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

Chrysanthos Dellarocas*

August 2004

Abstract

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 try to manipulate consumer perceptions by posting costly anonymous messages that praise their products or by offering incentives to consumers to do so. This paper offers a theoretical analysis of the impact of such behavior on firm profits and consumer surplus. We examine a setting where two firms simultaneously introduce imperfect substitute experience goods of different qualities and consumers obtain quality information from an online forum. The most striking result of our analysis is that strategic manipulation can either *decrease* or *increase* the information value of online forums to consumers relative to the case where no manipulation takes place. Specifically, there exist settings where the presence of honest consumer opinions induces firms to reveal their own, more precise, knowledge of product qualities by manipulating the forums at relative intensities that are proportional to their actual qualities. However, if a sufficiently large number of consumers post honest opinions online forum manipulation is harmful to firms because its cost outweighs its benefits. The social overhead of online manipulation can be reduced by developing technologies that increase the unit cost of manipulation and by encouraging higher participation of honest consumers.

*Address: Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139. This material is based upon work supported by the National Science Foundation under CAREER Grant No. 9984147. I am grateful to Vidyanand Choudhary, Il-Horn Hann, Terry Hendershott, Flavio Toxvaerd, Dimitri Vayanos and seminar participants at Berkeley, CMU, Irvine, USC, and Wharton for helpful comments.

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, in turn, 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.

In the past few years, online forums have been steadily growing in popularity and have become an indispensable component of portals (e.g. Yahoo! Movies), traditional news sites (e.g. BBC Talking Point), consumer report sites (e.g. CNET, Epinions, Citysearch, Moviefone), and online retailer websites (e.g. Amazon, Circuit City). Table 1 lists some common types of online forums in use today.

There is growing evidence that consumers are influenced by online forums before making a variety of purchase decisions (Chevalier and Mayzlin, 2003; Senecal and Nantel, 2003; Thompson 2003). Similar evidence suggests that forums play an increasingly important role in public opinion formation. Online 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 scalability and decentralized nature of these new information structures can lead to better informed and, ultimately, more efficient markets and societies. The above argument certainly has merits. From the law of large numbers, the aggregation of sufficiently many observations of "ordinary" consumers is eventually going to exceed the precision of any single specialist's opinion, however enlightened that may be. Furthermore, the decentralized nature of information exchange makes it difficult for any single stakeholder to directly control the signal that reaches the population of receivers.

Nevertheless, those same properties (decentralized architecture, relative anonymity of information sources) that make the Internet such a difficult medium to control directly, make it relatively easy for stakeholders to *indirectly* manipulate the probability distribution of information propagated through online forums by anonymously adding their own strategically biased messages to the total mix of posted opinions.

Online review sites. Includes consumer opinions published on the Internet by online merchants, by commercial web sites that specialize in posting consumer opinions, and by consumers who publish their product opinions on their own web sites, including "revenge" sites.

Mailbags. Includes customer and reader comments and feedback posted on web sites of such organizations as consumer products manufacturers, service providers, magazines, and news organizations.

Discussion forums. Includes bulletin boards, Usenet groups, and published ongoing discussions on specific topics.

Listservs. Includes consumer opinions sent by e-mail to the members of an e-mail list.

Chat rooms. Includes real-time conversations over the Internet between groups of people, often based on a particular topic.

Table 1: Common types of online forums (Adapted from Schindler and Bickart (2003)).

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 about them 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 takes place. For example, when, in February 2004, due to 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 those reviews were written by the books' own publishers, authors, and competitors (Harmon 2004). The music industry is known to hire professional "posters" who surf various online chat rooms and fan sites in order to post positive opinions on behalf of new albums (White 1999; Mayzlin, 2003).

Given the potential backlash of such activity[1], firms are experimenting with more indirect approaches. 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 that would start "conversations" in online forums praising singer Lucinda Williams's albums (see 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

[1]For example, as a result of the above-mentioned incident, Amazon has stopped accepting anonymous reviews.

favorable reviews by sending them free samples, inviting them to special events, etc. There is at least one professional marketing firm who conducts such campaigns on behalf of its clients (see www.electricartists.com).

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 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 that are affected by them. The results of such analyses will be relevant to policy decisions (should Internet forum manipulation be outlawed?), 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 towards Internet forums (how much should consumers trust online forums? how much should firms invest in trying to manipulate them? is consumer participation in such forums socially beneficial?).

This paper contributes to answering such questions by analyzing how the strategic manipulation of Internet opinion forums affects the payoffs of consumers and firms in a duopoly of vertically differentiated imperfect substitute experience goods. I assume that the qualities of the two competing goods are known to firms but, initially, unknown to consumers. I further assume that the main source of quality information for consumers is an online "product review" forum (such as Epinions.com) where past consumers post opinions about their experiences with the goods. New consumers read these opinions and form perceptions about the qualities of the competing products. Based on those 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. Furthermore, consumers are "smart"; even though they cannot directly distinguish "honest" opinions from "fake" opinions, they are aware that manipulation takes place and adjust their interpretation of online opinions accordingly.

The most striking result of our analysis is that strategic manipulation can either *decrease* or *increase* the informativeness of online forums (in the sense of decreasing or increasing the expected ex-ante payoff of consumers who base their decisions on information published in such forums) relative to the case where no manipulation takes place. Specifically, there exist settings where the presence of honest consumer opinions induces firms to reveal their own, more precise, knowledge of product qualities by manipulating at relative intensities that are proportional to their actual qualities (that is, the high quality firm inflates its, already higher, ratings by a higher amount than the low quality firm). This behavior accentuates the gap between the ratings of the two competitors and increases the ability of consumers to distinguish the high quality firm. Outcomes of this type arise in settings where the marginal payoff to a firm from shifting consumers' perceptions to its favor is an *increasing* function of its true quality. In contrast, in settings where the marginal payoff is a *decreasing* function of true quality, equilibrium firm manipulation intensities are inversely proportional to firms' true qualities (i.e. the low quality firm inflates its ratings more than the high quality firm, closing the gap that separates the aggregate ratings of the two firms). Manipulation activity then decreases the ability of consumers to distinguish the high quality firm.

In the duopoly setting that forms the center of our analysis, I show that, in markets where prices are endogenously determined (e.g. markets for consumer electronics, cars, etc.), if the marginal manipulation costs do not rise too steeply with actual quality, manipulation activity tends to *increase* forum informativeness. On the other hands, in markets where prices are exogenously fixed (e.g. markets for CDs, movies, etc.), manipulation activity usually *decreases* forum informativeness.

The preceding analysis shows that there are practically important cases where the impact of online manipulation on consumers will be positive. However, even in such cases, if a sufficiently large number of consumers post honest opinions online, I show that the cost of manipulation to firms outweighs its benefits: the impact of the incremental signal precision on firm revenues is lower than the corresponding manipulation costs. Nevertheless, firms have no choice: the high quality firm is forced to spend resources inflating its online reviews to prevent its competitor from catching up with it. The low quality firm must do the same to avoid being considered as even lower quality than it truly is. Online manipulation, thus, becomes an "arms race" that harms the profits of both firms.

To summarize, even though it can sometimes be beneficial to *consumers*, if the size of the market is large enough, online forum manipulation is always harmful to *firms*. The result is surprising as it is the opposite of what common sense might have suggested. The paper shows that the social cost of online manipulation can be reduced by developing technologies that increase the unit cost of manipulation and by encouraging higher participation of consumers who post honest opinions.

This work relates to a number of different streams of economics literature. In the signal-jamming literature (Holmstrom 1999; Riordan 1985; Fudenberg and Tirole 1986; Mirman, Samuelson and Schlee 1994) players that possess some private information (about their costs, ability, quality, etc.) attempt to influence the direction or degree to which their rivals update their beliefs by distorting prices or outputs relative to myopically optimal levels, sacrificing current profits but affecting the informational content of the market outcome in ways that increase future expected profits. In the quality signaling literature (Kihlstrom and Riordan 1984; Milgrom and Roberts 1986) producers of experience goods use costly signals (such as prices or advertising) to communicate their quality to consumers. Judd and Riordan (1994) in particular show that if consumers have some (possibly small amount of) information about a product's quality that firms don't possess, price signaling is possible even in the absence of cost asymmetries or repeat purchases. In a similar manner, our analysis shows that the presence of honest but noisy consumer ratings can sometimes induce firms to act in ways that signal their private information about their relative qualities to consumers. In such settings, online manipulation becomes an alternative form of quality signaling, similar in spirit to dissipative advertising.

In the marketing literature, Mayzlin (2003) offers a theoretical model of promotional chat in Usenet groups where consumers discuss products and services. Each consumer is exposed to a random subset of messages (in Mayzlin's model, a single message) posted in those groups. Firms masquerade as consumers and inject promotional messages in order to influence consumer opinions and thus increase their sales. 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, the low quality firm manipulates more, but overall, online chat remains somewhat informative: consumers would still be correct to follow the advice of the single random message they are exposed to, even though the probability of making a wrong choice is higher than if there was no promotional chat. Promotional chat thus *decreases* the informativeness of online forums. Our work generalizes Mayzlin's result, shows that there exist settings where manipulation can *increase* forum informativeness, identifies general conditions under which manipulation increases or decreases informativeness, and considers the impact of manipulation cost and degree of consumer participation on forum informativeness and firm profits.

The rest of the paper is organized as follows. Section 2 introduces the model. Sections 3 and 4 derive equilibrium strategies and comparative statics in settings where prices are endogenously and exogenously determined respectively. Section 5 shows how the results generalize for a broad class of payoff functions and signal distributions. Finally, Section 6 discusses the strategic implications of our findings for consumers, firms, and forum operators and concludes.

2 The setting

Two competing firms, A and B (denoted by the subscript j, $j = A, B$) simultaneously introduce imperfect substitute durable products (automobiles, digital cameras, laptops, etc.) The appeal of each product to consumers is the sum of two independent components: a "type" component, representing product characteristics whose valuation depends on each individual consumer's taste (color, shape, look-and-feel, etc.), and a "quality" component, representing characteristics whose valuation is identical among all consumers (ease of use, durability, etc.) A product's "type" can be reliably communicated to consumers, whereas a product's true "quality" can only become known after the good is bought and consumed.

The two products are assumed to be located at the [0, 1] taste interval. Product A is located at 0 and product B is located at 1. Qualities q_A, q_B are exogenously given and stay constant throughout the game. Firms know each other's quality. Consumers initially ignore this information. Variable costs are assumed to be zero or, alternatively, marginal costs are constant and identical for both firms, and prices are defined net of marginal costs[2].

Demand functions. Each period, N consumers enter the market and stay there for just one period. Consumers are uniformly distributed on the unit "taste interval" and face linear transportation costs. Consumers have unit demand. A consumer's utility from consuming products A, B is given by:

$$u_A^i = q_A - \frac{i}{2} - p_A \quad u_B^i = q_B - \frac{1-i}{2} - p_B \tag{1}$$

where $i \in [0, 1]$ is the consumer's location on the taste interval.

Let $q = q_A - q_B$ denote the *quality differential* between goods A and B. Let I_{it} denote the information available to consumer i in period t and $\theta_{it} = E(q|I_{it})$. Furthermore, let $\theta_t = E(\theta_{it})$.

[2]These assumptions characterize settings where quality, defined as the degree of matching between product characteristics and consumer preferences, is uncorrelated with cost.

Expected utility maximization and price-taking behavior given prices $p_{A,t}$, $p_{B,t}$ imply the following linear demand functions[3]:

$$x_{A,t} = N\left(\frac{1}{2} + \theta_t - p_{A,t} + p_{B,t}\right) \quad x_{B,t} = N\left(\frac{1}{2} - \theta_t - p_{B,t} + p_{A,t}\right) \tag{2}$$

Consumer information. I assume that the only source of consumer information regarding product qualities is an online forum (such as Epinions.com) where past consumers post anonymous numerical ratings about products[4]. At the beginning of each period the forum publishes the arithmetic mean of all consumer ratings posted for that firm during all past periods[5].

Throughout this paper I will consider a two-period version of the game. To highlight the effect of online ratings on consumer learning I also will assume that prices do not play a signaling role[6]. Denote the two periods 0 and 1. Each period's consumers observe only current prices and aggregate ratings. Consumers have identical normally distributed priors regarding the two firms' qualities with precision (inverse of variance) 2τ. Therefore, their prior beliefs regarding the quality differential q have mean zero and precision τ.

During period 0 there are no online ratings. Therefore $\theta = 0$. Since there is no cost asymmetry between firms A and B the only possible equilibrium is a pooling equilibrium. Given demand functions (2), profit maximization requires that both firms set prices equal to $p_A = p_B = 1/2$ and enjoy identical demand $x_A = x_B = N/2$.

At the end of period 0, an exogenously determined fraction r of consumers who bought each product during that period visits the online forum and posts quality ratings[7]. I assume that con-

[3]In the rest of the paper I will omit the time subscript whenever it is implied.

[4]In Appendix C I relax this assumption and show that the presence of additional sources of consumer information does not qualitatively change the results.

[5]Alternatively, the forum solicits anonymous text reviews; each review implicitly makes a quality judgment about a product and, thus, corresponds to a scalar quality rating. At the beginning of each period, the forum publishes the text of all reviews submitted in past periods. Each period's consumers then read all reviews, weigh them equally, and make (identical) inferences about a product's quality from the information contained therein.

[6]This assumption is without loss of generality if we assume that firms have perfect information about each other's quality and set prices directly. In the absence of cost asymmetries between firms, no separating equilibrium where prices alone perfectly signal quality is sustainable. If one assumes that prices only carry *imperfect* information about quality (for example, because firms themselves possess imperfect information about their qualities or because a retailer introduces noise to the final price) then it *is* possible to construct separating equilibria where both prices and online ratings affect consumer learning (see Judd and Riordan (1994) and Caminal and Vives (1996) for examples of such models). In such models the strategic considerations affecting online rating manipulation are qualitatively identical to those arising in models where prices do not signal quality. Since the focus of this paper is not the signaling of quality through price, to clarify the exposition we therefore assume that prices do not play a signaling role.

[7]Throughout the paper I am treating online word-of-mouth as an exogenous phenomenon and do not make an

sumer ratings are honest, although consumers make occasional mistakes in experiencing a product's true quality or in posting their opinion online[8]. Online ratings are thus noisy signals of a product's true quality. I assume that individual ratings are independent and follow a normal distribution with precision normalized to 2.

At the beginning of period 1, the forum aggregates individual ratings and publishes the arithmetic means y_A, y_B of all ratings posted at the end of period 0. In the absence of explicit manipulation, published ratings y_j correspond to normally distributed random variables with mean q_j and identical precision $\rho = N/2 \times r \times 2 = Nr$ ($N/2$ period-0 consumers; a fraction r posts ratings; the precision of each rating is equal to 2).

Period-1 consumers view the published ratings y_A, y_B and update their prior beliefs about the quality difference of the two firms. Firms view the same information and set prices. Finally, consumers observe prices and make purchasing decisions.

Forum Manipulation. At the end of period 0 firms attempt to manipulate period-1 consumer beliefs by posting anonymous ratings that praise their respective product (or by providing monetary incentives to consumers to do so)[9]. The forum cannot distinguish between genuine ratings and fake ratings, and thus includes all ratings in the published averages. We assume that the effect of manipulation is to inflate the mean of the probability distribution of a firm's published ratings from q_j to $q_j + \eta_j$.

Forum manipulation is a costly activity. Firms engaging in direct manipulation must put sufficient effort so that consumers do not identify their fake reviews as such. Firms engaging in indirect manipulation must put effort in identifying key consumers (and must often offer them monetary incentives). An important determinant of the results that follow is the cost of manipulation activity. To my knowledge there is no prior empirical study of this phenomenon and thus no reliable

attempt to rationalize the factors that determine the fraction of consumers that engage in it. Although this problem is important, it is orthogonal to the main questions addressed by this work. It has also been studied elsewhere (see Avery, Resnick, and Zeckhauser (1999) for a theoretical treatment and Dellarocas, Fan and Wood (2003) and Gu and Jarvenpaa (2003) for empirical studies of the drivers of consumer participation in online forums).

[8]The honesty of consumer reports can be weakly justified by the assumption that consumers only stay in the market for one period. They are, thus, indifferent between true and false reporting. Moreover, a number of side payment mechanisms that provide consumers with strict incentives to both participate in a reputation mechanism as well as to rate truthfully have been proposed in the literature (Miller, Resnick and Zeckhauser 2002). Such mechanisms can be easily combined with my model to provide a rational justification for consumer participation and honest reporting.

[9]I do not include the ability to bad-mouth one's competitors in the base model. Appendix C shows that adding the possibility of bad-mouthing to the model does not qualitatively change the essence of the results.

information about the nature of its costs. I, therefore, *assume* a cost function that satisfies the following three properties:

- Property 1: The cost of manipulation is proportional to the number of honest ratings Nr
- Property 2: The cost of manipulation is an increasing convex function of the amount η_j by which a firm wishes to inflate its ratings
- Property 3: The marginal cost of manipulation is an increasing function of a firm's true quality q_j

Appendix B provides intuition and partial justification for Properties 1-3 by presenting a simple setting where these properties arise as natural consequences of the common-sense assumption that the cost of manipulation is proportional to the total number of fake reviews that firms must post to inflate their average ratings by a given amount.

There are several cost functions that satisfy Properties 1-3. To retain mathematical tractability, I assume a quadratic cost function $c(q_j, \eta_j) = \lambda Nr(\eta_j^2 + \mu q_j \eta_j)$, where λ is a parameter that relates to the cost of posting a single fake message and μ captures the degree to which the marginal cost of manipulation depends on a product's true quality[10].

Forum informativeness. A fundamental question addressed by this paper is whether forum manipulation activity will increase or decrease the value of an online forum to consumers. A formal definition of the notion of *informativeness* is essential to making such comparisons. I conclude this section by providing such a definition.

Assume that a family of decision makers (the consumers) must make decisions whose payoff depends on an unknown state of the world q, drawn from a distribution $h(q)$. Let x and y be two noisy signals of q with corresponding information structures (conditional probability distributions) $f(x|q)$ and $g(y|q)$ respectively.

Definition 1 *Signal y is more informative than signal x for a class of decision problems if and only if the (ex-ante) average consumer surplus when consumers base their decisions on observations of y*

[10] In Section 5 I show how the results of this analysis generalize for arbitrary cost functions.

is higher than when consumers base their decisions on observations of x. Stated formally, y is more informative than x if and only if:

$$\int h(q)\left(\int E[u_i(a_i(y), y, q)]g(y|q)dy\right)dq > \int h(q)\left(\int E[u_i(a'_i(x), x, q)]f(x|q)dx\right)dq \qquad (3)$$

In the above equation $a_i(y), a'_i(x)$ represent the ith consumer's best response upon observing signal realizations y and x respectively, while $u_i(a, s, q)$ represents the utility function of the ith consumer if she takes action a when she observes signal s and the true state of the world is q.

3 Endogenous prices

This section analyzes the impact of strategic forum manipulation in a duopoly setting where prices are endogenous. The most striking result is that, if the marginal cost of manipulation does not grow too steeply with a firm's true quality, there exist separating equilibria in linear strategies where forum manipulation activity *increases* the informativeness of the forum ratings, and thus its value to consumers.

Consider the setting of Section 2. Let θ be the average consumer belief about the quality difference of firms A and B at the beginning of period 1. If prices do not signal quality then, given demand functions (1), maximization of expected sales revenues implies:

$$\begin{array}{lll} \text{prices} & p_A = \frac{1}{2} + \frac{\theta}{3} & p_B = \frac{1}{2} - \frac{\theta}{3} \\ \text{demand} & x_A = N\left(\frac{1}{2} + \frac{\theta}{3}\right) & x_B = N\left(\frac{1}{2} - \frac{\theta}{3}\right) \\ \text{sales revenues} & w_A = N\left(\frac{1}{2} + \frac{\theta}{3}\right)^2 & w_B = N\left(\frac{1}{2} - \frac{\theta}{3}\right)^2 \end{array} \qquad (4)$$

The notion of equilibrium I use is perfect Bayesian (PBE). That is, players maximize their expected payoffs at any point in time given the beliefs they have, and beliefs are consistent in the Bayesian sense with strategies. Beliefs are obtained from equilibrium strategies and observations using Bayes' rule. The analysis of this game is considerably simplified by the fact that, due to Gaussian noise, all feasible signals are observable with positive probability. There are, thus, no out-of-equilibrium paths and no need to consider equilibrium refinements.

The analysis assumes that consumers are aware that firms may be attempting to manipulate online ratings (although they cannot directly observe the exact amounts of manipulation). Specifically, consumers understand that published ratings are the sum of three (indistinguishable) components:

$$y_j = q_j + \eta_j(q_A, q_B) + \varepsilon_j$$

where q_j are true qualities, $\eta_j(q_A, q_B)$ is the amount by which each firm inflates its ratings and ε_j are normally distributed error terms with mean zero and precision $\rho = Nr$.

I will now show the existence of a PBE in linear strategies. Suppose that consumers in period 1 believe that the amounts by which firms inflate their online quality ratings at the end of period 0 are linear functions of the quality differential q:

$$\eta_A = g + hq \quad \eta_B = g - hq$$

where g, h are real numbers that, at equilibrium, correspond to correct conjectures. From the above expressions for y_A, y_B it follows that:

$$q = \frac{y_A - y_B}{2h+1} - \frac{\varepsilon_A - \varepsilon_B}{2h+1} \tag{5}$$

in other words, the publicly observable statistic $z = (y_A - y_B) / (2h+1)$ is a normally distributed unbiased estimator of q with precision $\rho_z = \rho\,(2h+1)^2 / 2$. If consumers update their beliefs using Bayes rule, given the normality of prior beliefs and all observable signals, standard theory (DeGroot, 1970) predicts that each consumer's posterior beliefs about q will be normally distributed with mean:

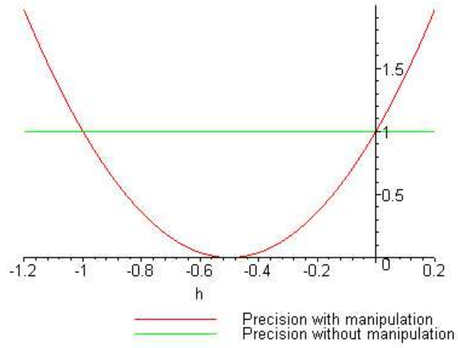
$$\theta = \frac{\rho_z z}{\tau + \rho_z} \tag{6}$$

The following theorem shows that the forum's informativeness is proportional to the precision of statistic z.

Theorem 1 *The ex-ante average consumer surplus obtained by basing inferences θ about the quality differential q on observations of z is an increasing function of that signal's precision ρ_z.*

Figure 1 shows how the precision $\rho_z = \rho\,(2h+1)^2 / 2$ of z compares to the precision $\rho/2$ of

Figure 1: Impact of manipulation on online forum informativeness.



$y_A - y_B$ in the absence of manipulation. There are five regions of interest.

1. If $h = 0$ then both firms inflate their ratings by the same amount g; the informativeness of online feedback regarding the quality differential is unaffected by manipulation.

2. If $h > 0$ then the high quality firm inflates its ratings more than the low quality firm. Given that, on the average, the high quality firm's honest ratings are higher than those of the low quality firm, manipulation activity accentuates that lead and helps consumers more clearly differentiate between the two competitors. This is a case in which feedback manipulation *increases* the informativeness of feedback forums.

3. If $-0.5 < h < 0$ then the low quality firm inflates its ratings more than the high quality firm and therefore decreases the "gap" between the two firms' expected aggregate ratings. The high quality firm still ends up with higher expected ratings; however, because the expected ratings are now closer together, the relative effect of noise is higher. This weakens the ability of consumers to infer the quality difference of the two competitors. In such cases feedback manipulation *decreases* the informativeness of online forums.

4. If $h = -0.5$ then the low quality firm exactly closes the gap between itself and its high quality competitor; the two firms become indistinguishable in terms of their average ratings ($E(y_A) = E(y_B)$). Such behavior completely *destroys* the informativeness of online forums; if consumers expect such behavior they will ignore online ratings.

5. If $h < -0.5$ then the low quality firm inflates its ratings substantially more than the high

quality firm so that the average ratings of the low quality firm exceed those of the high quality firm. If consumers expect such behavior, this becomes a form of reverse signaling; consumers understand that the firm with lower ratings is in fact the high quality firm and use equation (5) to estimate the true quality differential.

The following theorem shows that the equilibrium value of h depends on how steeply the marginal cost of manipulation grows with a firm's true quality.

Theorem 2 *Consider the duopoly setting of Section 2. Assume that prices are endogenous and do not signal quality. The following statements are true:*

1. *There exist Perfect Bayesian Equilibria in linear strategies (linear PBE) in which firm manipulation strategies are given by*

$$\eta_A = \left(\frac{N(2h+1)}{6\lambda(2\tau + Nr(2h+1)^2)} - \mu \frac{q_A + q_B}{4} \right) + hq$$

$$\eta_B = \left(\frac{N(2h+1)}{6\lambda(2\tau + Nr(2h+1)^2)} - \mu \frac{q_A + q_B}{4} \right) - hq$$

where h is a real solution of the following system of constraints:

$$\lambda = \frac{4N^2 r(2h+1)^3}{9(2\tau + Nr(2h+1)^2)^2(4h+\mu)} \quad \textit{and} \quad \frac{4h+\mu}{4(2h+1)} < 1$$

2. *The sign of admissible values of h depends on the magnitude of μ. Specifically,*

   (a) *if $\mu < \min(4+4h,\ 4N^2 r/9\lambda(2\tau+Nr)^2)$ then there exist linear PBE where $h > 0$*

   (b) *if $\mu = 4N^2 r/9\lambda(2\tau+Nr)^2 < 4$ then there exists a linear PBE where $h = 0$*

   (c) *otherwise, all linear PBE have $h < 0$*

3. *If $\lambda > 0$, there can never be a linear PBE where $h = -1/2$*

Theorem 2 shows that costly manipulation can never completely destroy the informativeness of a forum, (i.e. there can never be an equilibrium where $h = -1/2$). Even more importantly, however, if μ is not too large, there exist linear PBE where the high quality firm manipulates more intensely than the low quality firm. In such equilibria the, already higher, honest ratings of the high quality firm are inflated by a higher amount than the ratings of the low quality firm. The resulting ratings accentuate the gap between the high and the low quality firm and make it easier for consumers to differentiate between the two competitors. In such cases manipulation activity *increases* the informativeness of the online forum and becomes a form of *quality signaling.*

The intuition behind this result becomes clear once we study the payoff functions of the two competitors. From (4) recall that each firm's equilibrium price and demand are linear functions of consumer perceptions ($\pm\theta$) about that firm's quality difference relative to its competitor. Sales revenues $w_j = N\left(1/2 \pm \theta/3\right)^2$ are, therefore, quadratic functions of the perceived quality difference. The marginal revenues $\partial w_j/\partial(\pm\theta) = 2N\left(1/2 \pm \theta/3\right)$ that result from shifting consumer perceptions to each firm's favor (through manipulation) are, thus, higher for the firm for which, in the absence of manipulation, next-period's consumers would have higher opinions. The presence of honest consumer ratings guarantees that, on the average, this would be the high quality firm. The high quality firm, thus, expects to receive higher marginal revenues from inflating its ratings than the low quality firm. At the same time, the marginal cost of manipulation is higher for the high quality firm by a factor proportional to μ. If μ is not too large, the net marginal benefit of manipulation is higher for the high quality firm. The high quality firm will then manipulate more intensely than the low quality firm.

The presence of honest (but noisy) consumer ratings is crucial to the result. It is exactly the knowledge that this "base signal" exists that induces the two firms to reveal their own precise quality information by inflating their ratings in relative amounts that are exactly proportional to their quality difference. The more intensely firms manipulate, the more their perfectly informative signals will crowd out the noisy signal obtained through honest consumer opinions and thus increase forum informativeness.

The rest of the section will offer additional insights about the properties of manipulation equilibria by exploring their comparative statics. The following two theorems show how the intensity of manipulation and the resulting forum informativeness relate to the unit cost of manipulation and the degree of active consumer participation in the forum.

Theorem 3 *As the unit cost of manipulation λ grows:*

1. *Forum manipulation intensities η_j decline.*

2. *If, additionally, N is sufficiently large*[11] *then h and, thus, forum informativeness also decline.*

As expected, by making it costlier to manipulate (for example, by introducing technologies that make it more difficult to create multiple identities or to send large amounts of feedback from the same computer), forum operators can reduce the degree to which such activity takes place. The interesting side-effect is that, in settings where manipulation increases informativeness, less manipulation *decreases* the value of the forum to consumers.

An interesting relationship also exists between manipulation and the fraction of consumers who contribute honest online ratings. An increase in the fraction r of consumers who contribute ratings increases the cost of forum manipulation because the number of "fake" reviews that are required in order to inflate a firm's average ratings by a given amount is proportional to the number of honest reviews (see Appendix B). One expects that, at equilibrium, higher amounts of honest feedback contribution by consumers will result in lower levels of forum manipulation by firms. This intuition is confirmed by the following theorem.

Theorem 4 *If N is sufficiently large then, as the fraction of consumers r who submit feedback grows:*

1. *Forum manipulation intensities η_j decline.*

2. *h declines.*

As established by Theorem 1, forum informativeness is proportional to $\rho_z = Nr\,(2h+1)^2/2$. As r grows, the base precision of the forum signal $(Nr/2)$ grows but, by Theorem 4, the incremental precision $(2h+1)^2$ due to manipulation declines. The following theorem shows that, if the size of the market is large, the combined effect is positive: on the balance increased consumer participation increases the value of the forum.

Theorem 5 *If N is sufficiently large then, as the fraction of consumers r who submit feedback grows, forum informativeness grows.*

[11]The condition "N is sufficiently large" in Theorems 3-8 captures the requirement that the number of honest ratings must be large enough so that the precision $\rho = Nr$ of average honest ratings is comparable to or higher than the precision τ of prior beliefs. The precise thresholds differ for each theorem. The interested reader can find the details in the proofs.

Finally, if enough honest consumers submit ratings, firm profits are monotonically increasing with the unit cost of manipulation and consumer participation.

Theorem 6 *If* $\mu < (6 + 4q)/(3 + 4q + 2(q_A + q_B))$ *and* N *is sufficiently large, firm profits are monotonically increasing functions of the unit cost of manipulation* λ *and the rate of consumer participation* r*.*

In conjunction with Theorems 3 and 4, Theorem 6 has a remarkable implication:

Corollary 1 *If the size of the market is large, both firms are better off manipulating less.*

The intuition behind this conclusion is based on the fact that, if a sufficiently large number of consumers post ratings, honest ratings already possess sufficient precision to signal quality. The incremental signal precision due to manipulation then has a small effect on average posterior beliefs[12]. If N is high enough, the revenue gains obtained through manipulation are not enough to justify the manipulation cost for either firm. Nevertheless, firms have no choice: the high quality firm is forced to spend resources inflating its online reviews to prevent its competitor from catching up with it. The low quality firm must do the same to avoid being considered as even lower quality than it truly is. Online manipulation, thus, becomes an "arms race" that harms the profits of both firms. Both firms will be better off if high unit costs of manipulation or high levels of consumer participation induce them to keep such activities at low levels.

4 Exogenous prices

This section analyzes the impact of strategic forum manipulation in a duopoly setting where prices are exogenously fixed. In such a setting firms compete for market share. Let θ be the average consumer belief about the quality difference of firms A and B at the beginning of period 1 and let $p_A = p_B = p$. The analysis of Section 2 then implies:

$$\begin{array}{lll} \text{demand} & x_A = N\left(\frac{1}{2}+\theta\right) & x_B = N\left(\frac{1}{2}-\theta\right) \\ \text{sales revenues} & w_A = pN\left(\frac{1}{2}+\theta\right) & w_B = pN\left(\frac{1}{2}-\theta\right) \end{array} \quad (7)$$

[12]From (6) one can show that, for $\rho > \tau$, $\partial^2\theta/\partial\rho\partial h < 0$: the higher the base precision ρ of honest ratings, the lower the impact of a given h on the average posterior beliefs.

The marginal revenues that result from shifting consumer perceptions to each firm's favor (through manipulation) are now independent of the firm's true quality. If the marginal cost of manipulation grows with firm quality, high quality firms have a lower marginal benefit from manipulation and are, thus, expected to manipulate less intensely than low quality firms. This intuition is confirmed by the following theorem:

Theorem 7 *Consider the duopoly setting of Section 2. Assume that prices p are exogenously set. If $\mu < 2$ then there exist Perfect Bayesian Equilibria in linear strategies where firm manipulation strategies are given by:*

$$\eta_A = \left(p\frac{N(2-\mu)}{\lambda(8\tau + Nr(2-\mu)^2)} - \mu\frac{q_A + q_B}{4}\right) - \frac{\mu}{4}q, \quad \eta_B = \left(p\frac{N(2-\mu)}{\lambda(8\tau + Nr(2-\mu)^2)} - \mu\frac{q_A + q_B}{4}\right) + \frac{\mu}{4}q$$

We see that, if $\mu > 0$, $h = -\mu/4 < 0$. Therefore, if the marginal cost of manipulation grows with quality, manipulation activity *reduces* the informativeness of the forum. As μ tends to 2, h tends to -1/2 and the precision of the forum signal tends to zero. This means that, in settings with exogenously fixed prices, firm manipulation activity can severely reduce the value of an online forum to consumers. This result has practical importance because a lot of popular forums (e.g. for music, movies, etc.) are associated with markets where prices are essentially exogenously fixed. The result is consistent with similar findings by Mayzlin (2003).

Note that h is independent of the unit cost λ and the ratio of consumer participation r. This means that, as λ grows, forum informativeness (and, thus, consumer surplus) remains unchanged, whereas, as r grows, forum informativeness grows. In this setting, increases in consumer participation are more beneficial to consumers than the development of technologies that make it more difficult for firms to manipulate.

Since h is independent of λ and r, everything else being equal, firms would prefer to spend fewer resources on manipulation. In common with the case of endogenous prices we find that, as λ and r grow, manipulation intensities decline and firm profits grow.

Theorem 8 *If $\mu < 2$, as λ and/or r grow:*

1. *Manipulation intensities η_j decline*

2. *If, additionally, N is sufficiently high, firm profits grow.*

5 The general case

The results of the preceding sections were obtained by assuming specific payoff functions and signal distributions. This section distills general conditions that can be used to determine whether strategic manipulation of an Internet forum will increase or decrease the information value of that forum to its readers for a broad class of payoff functions and signal distributions.

Consider an Internet forum where users post honest but noisy opinions about some attribute of a set of firms. Denote by ω_j the true value of the target attribute of firm j, $j = 1, .., M$ (the firm's "type"), by ω_{-j} the vector of every other firm's type, and by ω the vector of all firms' types (the "state of the world"). Assume that each firm's type is independent of every other firm's type. Consumers read online opinions and use them to take a multidimensional action $a = (a_1, .., a_K)$ that affects both their own payoffs and those of the firms. A consumer's payoff depends both on her action and on the true state of the world. Firms manipulate the forum trying to influence consumers' posterior beliefs about their type.

In the absence of manipulation, average online opinions can be thought of as random variables x_j drawn from information structures $f_j(x_j|\omega_j)$. I assume that the result of a firm's manipulation activity $\xi_j(\omega_j, \omega_{-j})$ is to shift the probability distribution of that firm's average ratings so that it mimics the distribution of another type's ratings. Specifically, let the resulting distribution be $g_j(y_j|\omega_j, \omega_{-j}) = f_j(y_j|\omega_j + \xi_j(\omega_j, \omega_{-j}))$. The general question is, under what conditions will signal y_j be more (less) informative than x_j with respect to the decision problem faced by consumers.

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 payoff function has single-crossing incremental returns and posterior beliefs have the monotone likelihood ratio property. Building on Lehmann's work, I derive the following result:

Theorem 9 *Assume that consumers' payoff functions have single-crossing incremental returns*[13] *in*

[13]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.

(a_i, ω_j) for all $i = 1, .., K$, $j = 1, .., M$. Assume, further, that all $f_i(\cdot|\cdot)$ belong to location families that have the monotone likelihood ratio property (MLRP). The following statements are true:

1. *Online manipulation increases the value of the forum to consumers if the amount of manipulation $\xi_j(\omega_j, \omega_{-j})$ is a monotonically increasing function of each firm's own type (i.e. if $\partial\xi_j(\omega_j, \omega_{-j})/\partial\omega_j \geq 0$ for all j and all ω_j).*

2. *Online manipulation decreases the value of the forum to consumers if the amount of manipulation $\xi_j(\omega_j, \omega_{-j})$ is a monotonically decreasing function of each firm's own type (i.e. if $\partial\xi_j(\omega_j, \omega_{-j})/\partial\omega_j \leq 0$ for all j and all ω_j).*

Intuitively, if $\partial\xi_j(\omega_j, \omega_{-j})/\partial\omega_j \geq 0$ for all j, manipulation activity *spreads out* the means $\omega_j + \xi_j(\omega_j, \omega_{-j})$ 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 $\partial\xi_j(\omega_j, \omega_{-j})/\partial\omega_j \leq 0$ for all j, manipulation activity *condenses* the means 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.

Our next result connects the informativeness criterion of Theorem 9 to properties of firms' payoff functions. This connection allows us to understand in what settings we can expect online forum manipulation to benefit or harm consumers.

Theorem 10 *A necessary and sufficient condition for the existence of equilibria where firm j's manipulation intensity $\xi_j(\omega_j, \omega_{-j})$ is a monotonically increasing (decreasing) function of that firm's type ω_j is that the firm's expected payoff function (inclusive of the cost of manipulation) satisfies the single crossing property in (η_j, ξ_j) (or $(\eta_j, -\xi_j)$ respectively).*

If a firm's payoff function is differentiable, a sufficient condition for Theorem 10 to hold is that the cross-partial derivative of the payoff function with respect to ξ_j and ω_j is always positive (negative). Assume that a firm's payoff function is twice differentiable and can be written as:

$$v_j = w_j(a(\omega_j + \xi_j + \varepsilon_j, \omega_{-j} + \xi_{-j} + \varepsilon_{-j})) - c(\xi_j, \omega_j, \omega_{-j})$$

where $a(\cdot)$ is a scalar measure of the consumers' collective action after observing signal vector $y = (\omega_j + \xi_j + \varepsilon_j, \omega_{-j} + \xi_{-j} + \varepsilon_{-j})$, $w_j(a)$ is firm j's payoff from consumer action a, and $c(\cdot)$ is the cost of manipulation. Differentiating we obtain:

$$\partial^2 v_j / \partial \xi_j \partial \omega_j = w_j''(a(\omega_j + \xi_j + \varepsilon_j, \omega_{-j} + \xi_{-j} + \varepsilon_{-j}))[a_1(\omega_j + \xi_j + \varepsilon_j, \omega_{-j} + \xi_{-j} + \varepsilon_{-j})]^2 \\ + w_j'(a(\omega_j + \xi_j + \varepsilon_j, \omega_{-j} + \xi_{-j} + \varepsilon_{-j}))a_{11}(\omega_j + \xi_j + \varepsilon_j, \omega_{-j} + \xi_{-j} + \varepsilon_{-j}) - c_{12}(\xi_j, \omega_j, \omega_{-j})$$

Assuming $a_1(\cdot) \geq 0$, $w_j'(\cdot) \geq 0$, and $c_{12}(\cdot) \geq 0$, for the above cross-partial derivative to be positive it must be $w_j''(\cdot) > 0$ and/or $a_{11}(\cdot) > 0$. Therefore, forum manipulation is more likely to increase the information value of a forum in settings where firm payoffs are increasing convex functions of consumer actions and/or where consumer actions are increasing convex functions of observed signals. Such types of functions arise, for example, in settings where firms have increasing marginal returns from higher consumer actions and/or where consumer marginal returns from higher actions (or from higher values of the attribute being discussed in the forum) are supermodular in action and attribute value[14]. Settings with economies of scale and network effects often exhibit such properties.

The above results often enable us to characterize the consequences of strategic manipulation in a broad class of settings without having to derive the precise form of the corresponding equilibria. For example, some of the results of Theorem 2 can be derived by setting $\omega_1 = \omega_2 = q$, $\xi_1 = \eta_1$, and $\xi_2 = -\eta_2$. Given consumer actions $a \in \{\text{purchase A, purchase B}\}$, consumer utility functions (1) have single-crossing incremental returns in (a, q). Furthermore, all signals are normally distributed and, hence, satisfy the MLRP. According to Theorem 9, manipulation increases (decreases) forum informativeness if $\partial\xi_1/\partial q \geq 0$ and $\partial\xi_2/\partial q \geq 0$ ($\partial\xi_1/\partial q < 0$ and $\partial\xi_2/\partial q < 0$). If sellers follow linear manipulation strategies $\eta_j = g \pm hq$, the conditions of Theorem 9 are equivalent to $h \geq 0$ ($h < 0$). If buyers believe that sellers follow linear strategies, their posterior beliefs regarding q are normally distributed with mean given by (6). The expected value of firm profits in period 1 is then equal to:

$$v_j = -c_j + E[w_j] = -\lambda N r \left(\xi_j^2 + \mu \left(\frac{q}{2} \pm \frac{q_A + q_B}{2} \right) \xi_j \right) + N \left(\frac{1}{4} \pm \frac{E[\theta]}{3} + \frac{E[\theta^2]}{9} \right)$$

[14] Given consumer utility $u(a, \omega)$ supermodularity of marginal consumer returns in action and attribute value corresponds to $u_{112}(a, \omega) \geq 0$ and $u_{122}(a, \omega) \geq 0$. These conditions induce utility-maximizing consumer actions that are convex in ω.

where:

$$E\left[\theta\right]=\frac{\rho_z E[z]}{\tau+\rho_z}=\frac{\rho_z}{\tau+\rho_z}\frac{q+\xi_A+\xi_B}{2h+1},\quad E\left[\theta^2\right]=\left(E[\theta]\right)^2+V[\theta]=\left(E[\theta]\right)^2+\frac{\rho_z}{\left(\tau+\rho_z\right)^2}$$

Taking cross partial derivatives we obtain:

$$\frac{\partial^2 v_j}{\partial q \partial \xi_j}=-\frac{Nr\lambda\mu}{2}+\frac{2N\rho_z^2}{9(\tau+\rho_z)^2(2h+1)^2} \tag{8}$$

According to Theorem 10, equilibria where $\partial\xi_j/\partial q = h \geq 0$ exist if and only if $\partial^2 v_j/\partial q\partial\xi_j \geq 0$. From (8), for the above to be non-negative it must be $\mu \leq 4\rho_z^2/9\lambda r(\tau+\rho_z)^2(2h+1)^2$ for some non-negative h. Substituting $\rho_z = \rho\left(2h+1\right)^2/2$ and differentiating with respect to h one can verify that, for $\rho > \tau$, the right hand size of the above inequality is a decreasing function of h. For $h = 0$ the inequality becomes $\mu \leq 4\rho^2/9\lambda r(2\tau+\rho)^2$. Therefore, if $\mu \leq 4\rho^2/9\lambda r(2\tau+\rho)^2$, there exists some $\overline{h} \geq 0$ such that (8) is non-negative for $0 \leq h \leq \overline{h}$, whereas if $\mu > 4\rho^2/9\lambda r(2\tau+\rho)^2$ then $\partial^2 v_j/\partial q\partial\xi_j < 0$ for all positive h. This means that, for large μ, no linear PBE where $h \geq 0$ can exist. Substituting $\rho = Nr$ we can verify that these results are consistent with those of Theorem 2.

6 Discussion

This paper offers a systematic exploration of the issues surrounding the strategic manipulation of online opinion forums by firms whose products are being discussed in them. The principal results can be summarized as follows:

- Strategic manipulation of online forums can either increase or decrease the information value of a forum to consumers. Manipulation increases informativeness in settings where (i) the marginal revenue gains to firms from higher consumer beliefs about their quality are increasing functions of consumer beliefs, and (ii) the marginal cost of manipulation does not grow too steeply with a firm's true quality.

- If the number of consumers who post honest online reviews is sufficiently large, forum manipulation is harmful to firms because its cost outweighs its benefits. Nevertheless, competing firms are locked into an "arms race" and forced to spend resources on such activities to prevent each other from distorting consumer beliefs about their relative qualities.

- The cost of manipulation to firms can be reduced by developing technologies that make it more difficult to manipulate. Such technologies, however, do not necessarily increase the value of a forum to consumers.

- The most effective way to increase the value of a forum to all parties involved is the encouragement of higher levels of active consumer participation.

The results have interesting implications for all affected parties.

Consumers. Two results are of particular relevance to consumers. First, there exist settings where forum manipulation activity increases the information value of a forum to consumers. In such settings, the presence of extraordinary amounts of "hype" about a product is a signal that the product is, indeed, good. Second, by contributing online opinions, consumers increase the expected information value of a forum to future consumers, as well as help reduce "wasteful" firm manipulation activity. Consumers should therefore post more opinions in online forums and perhaps expect to be compensated for their efforts[15].

Producers. Contrary to intuition our analysis shows that, in large-scale markets, strategic manipulation of online forums is a wasteful activity that reduces the profits of all firms. Firms will benefit it they can manipulate less. Firms should, therefore, support technological initiatives that make forum manipulation more costly[16] and actively encourage their customers to contribute (honest) opinions to established forums[17].

Forum Operators. Online forums operators acquire a pivotal role in this new competitive environment. By investing in technologies that increase the cost of forum manipulation, operators can help reduce the manipulation expenditures of all competing firms in their industry. Even more importantly, by introducing mechanisms that motivate more consumers to post honest opinions

[15] Avery, Resnick and Zeckhauser (1999) note that online opinions constitute a public good and are, thus, expected to be undersupplied. Concrete incentives might therefore be needed to increase honest consumer participation in online forums.

[16] The development of such technologies is the topic of active research. Proposed approaches include statistical filtering of posted ratings (Dellarocas 2000), side-payment mechanisms that provide incentives to submit honest feedback (Miller, Resnick, and Zeckhauser 2002) 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).

[17] An example of a company that is already doing this is Tivo, the makers of a TV recording device. TiVo sends users to a specific, independently run site to register their compliments and grievances. TiVo personnel not only monitor the forum, but also occasionally post responses to complaints and concerns.

online, operators increase the information value of forums to consumers and decrease firm expenditures. The results of this paper provide a quantitative estimate of the business value of such investments as well as hints as to who should bear their cost (firms are the primary beneficiaries from increases in the cost of manipulation; both future consumers and firms benefit from increased consumer participation).

The results of this paper are robust to several perturbations of the modeling assumptions. Appendix C shows that the qualitative nature of the results does not change if we: (i) also allow firms to anonymously bad-mouth one another, (ii) assume that, in addition to the forum, consumers have access to outside sources of quality information, and (iii) assume that several, smaller forums (e.g. web logs, newsgroups) co-exist and that each consumer only accesses a random subset of them.

There are several promising directions for future research. The current model assumes that the only decision variable of the firm is its forum manipulation strategy. In actual practice firms are using a variety of quality signals. It will be interesting to study how the increasing importance of online forums affects the use of more traditional quality signals, such as prices and advertising. Furthermore, my analysis makes the extreme assumption that consumers cannot distinguish the source of individual opinions and, therefore, weigh all opinions equally in their inference process. Many forums provide social cues that allow readers of anonymous reviews to gauge the trustworthiness and level of expertise of their authors[18]. Exploring the consequences of strategic manipulation in such, more complex, environments is a natural next step of this line of work. Finally, this paper analyzed 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 question for future research.

References

Athey, S. (2001) Single Crossing Properties and the Existence of Pure-Strategy Equilibria in Games of Incomplete Information. *Econometrica*, 69(4), 861-889.

[18]For example, some forums allow readers to rate the usefulness of posted reviews and display average "usefulness" ratings alongside each review. Other forums post reputation scores next to the usernames of review authors and highlight authors whose past contributions have been particularly well received by the user community.

Athey, S. and Levin, J. (2001) The Value of Information in Monotone Decision Problems. Working Paper.

Avery, C., Resnick, P., Zeckhauser, R. (1999) The Market for Evaluations. *American Economic Review.* 89 (June), 564-84.

Blackwell, D. (1951) Comparison of Experiments, in J. Neyman (ed.) Proc. Second Berkeley Symposium on Mathematical Statistics and Probability. Univ. California Press.

Caminal, R. and Vives, X. (1996) Why Market Shares Matter: An Information-Based Theory. *RAND Journal of Economics*, 27 (2), 221-239.

Chevalier, J. and Mayzlin, D. (2003) The Effect of Word of Mouth on Sales: Online Book Reviews, Yale School of Management Working Paper.

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. *Proceedings of the 2nd ACM Conference on Electronic Commerce.* Association for Computing Machinery, Minneapolis, MN, 150-157.

Dellarocas, C., Fan, M., and Wood, C. (2003) Self-Interest, Reciprocity, and Participation in Online Reputation Systems. 2003 *Workshop in Information Systems and Economics* (WISE), Seattle, WA.

Friedman, E., P. Resnick (2001) The Social Cost of Cheap Pseudonyms. *Journal of Economics and Management Strategy* **10** (1), 173-199.

Fudenberg, D. and Tirole J. (1986) A Signal-Jamming theory of Predation. *Rand Journal of Economics* 17, 366-376.

Gu, B. and Jarvenpaa, S. (2003) Online Discussion Boards for Technical Support: The Effect of Token Recognition on Customer Contributions. *Proceedings of the 2003 International Conference on Information Systems* (ICIS), Association for Information Systems, Seattle, WA.

Harmon, A. (2004) Amazon Glitch Unmasks War of Reviewers. *The New York Times*, February 14, 2004.

Holmstrom, B. 1999. Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies* **66** (1) 169-182.

Judd, K. L. and Riordan, M. H. (1994) Price and Quality in a New Product Monopoly. *Review of Economic Studies*, 61, 773-789.

Kihlstrom, R. and Riordan, M. (1984) Advertising as a Signal. *Journal of Political Economy.* 92 (4), 427-450.

Lehmann E. L. (1988) Comparing Location Experiments. *The Annals of Statistics*, 16 (2), 521-533.

Mayzlin, D. (2003) Promotional Chat on the Internet. Yale School of Management Working Paper #MK-14, January 31, 2003.

McAdams, D. (2002) Isotone Equilibrium in Games of Incomplete Information, MIT Sloan Working Paper #4248-02.

Milgrom, P. and Roberts, J. (1986) Price and Advertising Signals of Product Quality. *Journal of Political Economy*, 94 (5), 796-821.

Milgrom, P. and Shannon, C. (1994) Monotone Comparative Statics. *Econometrica* 62 (1),157-180.

Miller, N., Resnick, P., R. Zeckhauser. (2002) Eliciting Honest Feedback in Electronic Markets, Research Working Paper RWP02-039, Harvard Kennedy School, Cambridge, MA

Mirman, L. J., Samuelson, L. and Schlee, E.E. (1994) Strategic Information Manipulation in Duopolies. *Journal of Economic Theory* 62, 363-384.

Riordan M. (1985) Imperfect Information and Dynamic Conjectural Variations. *Rand Journal of Economics* 16, 41-50.

Schindler, R. and Bickart, B. (2003) Published 'Word of Mouth': Referable, Consumer-Generated Information on the Internet, in C. Hauvgedt, K. Machleit and R. Yalch (eds.), *Online Consumer Psychology: Understanding and Influencing Behavior in the Virtual World.* Lawrence Erlbaum Associates.

Senecal, S. and Nantel, J. (2003) The Influence of Online Product Recommendations on Consumers' Online Choices. Working Paper.

Thompson, N. (2003) More Companies Pay Heed to Their 'Word of Mouse' Reputation. *The New York Times*, June 23, 2003.

White, E. (1999) Chatting a Singer Up the Pop Charts. *The Wall Street Journal*, October 5, 1999.

A Proofs

Theorem 1:

Substituting $M = q_A + q_B$, $p_A = \frac{1}{2} + \frac{\theta}{3}$, $p_B = \frac{1}{2} - \frac{\theta}{3}$, $\theta = \frac{\rho_z z}{\tau+\rho_z}$ the utility functions (1) of individual consumers can be written as:

$$u_A^i = \frac{M+q}{2} - \frac{i}{2} - \frac{1}{2} - \frac{\rho_z z}{3(\tau+\rho_z)}, \quad u_B^i = \frac{M-q}{2} - \frac{1-i}{2} - \frac{1}{2} + \frac{\rho_z z}{3(\tau+\rho_z)}$$

After observing signal z, consumers with indices $i \in [0, 1/2 + \theta/3]$ will purchase product A, whereas consumers with indices in $i \in (1/2+\theta/3, 1]$ will purchase product B. The average consumer utility given signal realization z and true quality difference q is:

$$E[u^i(q,z)] = \int_0^{1/2+\theta/3} u_A^i di + \int_{1/2+\theta/3}^{1} u_B^i di = \frac{M}{2} - \frac{5}{8} + \frac{\rho_z z}{3(\tau+\rho_z)} q - \frac{5\rho_z^2 z^2}{18(\tau+\rho_z)^2}$$

The average utility given true quality difference q and a normal signal $f(z|q)$ with mean q and precision ρ_z is:

$$\int E[u^i(q,z)]f(z|q)dz = \frac{M}{2} - \frac{5}{8} + \frac{\rho_z}{3(\tau+\rho_z)}q^2 - \frac{5\rho_z^2}{18(\tau+\rho_z)^2}\left(q^2 + \frac{1}{\rho_z}\right)$$

The ex-ante average utility given quality differences q normally distributed according to $h(q)$ with mean zero and precision τ is:

$$U = \int h(q) \int E[u^i(q,z)]f(z|q)dzdq = \frac{M}{2} - \frac{5}{8} + \frac{\rho_z}{3(\tau+\rho_z)}\frac{1}{\tau} - \frac{5\rho_z^2}{18(\tau+\rho_z)^2}\left(\frac{1}{\tau} + \frac{1}{\rho_z}\right)$$

Differentiating with respect to ρ_z we obtain $\partial U/\partial \rho_z = 1/18(\tau+\rho_z)^2 > 0$. Therefore, the ex-ante average consumer utility grows with the precision of the signal z.

Theorem 2:

At the end of period 0 firms select manipulation strategies that maximize the expected value of profits in period 1:

$$v_j = -c_j + E\left[w_j\right] = -\lambda N r\left(\eta_j^2 + \mu\left(\frac{M}{2} \pm \frac{q}{2}\right)\eta_j\right) + N\left(\frac{1}{4} \pm \frac{E\left[\theta\right]}{3} + \frac{E\left[\theta^2\right]}{9}\right)$$

where:

$$E\left[\theta\right] = \frac{\rho_z E[z]}{\tau+\rho_z} = \frac{\rho_z}{\tau+\rho_z}\frac{q+\eta_A-\eta_B}{2h+1}, \quad E\left[\theta^2\right] = (E[\theta])^2 + V[\theta] = (E[\theta])^2 + \frac{\rho_z}{(\tau+\rho_z)^2}$$

$$M = q_A + q_B, \quad q = q_A - q_B, \quad \eta_j \geq 0, \quad v_j \geq 0, \quad \rho_z = \rho\left(2h+1\right)^2/2$$

First-order conditions result in a system of two linear equations on η_A, η_B. Solving these equations we obtain:

$$\eta_A = \left(\frac{\rho_z}{6(\tau+\rho_z)(2h+1)r\lambda} - \frac{\mu M}{4}\right) + hq, \quad \eta_B = \left(\frac{\rho_z}{6(\tau+\rho_z)(2h+1)r\lambda} - \frac{\mu M}{4}\right) - hq \qquad (9)$$

Therefore, if consumers believe that firms follow linear manipulation strategies $\eta_j = g \pm hq$, firms will indeed find it optimal to adopt such strategies[19]. Consistency of beliefs requires that the conjectured and optimal h are equal. From the above system of linear equations, this is equivalent to requiring that h is the solution of the equation:

$$\lambda = \frac{4\rho_z^2}{9(\tau + \rho_z)^2(2h+1)(4h+\mu)r} \tag{10}$$

Substituting $\rho_z = \rho\,(2h+1)^2/2$, (10) becomes equivalent to a fifth degree polynomial equation:

$$\lambda = \frac{4N^2 r(2h+1)^3}{9(2\tau + Nr(2h+1)^2)^2(4h+\mu)} \tag{11}$$

The right-hand side of (11) tends to zero as $h \to -1/2$. Therefore, for $\lambda > 0$, $h = -1/2$ can never be an equilibrium solution (i.e. the forum always remains somewhat informative).

Rearranging (10) we see that all of its solutions must also solve:

$$8h^2 + 2(2+\mu)h + \left(\mu - \frac{4\rho_z^2}{9\lambda r(\tau+\rho_z)^2}\right) = 0 \tag{12}$$

Given that the coefficients of h^2 and h are both positive, equation (12) has one positive real solution if and only if the constant term is negative. This requires that $\mu < \frac{4\rho_z^2}{9\lambda(\tau+\rho_z)^2 r}$. Substituting $\rho_z = \rho\,(2h+1)^2/2$ and differentiating with respect to h one can verify that, for $\rho > \tau$, the right hand size of the above inequality is a decreasing function of h. For $h = 0$ the inequality becomes $\mu \leq 4\rho^2/9\lambda r(2\tau+\rho)^2$. Therefore, if $\mu < 4\rho^2/9\lambda r(2\tau+\rho)^2$, there always exists a positive h that solves (12). Substituting $\rho = Nr$ the requirement becomes $\mu < \frac{4N^2 r}{9\lambda(2\tau+Nr)^2}$. If $\mu = 4\rho^2/9\lambda r(2\tau+\rho)^2$ the only solution of (12) is $h = 0$. If $\mu > 4\rho^2/9\lambda r(2\tau+\rho)^2$ the constant term of (12) becomes positive for all positive h, hence there can be no positive real roots. It can be shown that the discriminant also remains positive. Therefore, for large μ equation (12) has two negative real roots, corresponding to situations where the low quality firm manipulates more intensely than the high quality firm.

Second order conditions require that $\lambda > \frac{\rho_z^2}{9(\tau+\rho_z)^2(2h+1)^2 r}$. Combining this condition with (10) we obtain the equivalent condition $\frac{4h+\mu}{4(2h+1)} < 1$. For $h > -1/2$ this, in turn, is equivalent to the

[19]We only consider interior solutions and implicitly assume that parameters are chosen to satisfy $\eta_j \geq 0$.

simpler condition $\mu < 4(h+1)$.

Theorems 3 and 4:

Taking cross-partial derivatives $\partial^2 v_j/\partial\eta_j\partial\lambda = -Nr(2\eta_j + \mu q_j) < 0$. By Milgrom and Shannon (1994) this implies that the optimum η_j declines with λ. Likewise, it is $\partial^2 v_A/\partial h\partial\lambda = -Nr(\frac{\partial g}{\partial h} + q)(2\eta_A + \mu q_A) < 0$, which implies that h declines with λ as long as $\partial g/\partial h > 0$, when g is expressed as a function of h only. Eliminating λ from the expression for g in equation (9) by substituting (10) and differentiating with respect to h we can verify that $\partial g/\partial h > 0$ if $Nr > \max(2\tau/3, 2\tau(\mu-1))$.

The cross-partial derivatives $\partial^2 v_j/\partial\eta_j\partial r$ and $\partial^2 v_j/\partial h\partial r$ are more complicated. However, they can be reduced to the ratio of two polynomial expressions of the market size N such that the denominator is always positive and the leading coefficient of the numerator is negative. This implies that, for N sufficiently large, $\partial^2 v_j/\partial\eta_j\partial r < 0$ and $\partial^2 v_j/\partial h\partial r < 0$. By Milgrom and Shannon (1994), this, in turn, implies that, if the size N of the market is sufficiently large, increases in the level of participation r decrease the firms' intensity of manipulation.

Differentiating (11) we can obtain direct expressions for the derivatives of h with respect to λ and r:

$$\frac{\partial h}{\partial\lambda} = -\frac{9(4h+\mu)^2(Nr(2h+1)^2+2\tau)^3}{8rN^2(2h+1)^2(Nr(2h+1)^2(8h+2+\mu)-2\tau(3\mu+8h-2))} \tag{13}$$

Negative if $Nr > \max(4\tau/(4+\mu), (6\mu-4)\tau/(\mu+2))$

$$\frac{\partial h}{\partial r} = -\frac{(2h+1)(4h+\mu)(Nr(2h+1)^2-2\tau)}{2r(2h+1)^2(Nr(2h+1)^2(8h+2+\mu)-2\tau(3\mu+8h-2))} \tag{14}$$

Negative if $Nr > \max(2\tau, (6\mu-4)\tau/(\mu+2))$

If $\mu < 2$ the above conditions are satisfied if $Nr/2 > \tau$, i.e. if the precision of honest online ratings is at least as high as that of the prior.

Theorem 5:

The precision of the forum signal is equal to $\rho_z = \rho\left(2h(r)+1\right)^2/2 = Nr\left(2h(r)+1\right)^2/2$. Differentiating with respect to r we obtain:

$$\frac{\partial\rho_z}{\partial r} = \frac{N(2h+1)}{2}\left(2h+1+4r\frac{\partial h}{\partial r}\right)$$

The above is positive iff $\Xi = 2h + 1 + 4r\frac{\partial h}{\partial r} > 0$. Substituting (14) we obtain:

$$\Xi = \frac{(2-\mu)(2h+1)(Nr(2h+1)^2 + 2\tau)}{Nr(2h+1)^2(8h+2+\mu) - 2\tau(3\mu+8h-2)}$$

which is positive for $\mu < 2$ and $Nr > \max(4\tau/(4+\mu), (6\mu-4)\tau/(\mu+2))$. For $\mu < 2$ the latter condition of Nr is subsumed by $Nr/2 > \tau$.

Theorem 6:

Substituting (10) into the expressions for η_j we can eliminate λ and write the profit functions v_j as expressions of h. Given that, for N sufficiently large, we have established that $\partial h/\partial\lambda < 0$, to prove that $\partial v_j/\partial\lambda > 0$ it suffices to show that $\partial v_j/\partial h < 0$. The derivative $\partial v_j/\partial h$ can be expressed as the ratio of two high degree polynomials of N. The denominator polynomial is always positive. The leading coefficient of the numerator polynomial is negative if $\mu < \frac{6+4q}{3+4q+2(q_A+q_B)} \leq 2$. This means that if μ is not too large, $\partial v_j/\partial h < 0$ for N sufficiently large.

The derivative $\partial v_j/\partial r$ can be calculated directly. Similar to $\partial v_j/\partial h$, it can be expressed as the ratio of two high degree polynomials of N. The denominator polynomial is positive if $Nr > \max(4\tau/(4+\mu), (6\mu-4)\tau/(\mu+2))$. The leading coefficient of the numerator polynomial is positive if $\mu < \frac{6+4q}{3+4q+2(q_A+q_B)} \leq 2$. For $\mu < 2$ the above condition of Nr is subsumed by $Nr/2 > \tau$. This means that if μ is not too large, $\partial v_j/\partial r > 0$ for N sufficiently large.

Theorem 7:

At the end of period 0 each firm maximizes the expected value of its profits in period 1:

$$v_j = -c_j + E\left[w_j\right] = -\lambda Nr\left(\eta_j^2 + \mu\left(\frac{M}{2} \pm \frac{q}{2}\right)\eta_j\right) + Np\left(\frac{1}{2} \pm E\left[\theta\right]\right)$$

where

$$E\left[\theta\right] = \frac{\rho_z E[z]}{\tau + \rho_z} = \frac{\rho_z}{\tau + \rho_z}\frac{q + \eta_A - \eta_B}{2h+1},\ M = q_A + q_B,\ q = q_A - q_B,\ \rho_z = \rho\left(2h+1\right)^2/2$$

First-order conditions result in a system of two linear equations on η_A, η_B. Solving these equations gives:

$$\eta_A = \left(p\frac{N(2-\mu)}{\lambda(8\tau + Nr(2-\mu)^2)} - \frac{\mu M}{4}\right) - \frac{\mu}{4}q, \quad \eta_B = \left(p\frac{N(2-\mu)}{\lambda(8\tau + Nr(2-\mu)^2)} - \frac{\mu M}{4}\right) + \frac{\mu}{4}q \quad (15)$$

Second-order conditions $-2Nr\lambda < 0$ are trivially satisfied. We implicitly assume that parameters satisfy $\eta_A, \eta_B \geq 0$. Among other things this requires that $\mu < 2$. It follows that, if consumers believe that firms follow linear manipulation strategies $\eta_j = g \pm hq$, firms will indeed find it optimal to adopt such strategies. Notice that $h = -\mu/4 < 0$, which means that, for $\mu > 0$, the low quality firm manipulates more than the high quality firm, reducing the informativeness of the forum. For $\mu = 0$ manipulation does not affect informativeness.

Theorem 8:

Since $h = -\mu/4$ is independent of λ and r, for $\mu < 2$:

$$\frac{\partial \eta_j}{\partial \lambda} = \frac{\partial g}{\partial \lambda} = -p\frac{N(2-\mu)}{\lambda^2(8\tau + Nr(2-\mu)^2)} < 0 \quad \frac{\partial \eta_j}{\partial r} = \frac{\partial g}{\partial r} = -p\frac{N^2(2-\mu)^3}{\lambda(8\tau + Nr(2-\mu)^2)^2} < 0$$

Substituting (15) into the expressions of v_j and differentiating with respect to λ and r, the derivatives $\partial v_j/\partial \lambda$, $\partial v_j/\partial r$ can be expressed as the ratio of two high degree polynomials of N where the denominator polynomial and the leading coefficient of the numerator polynomial are positive. This implies that for N sufficiently large $\partial v_j/\partial \lambda > 0$ and $\partial v_j/\partial r > 0$.

Theorem 9

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:

$$G^{-1}[F(x|\omega)|\omega] \text{ is a nondecreasing function of } \omega \text{ for all } x \quad (16)$$

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 (16) can be equivalently written as requiring that the function:

$$y(\omega, x) \text{ such that } G(y(\omega, x)|\omega) = F(x|\omega) \text{ is a nondecreasing function of } \omega \text{ for all } x \qquad (17)$$

Let $x = (x_1, .., x_n)$ and $y = (y_1, .., y_n)$ 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 Lehman's 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:

$$y_j(\omega_j, x_j) \text{ such that } G_j(y_j(\omega_j, x_j)|\omega_j) = F_j(x_j|\omega_j) \text{ is a nondecreasing function of } \omega_j \text{ for all } x_j \qquad (18)$$

In our setting, $G_j(y_j|\omega_j) = F_j(y_j|\omega_j + \xi_j(\omega_j, \omega_{-j}))$. If we make the additional assumption that distributions F and G belong to location families then $F_j(x_j|\omega_j) = F_j(x_j - \omega_j)$, $G_j(y_j|\omega_j) = F_j(y_j - \omega_j - \xi_j(\omega_j, \omega_{-j}))$. Substituting into (18):

$$G_j(y_j(\omega_j, x_j)|\omega_j) = F_j(x_j|\omega_j) \Leftrightarrow F_j(y_j(\omega_j, x_j) - \omega_j - \xi_j(\omega_j, \omega_{-j})) = F_j(x_j - \omega_j) \Leftrightarrow$$
$$y_j(\omega_j, x_j) = x_j + \xi_j(\omega_j, \omega_{-j})$$

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

Theorem 10:

The theorem is a simple corollary of the results of Athey (2001) and McAdams (2002).

B Properties of forum manipulation cost function

This appendix presents a setting in which the properties of the manipulation cost function postulated in Section 2 arise as consequences of the common-sense assumption that the cost of manipulation is proportional to the number of fake reviews that a firm must post to inflate its average ratings by a fixed amount.

Suppose that online ratings are binary assessments of a consumer's satisfaction with a product, that is, that consumers rate a product as either "satisfactory" or "unsatisfactory". Suppose, further, that each product's quality $q_j \in [0, 1]$ can be interpreted as that product's reliability (probability of performing in a satisfactory manner). Online forums publish the fraction of "satisfactory" ratings over the total number of ratings posted for each product during a given period. Let Nr be the total number of consumers who post ratings for firm j during a given period. In the absence of manipulation, the expected number of satisfactory ratings posted for firm j will be equal to Nrq_j. In order for firm j to inflate the expected quality rating of its product by η_j, it must post z promotional ratings, so that:

$$E[y_j] = \frac{Nrq_j + z}{Nr + z} = q_j + \eta_j$$

This gives:

$$z = Nr\frac{\eta_j}{1 - q_j - \eta_j} \approx Nr\eta_j\left(1 + q_j + \eta_j\right) + ...$$

that is, the required number of promotional ratings is (i) proportional to Nr (Property 1) and (ii) a convex function of η_j (Property 2). Furthermore, for a fixed η_j the required number of promotional ratings is increasing in q_j (Property 3). Under the assumption that each promotional rating has a fixed cost λ, the total cost of manipulation exhibits the three properties postulated in Section 2.

C Model extensions

This appendix shows that the results of Section 3 are qualitatively robust to a number of perturbations of the original set of assumptions. Specifically, I consider the following three extensions of my original set of assumptions:

1. In addition to positive reviews about their own product, firms can anonymously post fake negative reviews about their competitor's product.

2. In addition to visiting online forums, each period's consumers obtain additional information about a product's quality from friends, magazine articles and other sources that firms cannot manipulate.

3. There are several forums where the products of the competing firms are discussed; each consumer only visits a random subset of those forums.

Negative manipulation. Assume that, in addition to promotional reviews that praise their own product, firms A and B can post negative reviews about their competitor's product with the objective of *reducing* the competitor's average ratings. Denote by ζ_A, ζ_B the amounts by which firms B and A wish to reduce the average ratings of their competitors respectively[20]. Assume, further, that the cost of negative manipulation of firm j's ratings (borne by its competitor) is given by $c'(q_j, \zeta_j) = \alpha\lambda Nr(\zeta_j^2 + \beta\mu q_j \zeta_j)$, where the new constants α, β intend to capture the fact that the cost of negative manipulation may have different parameters than the cost of positive manipulation.

As in the original model, I assume normally distributed prior beliefs with mean zero and linear manipulation strategies $\eta_A = g^+ + h^+ q$, $\eta_B = g^+ - h^+ q$, $\zeta_A = g^- + h^- q$, $\zeta_B = g^- - h^- q$. Then, published ratings are equal to $y_j = q_j + \eta_j - \zeta_j + \epsilon_j$ and the statistic $z = (y_A - y_B)/(2(h^+ - h^-) + 1)$ is an unbiased estimator of q with precision $\rho_z = Nr(2(h^+ - h^-) + 1)^2/2$. Substituting the above and following the steps of the proof of Theorem 2 we can verify the existence of linear PBE where manipulation strategies follow the postulated linear form and where $h = h^+ - h^-$ must solve:

$$\lambda \frac{\alpha}{1+\alpha} = \frac{4\rho_z^2}{9(\tau + \rho_z)^2(2h+1)(4h + (1-\beta)\mu)r} \tag{19}$$

subject to the second order condition $\lambda > \rho_z^2/9\min(1,\alpha)(\tau+\rho_z)^2(2h+1)^2 r$. Manipulation increases forum informativeness if and only if $h > 0$. Such solutions exist if

$$(1-\beta)\mu < \frac{4rN^2}{9\lambda\frac{\alpha}{1+\alpha}(2\tau + Nr)^2}$$

Comparison of the above equations with the results of Section 3 reveals that the assumption of negative manipulation is mathematically equivalent to replacing the parameters λ, μ of the original

[20] The subscript indicates the firm whose ratings are being manipulated (by the other firm).

cost function with $\lambda' = \lambda\frac{\alpha}{1+\alpha}$ and $\mu' = (1-\beta)\mu$ respectively. Intuitively, the two firms will use a combination of positive and negative manipulation. Observe that, in terms of affecting consumer perceptions about the quality difference, posting negative reviews about one's competitor is qualitatively equivalent to posting positive reviews about one's own product. The following points then describe the impact of parameters α and β:

- If α is small then firms will predominantly use negative manipulation. The cumulative unit cost parameter then becomes $\lambda' \simeq \lambda\alpha$.
- If α is large then firms will predominantly use positive manipulation. The cumulative unit cost parameter then becomes $\lambda' \simeq \lambda$.
- If $\beta > 0$, the marginal cost of negative manipulation increases with the competitor's quality. It is, therefore, more expensive for the low quality firm to badmouth the high quality firm, whereas the high quality firm will find it easier to badmouth the low quality firm. The cumulative effect is to reduce the rate at which the cumulative marginal cost of manipulation grows with a firm's quality and, thus, make it more likely that manipulation will increase informativeness.
- If $\beta < 0$, the marginal cost of negative manipulation decreases with the competitor's quality. It is, therefore, less expensive for the low quality firm to badmouth the high quality firm, whereas the high quality firm will find it more difficult to badmouth the low quality firm. The cumulative effect is to increase the rate at which the cumulative marginal cost of manipulation grows with a firm's quality and, thus, make it *less likely* that manipulation will increase informativeness.

Outside information sources. In addition to visiting online forums, consumers often obtain product quality information from a variety of other sources, such as discussions with friends, magazine articles, etc. If one can assume that these sources are beyond the control of the competing firms (and thus cannot be manipulated), an elegant way in which their influence can be factored into our model is by incorporating it into the prior beliefs of consumers. Specifically, instead of assuming that prior beliefs about the quality difference q have zero mean, one can assume that each period's prior beliefs have mean z_t, where z_t is exogenous and known to firms.

The assumption of non-zero prior beliefs implies that period 0's demand will be different for the two firms and equal to $x_{j,0} = N\left(\frac{1}{2} \pm \frac{1}{3}z_0\right)$. This also implies that the number of (honest) ratings that will be posted for firms A and B at the end of period 0 will also be different. Each firm's manipulation cost function then becomes a function of period 0's demand:

$$c_A(q_A, \eta_A) = \lambda N\left(1 + \frac{2}{3}z_0\right) r(\eta_A^2 + \mu q_A \eta_A) \quad c_B(q_B, \eta_B) = \lambda N\left(1 - \frac{2}{3}z_0\right) r(\eta_B^2 + \mu q_B \eta_B)$$

Assume linear manipulation strategies $\eta_A = g_A + h_A q$, $\eta_B = g_B - h_B q$. Then the statistic $z = (y_A - y_B - (g_A - g_B))/(h_A + h_B + 1)$ is an unbiased estimator of q with precision $\rho_z = Nr\left(1 - \frac{4}{9}z_0^2\right)(h_A + h_B + 1)^2/2$. Finally, the expression for the mean of posterior beliefs regarding q in period 1 becomes $\theta = (\tau z_1 + \rho_z z)/(\tau + \rho_z)$. Substituting the above and following the steps of the proof of Theorem 2 we can verify the existence of linear PBE where manipulation strategies follow the postulated linear form and where $h = h_A + h_B$ must solve:

$$\lambda = \frac{4\rho_z^2}{9(\tau + \rho_z)^2(h+1)(2h+\mu)r(1 - \frac{4}{9}z_0^2)} \tag{20}$$

subject to the second order condition $\lambda > \rho_z^2/3\min(3+2z_0, 3-2z_0)(\tau+\rho_z)^2(h+1)^2 r$. Manipulation increases forum informativeness if and only if $h > 0$. Such solutions exist if

$$\mu < \frac{4rN^2(1 - \frac{4}{9}z_0^2)}{9\lambda(2\tau + Nr(1 - \frac{4}{9}z_0^2))^2}$$

Comparison of the above equations with the results of Section 3 reveals that the assumption of non-zero prior beliefs is mathematically equivalent to replacing the fraction r of consumers who submit feedback in the original model with $r' = r(1 - \frac{4}{9}z_0^2)$. The extended model reduces to the model of Section 3 for $z_0 = 0$.

Multiple online forums. The base model assumes that all consumers visit a single online forum and thus access (and contribute to) the same history of quality ratings. This set of assumptions describes settings where large amounts of consumer opinions are concentrated in a small number of popular websites (such as Amazon, Epinions, Citysearch, Yahoo, etc.). Although in some settings these assumptions correspond nicely to the current reality, there are other settings where a variety of smaller forums co-exist and where each consumer visits and contributes only to a subset of them. If we assume that prices are signal-free and that a consumer's decision to visit a forum is uncorrelated with her decision to visit any other forum as well as with her location in the taste interval, the presence of a set F of forums (indexed by $k = 1, ..., K$) does not substantially change our model. In fact, if we make the assumptions that, in addition to the total market size N, firms (i) know the number N_k of consumers who access each forum and the fraction r_k of each forum's visitors who contribute opinions, and (ii) search all forums at the beginning of each period and, thus, have precise knowledge of the aggregate ratings posted on each of them before they set prices, then a simple extension of the model of Section 3 applies to a setting with multiple forums.

The idea is that firms follow separate manipulation strategies η_{jk} for each individual forum and that each of these strategies is qualitatively similar to the equilibrium strategies derived in Section 3. Let $\rho_k = N_k r_k$ denote the base precision of forum k's aggregate (honest) ratings. Assume linear manipulation strategies $\eta_{jk} = g_k \pm h_k q$ and cost functions $c_k(q_j, \eta_{jk}) = \lambda_k N_k r_k(\eta_{jk}^2 + \mu_k q_j \eta_{jk})$. Given published ratings y_{jk} the posterior beliefs (regarding the quality difference q) of a consumer who visits a subset $S \subseteq F$ of forums have mean value $\theta_S = \sum_{k \in S} z_k \rho_{zk} / (\tau + \sum_{k \in S} \rho_{zk})$ where $\rho_{zk} = \rho_k \left(2h_k + 1\right)^2 / 2$ is the precision of published ratings obtained from forum k and $z_k = (y_{Ak} - y_{Bk})/(2h_k + 1)$ is an unbiased estimator of q based on ratings published in forum k. Given demand functions (2), maximization of expected sales revenues implies:

$$\begin{array}{ll} \text{prices} & p_j = \frac{1}{2} \pm \frac{1}{3} \sum_{S \in 2^F} n_S \theta_S \\ \text{sales revenues} & w_j = N \left(\frac{1}{2} \pm \frac{1}{3} \sum_{S \in 2^F} n_S \theta_S \right)^2 \end{array}$$

where 2^F denotes the powerset of F and $n_S = N^{-K} \prod_{k \in S} N_k \prod_{l \in F-S} (N - N_l)$ is the probability that a random consumer will visit forum subset S. Sales revenues can be equivalently expressed as:

$$w_j = N \left(\frac{1}{2} \pm \frac{1}{3} \sum_{k=1}^{K} m_k z_k \right)^2 \quad \text{where} \quad m_k = \rho_{zk} \sum_{S \in 2^F \wedge k \in S} \frac{n_S}{\tau + \sum_{l \in S} \rho_{zl}}$$

Substituting the above and following the steps of the proof of Theorem 2 we can verify the existence of linear PBE where manipulation strategies η_{jk} follow the postulated linear form.