)^2(h + 1)^2 - \rho_z^2}.$$

Consistency of beliefs requires that the coefficients g and h that are optimal for the firm are equal to those conjectured

by consumers. Solving the above two equations for g, h , we obtain

$$g = \frac{\tau mh}{\rho_z} \quad h = \frac{\rho_z^2}{(\tau + \rho_z)^2(h+1)\lambda}. \quad (7)$$

The value of h is determined from the equation $h(h+1) = \rho_z^2/((\tau + \rho_z)^2\lambda) > 0$, $\rho_z = \rho_x(h+1)^2$. Elementary algebra shows that this equation always has at least one real positive root. The second-order condition $\partial^2 v/\partial \eta^2 < 0$ requires that

$$-2\lambda + \frac{2\rho_z^2}{(\tau + \rho_z)^2(h+1)^2} < 0.$$

Substituting

$$\lambda = \frac{\rho_z^2}{(\tau + \rho_z)^2(h+1)h}$$

from (7), the requirement becomes

$$-\frac{2\rho_z^2}{(\tau + \rho_z)^2(h+1)^2h} < 0,$$

which implies $h > 0$. Thus, only positive roots are admissible.

PROOF OF PROPOSITION 2. Expected firm revenues are now given by

$$\begin{aligned} E[w(y | q, \eta)] &= E\left[\frac{\tau m + \rho_z(y - g)/(h+1)}{\tau + \rho_z} \mid q, \eta\right] \\ &= \frac{\tau m + \rho_z(q + \eta - g)/(h+1)}{\tau + \rho_z}. \end{aligned}$$

If we follow the procedure of Proposition 1, solution of the linear equation that arises from first-order conditions gives $g = \rho_x/2(\tau + \rho_x)\lambda$ and $h = 0$. The second-order condition $\partial^2 v/\partial \eta^2 = -2\lambda < 0$ is always satisfied.

PROOF OF PROPOSITION 3. Expected firm revenues are now given by

$$\begin{aligned} E[w(y | q, \eta)] &= E\left[\frac{\tau m + \rho_z(y - g)/(h+1)}{\tau + \rho_z} \right. \\ &\quad \left. - \left(\frac{\tau m + \rho_z(y - g)/(h+1)}{\tau + \rho_z}\right)^2 \mid q, \eta\right] \\ &= E\left[\frac{\tau m + \rho_z(y - g)/(h+1)}{\tau + \rho_z}\right] \\ &\quad - \left(E\left[\frac{\tau m + \rho_z(y - g)/(h+1)}{\tau + \rho_z}\right]\right)^2 \\ &\quad - V\left[\frac{\tau m + \rho_z(y - g)/(h+1)}{\tau + \rho_z}\right] \\ &= \frac{\tau m + \rho_z(q + \eta - g)/(h+1)}{\tau + \rho_z} \\ &\quad - \left(\frac{\tau m + \rho_z(q + \eta - g)/(h+1)}{\tau + \rho_z}\right)^2 - \frac{\rho_z}{(\tau + \rho_z)^2}. \end{aligned}$$

If we follow the procedure of Proposition 1, the first-order condition $\partial v/\partial \eta = 0$ results in a linear equation on η . Solving for η verifies that the firm's optimal manipulation strategy is a linear function $\eta = g + hq$. Consistency of beliefs requires that the coefficients g and h that are optimal for the firm are equal to those conjectured by consumers. Solving for g, h , we obtain

$$g = \left(\frac{\tau m}{\rho_z} - \frac{\tau}{2\rho_z} - \frac{1}{2}\right)h \quad h = -\frac{\rho_z^2}{(\tau + \rho_z)^2(h+1)\lambda}. \quad (8)$$

The value of h is determined from the equation $h(h+1) = -\rho_z^2/((\tau + \rho_z)^2\lambda) < 0$, $\rho_z = \rho_x(h+1)^2$. Elementary algebra shows that this equation always has at least one negative real root $-1 < h < 0$. The second-order condition $\partial^2 v/\partial \eta^2 < 0$ requires that

$$-2\lambda - \frac{2\rho_z^2}{(\tau + \rho_z)^2(h+1)^2} < 0.$$

Substituting

$$\lambda = -\frac{\rho_z^2}{(\tau + \rho_z)^2(h+1)h}$$

from (8), the requirement becomes

$$\frac{2\rho_z^2}{(\tau + \rho_z)^2(h+1)^2h} < 0,$$

which implies $h < 0$. Thus, all negative roots are admissible.

PROOF OF PROPOSITION 4. Substituting the equilibrium values of g and h into $\eta = g + hq$, the firm's value function (1) becomes

$$\begin{aligned} v(h) &= -\lambda \left(\frac{\tau m}{\rho_x(h+1)^2} + q\right)^2 h^2 + \left(\frac{\tau m + \rho_x(h+1)^2 q}{\tau + \rho_x(h+1)^2}\right)^2 \\ &\quad + \frac{\rho_x(h+1)^2}{(\tau + \rho_x(h+1)^2)^2}. \end{aligned} \quad (9)$$

The firm's value function in an environment where, for exogenous reasons, manipulation is not possible is simply $v(0)$. Our goal is to determine the sign of $\Delta v(h) = v(h) - v(0)$ for large values of ρ_x and q .

The expression $\Delta v(h)$ can be written as the ratio of two, rather lengthy, polynomial expressions (omitted). The denominator of $\Delta v(h)$ is always positive. Expressed as a function of ρ_x , the numerator is a sixth-degree polynomial. The highest power coefficient is simply $-h^2\lambda q^2(h+1)^8 < 0$ for all parameter values h, λ, q . This shows that for sufficiently high ρ_x , $\Delta v(h) < 0$: the firm is strictly worse off in environments where manipulation is possible, irrespective of its true quality and the unit cost of manipulation.

When we express the numerator of $\Delta v(h)$ as a function of q , it becomes a second-degree polynomial. Substituting

$$\lambda = \frac{\rho_z^2}{(\tau + \rho_z)^2(h+1)h},$$

the highest power coefficient of q becomes

$$\rho_x^4 h(h+1)^4 (-(h+1)^3 \rho_x^2 + 2\tau(h+1)^2 \rho_x + \tau^2(h^2 + 3h + 3)).$$

The above coefficient is positive when ρ_x lies between the roots of the second-degree polynomial equation $-(h+1)^3 \cdot \rho_x^2 + 2\tau(h+1)^2 \rho_x + \tau^2(h^2 + 3h + 3) = 0$. For such values of ρ_x , if q is sufficiently high, then $\Delta v(h) > 0$: the firm is strictly better off in environments where manipulation is possible. From elementary algebra, given the signs of its coefficients, the above equation has one negative and one positive root. Since, by definition, $\rho_x \geq 0$, this implies that if ρ_x is smaller than the positive root, then $\Delta v(h) > 0$. Solving the equation, this gives the requirement

$$\rho_x < \frac{h+1 + \sqrt{h^3 + 5h^2 + 8h + 4}}{(h+1)^2} \tau.$$

PROOF OF PROPOSITION 5.

Part (a). From (7), differentiating we obtain

$$\frac{\partial h}{\partial \lambda} = -\frac{\rho_x^2(h+1)^4(\tau + \rho_x(h+1)^2)}{F(\rho_x)},$$

where $F(\rho_x)$ is a third-degree polynomial in ρ_x , whose highest order term is $\lambda^2(2h+1)(h+1)^6 > 0$. Therefore, for sufficiently high values of ρ_x , the denominator of the above derivative will always be positive, and the derivative itself, negative.

Part (b). At the limit $\rho_x \rightarrow \infty$, the firm's equilibrium value function (9) simplifies to

$$v(h) = q^2(1 - \lambda h^2).$$

At the same limit, Equation (7) gives

$$h(h+1) = 1/\lambda \Rightarrow h = (\sqrt{\lambda^2 + 4\lambda} - \lambda)/2\lambda.$$

Substituting h into $v(h)$ and differentiating with respect to λ , we obtain

$$\frac{\partial v(h)}{\partial \lambda} = \frac{q^2(\sqrt{\lambda(\lambda+4)} - \lambda)^2}{4\lambda\sqrt{\lambda(\lambda+4)}} > 0.$$

Since both $v(h)$ and h are continuous functions of ρ_x , positivity of the derivative $\partial v(h)/\partial \lambda$ at the limit $\rho_x \rightarrow \infty$ implies that there exists a threshold R such that the derivative is also positive for all $\rho_x > R$.

PROOF OF THEOREM 1. Lehmann's (1988) result can be stated as follows. Let x and y be scalar random variables drawn from information structures $F(x | \omega)$ and $G(y | \omega)$, where ω is an unknown scalar "state of nature" and F and G are probability distributions whose densities $f(x | \omega)$, $g(y | \omega)$ satisfy the monotone likelihood ratio property in x and y , respectively. Then, observing y is more effective than observing x with respect to any decision problem where the decision makers' payoff function has single-crossing incremental returns in ω if and only if

$$G^{-1}[F(x | \omega) | \omega] \text{ is a nondecreasing function of } \omega \text{ for all } x. \quad (10)$$

A utility function $u(\omega, a)$ has single-crossing incremental returns in ω if, for any action $a' > a$, the function $R(\omega) = u(\omega, a') - u(\omega, a)$ satisfies the single-crossing property (crosses zero only once and from below as ω grows). Condition (10) can be equivalently written as requiring that the function

$$y(\omega, x) \text{ such that } G(y(\omega, x) | \omega) = F(x | \omega) \text{ is a nondecreasing function of } \omega \text{ for all } x. \quad (11)$$

Let $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ be vectors of independent random variables drawn from information structures $F_j(x_j | \omega_j)$ and $G_j(y_j | \omega_j)$, where the components of the state of nature ω_j are also independent. A corollary of Lehmann's (1988) result is that y is more informative than x if the decision maker's payoff function has single crossing incremental returns for each ω_j and, in addition, the following condition holds for all j :

$$y_j(\omega_j, x_j) \text{ such that } G_j(y_j(\omega_j, x_j) | \omega_j) = F_j(x_j | \omega_j) \text{ is a nondecreasing function of } \omega_j \text{ for all } x_j. \quad (12)$$

In our setting, $G_j(y_j | \omega_j) = F_j(y_j | \omega_j + \eta_j(\omega_j, \omega_{-j}))$. If we make the additional assumption that distributions F and G belong to location families, then $F_j(x_j | \omega_j) = F_j(x_j - \omega_j)$, $G_j(y_j | \omega_j) = F_j(y_j - \omega_j - \eta_j(\omega_j, \omega_{-j}))$. Substituting into (12)

$$\begin{aligned} G_j(y_j(\omega_j, x_j) | \omega_j) &= F_j(x_j | \omega_j) \\ \Leftrightarrow F_j(y_j(\omega_j, x_j) - \omega_j - \eta_j(\omega_j, \omega_{-j})) &= F_j(x_j - \omega_j) \\ \Leftrightarrow y_j(\omega_j, x_j) &= x_j + \eta_j(\omega_j, \omega_{-j}). \end{aligned}$$

Differentiating with respect to ω_j , it is $\partial y_j(\omega_j, x_j)/\partial \omega_j = \partial \eta_j(\omega_j, \omega_{-j})/\partial \omega_j$. Therefore $y_j(\omega_j, x_j)$ is nondecreasing if and only if $\partial \eta_j(\omega_j, \omega_{-j})/\partial \omega_j \geq 0$. Thus, if $\partial \eta_j(\omega_j, \omega_{-j})/\partial \omega_j \geq 0$ for all ω_j and all j , observing y is preferable to observing x . Following a similar procedure, we find that if $\partial \eta_j(\omega_j, \omega_{-j})/\partial \omega_j \leq 0$ for all ω_j and all j , observing x is preferable to observing y .

PROOF OF THEOREM 3. I will show that under the conditions stated in the Theorem, at the limit $\rho_j \rightarrow \infty$, in all separating equilibria firms of type j and society would be strictly better off if manipulation were not possible. Continuity of payoffs with respect to ρ_j then implies that under these conditions, there exists a threshold R_j such that the property holds for all $\rho_j > R_j$.

If word of mouth has infinite precision, the mapping between firm types and word-of-mouth signals becomes deterministic. Let ω_j be the firm's type, $x_j = \omega_j$ its baseline signal, and $\eta_j(\omega_j)$ its manipulation strategy. Consumers then observe $y_j = \omega_j + \eta_j(\omega_j)$. Anticipating that firms manipulate, and that (as I will prove below) manipulation leads to a separating equilibrium, consumers discount what they see by $\zeta_j(y_j)$, and thus infer $\theta_j(y_j) = y_j - \zeta_j(y_j)$. At equilibrium, consumer inferences are correct, which implies that $\zeta_j(\omega_j + \eta_j(\omega_j)) = \eta_j(\omega_j)$. In such a setting, the firm's payoff function (5) becomes

$$v_j = w_j(\omega_j + \eta_j - \zeta_j(\omega_j + \eta_j), \theta_{-j}(y_{-j})) - c(\eta_j, \omega_j, \omega_{-j}). \quad (13)$$

In the rest of the proof, to reduce the size of the formulae, I will omit the dependence of $w_j(\cdot)$ on $\theta_{-j}(y_{-j})$ and that of $c(\cdot)$ on ω_{-j} since they do not play a role on the results. The firm selects η_j to maximize v_j . The first-order condition with respect to η_j gives

$$w'_j(\omega_j + \eta_j - \zeta_j(\omega_j + \eta_j))(1 - \zeta'_j(\omega_j + \eta_j)) - c_1(\eta_j, \omega_j) = 0. \quad (14)$$

The second-order condition requires

$$\begin{aligned} w''_j(\omega_j + \eta_j - \zeta_j(\omega_j + \eta_j))(1 - \zeta'_j(\omega_j + \eta_j))^2 \\ - w'_j(\omega_j + \eta_j - \zeta_j(\omega_j + \eta_j))\zeta''_j(\omega_j + \eta_j) - c_{11}(\eta_j, \omega_j) < 0. \end{aligned} \quad (15)$$

Consistency of consumer beliefs requires that at equilibrium, $\zeta(\omega_j + \eta_j(\omega_j)) = \eta_j(\omega_j)$. Differentiating, we obtain

$$\zeta'(\omega_j + \eta_j(\omega_j)) = \frac{\eta'_j(\omega_j)}{1 + \eta'_j(\omega_j)} \quad \zeta''(\omega_j + \eta_j(\omega_j)) = \frac{\eta''_j(\omega_j)}{(1 + \eta'_j(\omega_j))^3}.$$

Substituting into (13)–(15), we obtain

$$v_j = w_j(\omega_j) - c(\eta_j(\omega_j), \omega_j), \quad (16)$$

$$\frac{w'_j(\omega_j)}{1 + \eta'_j(\omega_j)} - c_1(\eta_j(\omega_j), \omega_j) = 0, \quad (17)$$

$$\frac{w''_j(\omega_j)}{(1 + \eta'_j(\omega_j))^2} - \frac{w'_j(\omega_j)\eta''_j(\omega_j)}{(1 + \eta'_j(\omega_j))^3} - c_{11}(\eta_j(\omega_j), \omega_j) < 0. \quad (18)$$

Equation (16) gives the firm's net equilibrium payoff. If manipulation were not possible, firm- j 's net payoff would simply be $w_j(\omega_j)$. We see, therefore, that a firm will be strictly worse off manipulating if at equilibrium, $\eta_j(\omega_j) \neq 0$. Under the assumption of separating equilibria, consumer surplus is identical in settings with and without manipulation because they are able to perfectly infer the firm's true quality in either case. Therefore, in addition to firm profits, social welfare will also be strictly lower if at equilibrium, $\eta_j(\omega_j) \neq 0$.

The firm's equilibrium manipulation strategy $\eta_j(\omega_j)$ is any solution of the first-order differential Equation (17) that satisfies (18). From Equation (17), we see that if $\eta_j(\omega_j) \neq 0$, it will be $\eta_j(\omega_j) \neq 0$ for all firm types, except for isolated types ω_j that correspond to the points where function $\eta_j(\omega_j)$ might cross zero. From (17) and (18), we see that equilibrium strategies, where $\eta_j(\omega_j) = 0$ (and thus $\eta_j'(\omega_j) = \eta_j''(\omega_j) = 0$) for all $\omega_j \in [\underline{\omega}, \bar{\omega}]$, $\underline{\omega} < \bar{\omega}$, can only arise if the profit and manipulation cost functions satisfy

$$w_j'(\omega_j) = c_1(0, \omega_j) \quad w_j''(\omega_j) < c_{11}(0, \omega_j) \quad \text{for all } \omega_j \in [\underline{\omega}, \bar{\omega}].$$

An example function pair that satisfies the above constraints for all $\omega_j \geq 0$ is $w(\omega_j) = \sqrt{\omega_j}$ and $c(\eta_j, \omega_j) = e^{\eta_j w_j'(\omega_j)} = e^{\eta_j/2\sqrt{\omega_j}}$.

The proof assumes the existence of separating equilibria where consumers can perfectly infer a firm's type from the signal they observe. We complete the proof by showing that our assumption is justified, given our assumptions on the firm's profit and manipulation cost functions. Assumption of a separating equilibrium implies that there is a one-to-one mapping between firm types ω_j and observable signals $\omega_j + \eta_j(\omega_j)$. If $w_j'(\omega_j) > 0$ and $c_1(\eta_j, \omega_j) > 0$, Equation (17) gives $1 + \eta_j'(\omega_j) = \partial(\omega_j + \eta_j(\omega_j))/\partial\omega_j > 0$. Equilibrium observable signals are then strictly monotonic in firm types, which is consistent with our assumption of a separating equilibrium.

References

- Athey, S. 2001. Single crossing properties and the existence of pure-strategy equilibria in games of incomplete information. *Econometrica* 69(4) 861–889.
- Athey, S., J. Levin. 2001. The value of information in monotone decision problems. Working Paper No. 98-24, Department of Economics, MIT, Cambridge, MA.
- Blackwell, D. 1951. Comparison of experiments. J. Neyman, ed. *Proc. Second Berkeley Sympos. Math. Statist. Probab.* University of California Press, Berkeley, CA.
- Chevalier, J., D. Mayzlin. 2006. The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3) 345–354.
- DeGroot, M. H. 1970. *Optimal Statistical Decisions*. McGraw-Hill, New York.
- Dellarocas, C. 2000. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. *Proc. 2nd ACM Conf. Electronic Commerce*. Association for Computing Machinery, New York, 150–157.
- Dellarocas, C., M. Fan, C. Wood. 2003. Self-interest, reciprocity, and participation in online reputation systems. 2003 *Workshop Inform. Systems and Econom.* (WISE), Seattle, WA.
- Dichter, E. 1966. How word-of-mouth advertising works. *Harvard Bus. Rev.* 44(Nov–Dec) 147–166.
- Engel, J. F., R. D. Blackwell, P. W. Miniard. 1993. *Consumer Behavior*, 8th ed. The Dryden Press, Orlando, FL.
- Friedman, E., P. Resnick. 2001. The social cost of cheap pseudonyms. *J. Econom. Management Strategy* 10(1) 173–199.
- Fudenberg, D., J. Tirole. 1986. A signal-jamming theory of predation. *RAND J. Econom.* 17 366–376.
- Gu, B., S. Jarvenpaa. 2003. Online discussion boards for technical support: The effect of token recognition on customer contributions. *Proc. 2003 Internat. Conf. Inform. Systems (ICIS)*. Association for Information Systems, Atlanta, GA.
- Harmon, A. 2004. Amazon glitch unmasks war of reviewers. *New York Times* (February 14).
- Hennig-Thurau, T., K. P. Gwinner, G. Walsh, D. D. Gremler. 2004. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *J. Interactive Marketing* 18(1) 38–52.
- Holmstrom, B. 1999. Managerial incentive problems: A dynamic perspective. *Rev. Econom. Stud.* 66(1) 169–182.
- Kihlstrom, R., M. Riordan. 1984. Advertising as a signal. *J. Political Econom.* 92(4) 427–450.
- Lehmann, E. L. 1988. Comparing location experiments. *Ann. Statist.* 16(2) 521–533.
- Mayzlin, D. 2006. Promotional chat on the Internet. *Marketing Sci.* 25(2) 157–165.
- McAdams, D. 2003. Isotone equilibrium in games of incomplete information. *Econometrica* 71(4) 1191–1214.
- Milgrom, P., J. Roberts. 1986. Price and advertising signals of product quality. *J. Political Econom.* 94(5) 796–821.
- Miller, N., P. Resnick, R. Zeckhauser. 2005. Eliciting informative feedback: The peer-prediction method. *Management Sci.* 51(9) 1359–1373.
- Mirman, L. J., L. Samuelson, E. E. Schlee. 1994. Strategic information manipulation in duopolies. *J. Econom. Theory* 62 363–384.
- Riordan, M. 1985. Imperfect information and dynamic conjectural variations. *RAND J. Econom.* 16 41–50.
- Schindler, R., B. Bickart. 2003. Published “word of mouth”: Referable, consumer-generated information on the Internet. C. Hauvgedt, K. Machleit, R. Yalch, eds. *Online Consumer Psychology: Understanding and Influencing Behavior in the Virtual World*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Senecal, S., J. Nantel. 2004. The influence of online product recommendations on consumers' online choices. *J. Retailing* 80 159–169.
- Sundaram, D. S., K. Mitra, C. Webster. 1998. Word-of-mouth communications: A motivational analysis. *Adv. Consumer Res.* 25 527–531.
- Thompson, N. 2003. More companies pay heed to their “word of mouse” reputation. *New York Times* (June 23).
- Walker, R. 2004. The hidden (in plain sight) persuaders. *New York Times Magazine* (December 5).
- White, E. 1999. Chatting a singer up the pop charts. *The Wall Street Journal* (October 5).