Cooperation Without Enforcement? A comparative analysis of litigation and online reputation as quality assurance mechanisms

Yannis Bakos[*]

Stern School of Business
New York University
New York, NY 10012

Chrysanthos Dellarocas

Sloan School of Management
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

Online reputation mechanisms are emerging as a promising alternative to more traditional mechanisms for promoting trust and cooperative behavior, such as legally enforceable contracts. As information technology dramatically reduces the cost of accumulating, processing and disseminating consumer feedback, it is plausible to ask whether such mechanisms can provide an economically more efficient solution to a wide range of moral hazard settings where societies currently rely on the threat of litigation in order to induce cooperation. In this paper we compare online reputation to legal enforcement as institutional mechanisms in terms of their ability to induce cooperative behavior. Furthermore, we explore the impact of information technology on their relative economic efficiency. We find that although both mechanisms result in losses relative to the maximum possible social surplus, under certain conditions online reputation outperforms litigation in terms of maximizing the total surplus, and thus the resulting social welfare.

Draft of 10/15/02 4:22 AM

---

[*] The authors' names appear in alphabetical order.

1. Introduction

Economic activity requires economic agents to abide by the terms of explicit or implicit promises. For example, a merchant is expected to ship a good that has been purchased and paid for, or to provide the quality explicitly or implicitly promised to the customer.

Most commercial transactions rely on the legal system to assure performance of promises, which are written into explicit or implicit contracts. The legal system is expensive, however, in terms of the cost of the institutions necessary to adjudicate claims (lawyers, courts, etc.) and to enforce decisions (police, correctional facilities, etc.). It is also dependent on access to the enforcement power of a sovereign state.

Electronic markets operate on a global scale and typically span multiple jurisdictions. Litigation across jurisdictions is very costly and often infeasible. Online reputation mechanisms (Resnick, et. al., 2000) have emerged as a viable alternative to the legal system in such settings. On eBay, for instance, an online feedback mechanism that encourages buyers and sellers to rate one another seems to have succeeded in encouraging cooperative behavior in an otherwise very risky trading environment (Bajari and Hortascu, 2000; Dewan and Hsu, 2001; Houser and Wooders, 2000; Lucking-Reiley et. al. 2000; Resnick and Zeckhauser, 2001).

The potential applications of online reputation mechanisms go beyond the relatively narrow domain of trust building in electronic marketplaces. The appeal of reputation mechanisms is that, when they work, they facilitate cooperation without the need for costly enforcement institutions (Wilson 1985). They have, therefore, the potential of providing more economically efficient outcomes in a wide range of moral hazard settings where societies currently rely on the threat of litigation in order to induce cooperation.

The concept of reputation is as old as society itself. In the early middle ages, before the emergence of sovereign states of substantial geographical span, reputation networks were the primary mechanism for inducing cooperative behavior in European trade (Milgrom, North and Weingast, 1990). It was only during the sixteenth century that state-enforced commercial law took over as the primary mechanism for adjudicating trade disputes (Benson, 1989). Milgrom et.

al. argue that the primary reason for this evolution was economic: at those times, communication of information about a trader's past record was costly and error-prone. Therefore, once state enforcement became possible because of the emergence of extended sovereign states, it provided an economically more *efficient* solution to the problem of policing exchange.

Information technology is having dramatic impacts on the cost, impact and performance of reputation mechanisms. Online systems greatly reduce the cost of collecting and disseminating feedback information on a worldwide scale, thus facilitating wider participation and enabling the pooling of experiences of unrelated individuals into a single, easily accessible repository. This increases the likelihood that a feedback report for a specific transaction will have an influence on large numbers of future transactions, thus strengthening the impact of reputation effects. Finally, in contrast to the ad-hoc nature of traditional word-of-mouth networks, information technology allows the precise control of who can participate, what type of feedback is solicited, how it is aggregated and what type of reputation information is disseminated to the community. These design dimensions can be properly engineered in order to build systems that elicit honest feedback, minimize efficiency losses due to noisy reports and maintain robust outcomes in the presence of boundedly rational participants or strategic manipulation (see Dellarocas (2002a) for a survey of issues and results in these areas).

On the other hand, the impact of information technology on the cost of traditional enforcement mechanisms, such as courts, lawyers and the police, is likely to be moderate, at least as long as these remain primarily dependent on human labor. Given this difference in the likely impact of technology, it is timely and appropriate for IS researchers and economists to reconsider the relative merits of these two important classes of mechanisms.

The novelty of our work stems from its comparative focus; while both reputation and litigation mechanisms have been previously studied in the economics literature, ours is the first work we are aware of that studies these two types of institutional mechanisms in the same setting. This allows us to compare their ability to induce cooperative behavior, and to explore the impact of information technology on their relative economic efficiency.

Game theoretic analyses of litigation are the focus of a rich, and growing, body of literature (see Cooter and Rubinfeld (1989) and Chapter 8 of Baird, Gertner and Picker (1995) for literature surveys). The litigation model developed in this paper is a simple version of a one-sided private information model: a model in which the defendant has information that is not available to the plaintiff (in our case, the true level of effort exerted) but the plaintiff has no information to which the defendant does not have access. Examples of such models in the literature include Bebchuk (1984), Nalebuff (1987), Reinganum and Wilde (1986) and P'ng (1983).

Reputation formation has also been extensively studied by economists (see Wilson (1985) and Dellarocas (2002a) for surveys). Most traditional reputation models make the assumption that the entire public history of past play is available to all players and that new players infer a long-run player's reputation by repeated application of Bayes' rule on that information. The model of this paper is novel, inspired primarily by eBay's feedback mechanism. Our model emphasizes the ability of online reputation mechanisms to succinctly summarize large volumes of past feedback into properly designed statistics that facilitate decision-making while not sacrificing efficiency. Some representative papers studying aspects of reputation formation in settings where player strategies are imperfectly observed by their opponents include Holmstrom (1982), Diamond (1989), Fudenberg and Levine (1992) and Mailath and Samuelson (1998).

2. The Setting

In this paper we offer a comparative analysis of reputation and enforcement-based mechanisms for quality assurance, focusing on the likely impact of information technology on the relative efficiency of these types of institutional mechanisms. We study these two types of mechanisms in a setting with a merchant ("seller") who in each period provides one unit of a product or a service ("good") to one of multiple consumers ("buyers"). The good is either "high quality" or "low quality", but only high quality is acceptable to the buyers. Following receipt of payment, the seller can exert either "high effort" ("cooperate") or "low effort" ("cheat"). The buyer observes the quality of the good delivered, but not the effort exerted by the seller. Moral hazard is introduced because high effort is costlier to the seller, who can reduce his costs by failing to exert high effort, providing the buyer with a good of lower expected quality.

More formally, we analyze a setting with a monopolist seller who each period offers for sale a single unit of a good to *m* buyers. Buyer i has valuation w_i for a high quality good and all buyers value a low quality good at zero. Buyer lifetime is exactly one period and in each period the *m* buyers are drawn from the same probability distribution, thus buyer valuations are independent and identically distributed within and across periods. There is an infinite number of periods and the seller has a period discount factor δ reflecting the frequency of transactions within the community, or the probability that the game will end after each period. Seller effort determines the probability that the good provided will be of low quality: if the seller exerts low effort, the good will be of low quality with probability β, whereas if the seller exerts high effort he will incur an additional cost c and the good will be of low quality with a smaller probability α ($\alpha < \beta$). The seller's objective is to maximize the present value of his payoffs over the entire span of the game, while the buyers' objective is to maximize their short-term (stage game) payoff.

In each period a mechanism is used to allocate the good among the *m* buyers by determining the buyer that receives the good and the price she pays to the seller. Without loss of generality we assume that buyers are indexed according to their valuations ($w_1 \geq w_2 \geq ... \geq w_m$). We further assume that a second price Vickrey auction is used to award the good to the buyer with the highest valuation w_1 for a high quality good. The winning bidder pays a price equal to the second-highest bid G; the valuation of the second-highest bidder for a high quality good is w_2.[1]

While stylized, the above setting captures the essential properties of a large number of important real-life economic settings, ranging from the provision of professional services, to online purchasing and auctions like eBay. In professional services (medical consultations, auditing, construction projects, etc.) there are well defined standards of high quality service and the uncertainty is focused on whether the provider will adhere to those standards or try to "cut

[1] Our results qualitatively apply to any mechanism for determining the buyer that will receive the good in each period and the price paid by that buyer, as long as this mechanism is reasonably efficient in awarding the good to a high valuation buyer and at a price increasing in line with her valuation.

corners". In mail order or online purchasing the moral hazard is focused on whether, following receipt of payment, the seller will provide a good of the quality advertised.

3. Reputation Framework

Modeling a reputation mechanism

In the above setting, we consider a reputation mechanism that allows buyers to rate the seller based on the quality of the good received. Buyers report the outcome of a transaction as either "positive" or "negative", with positive ratings indicating high quality good received, and negative ratings indicating low quality. The mechanism aggregates past ratings and publishes a summary of the seller's most recent ratings. Specifically, buyers can see the total number of each type of rating received by the seller during the most recent N transactions, while earlier ratings are discarded. This mechanism is modeled after eBay's "ID card", which summarizes ratings received during the most recent six-month period (Figure 1). Since we have assumed a binary feedback mechanism (ratings are either positive or negative), a seller's feedback profile can be represented as (x, N), where $x \in \{0,1,...,N\}$ is the number of negative ratings currently contained within that window. At the end of each period, the rating received during the current period is added to the profile whereas the rating received N periods ago is discarded.

Several characteristics of this model capture the role of information technology in online reputation systems: First, once an online system has been developed, the per period cost of collecting, processing, and communicating ratings information is much lower compared to a traditional off-line system. In the setting we consider in this paper, we assume that this per period cost is zero; this assumption is only appropriate for online systems. Second, the type of structured design that we assume for the reputation mechanism in our setting is only feasible in the context of an online system. Third, we assume that consumers provide truthful feedback on the quality of a good received, which hinges on the cost of providing this feedback being low enough so that consumers can be given incentives that induce participation and truth-telling. For instance, this can be accomplished through side-payment mechanisms like the one proposed by (Miller, Resnick and Zeckhauser, 2002). While mechanisms of this type might be infeasible in traditional reputation settings, they can be easily incorporated into online systems. Finally, as information technology makes the outcome of any single transaction immediately known to the entire population of prospective buyers, it increases the proportion of transactions affected by the seller's reputation, which in turn increases the seller's discount factor. As we discuss in Section 5, this affects significantly the ability of the reputation mechanism to promote cooperative behavior, and thus is central to our comparative analysis of reputation and litigation based mechanisms.

Figure 1: Sample eBay feedback profile

Dellarocas (2002b) has shown that the maximum efficiency attainable by such “eBay-like” mechanisms in two-outcome settings with moral hazard is independent of the size N of the time window. In this paper we therefore focus on the special case where $N = 1$: this corresponds to a reputation mechanism that only publishes the *single* most recent rating received for the seller and discards all past ratings. A seller’s reputation profile can then be denoted by a binary state variable $x \in \{0,1\}$ ($x = 0,1$ corresponds to “good” and “bad” reputation respectively). The corresponding stage game is summarized in Figure 2. Even this very simple mechanism allows us to illustrate the ability of reputation mechanisms to induce cooperative behavior. More sophisticated mechanisms are likely to perform even better, especially in more complex settings.

1. Seller offers a single unit of a good, promising to deliver a high quality good (as there is no demand for a low quality good).
2. System provides a binary (positive or negative) rating for the seller, based on the rating received by the buyer in the most recent period. The rating is positive if the buyer in the most recent period received a high quality good, and negative otherwise.
3. Buyers bid their expected valuations for the good in a second price Vickrey auction; the winning bidder pays G, which is the second-highest bid; we denote by w_1 and w_2 the respective valuations for a high quality good of the winning bidder and the second-highest bidder.
4. Seller decides on whether to exert high effort at cost c, or low effort at cost 0, with corresponding probabilities that the resulting good is of low quality being α and β ($\alpha < \beta$).
5. Buyer receives the good, experiences its quality, and realizes the corresponding valuation w_1 for a high quality good or 0 for a low quality good. Buyer reports the quality of the good received to the system, and the rating of the seller reported in the next period is changed accordingly.

Figure 2. Stage game for reputation mechanism.

An important assumption in our model is that all buyers leave truthful feedback to the system. This assumption merits some discussion. Although feedback submission through the Internet incurs drastically lower costs than through more traditional channels, it still does incur a, however small, cost related to connecting to the network and clicking through in order to get to the feedback submission page. Short-term buyers gain nothing from feedback submission; they should therefore be provided with incentives in order to be willing to incur that cost.

A simple bond/side payment mechanism can easily address this issue and induce full participation to the feedback mechanism: the system levies a fee from all prospective buyers before they are allowed to bid in an auction. That fee is refunded to everybody except the winning bidder upon completion of the auction. The winning bidder, on the other hand, only gets back her money after she submits feedback for the seller. Alternatively, the system can incent feedback submission by transferring a fraction of the listing fee it collects from the seller to the winning bidder upon submission of feedback to the system. If the fee/side payment is greater than the cost of feedback submission, it is easy to see how such simple schemes can induce full participation to the feedback mechanism.

Since, in our model, buyers are short-run they gain nothing from strategic manipulation of their feedback reports. Therefore, under the assumption that truthful feedback has equal cost to random feedback (a reasonable assumption if the solicited feedback is as simple as it is on eBay), truth-telling is a weakly dominant action for buyers. Miller, Resnick and Zeckhauser (2002) have recently proposed a more elaborate side-payment mechanism that provides strict incentives for participation and truth-telling even in the presence of strategic incentives to distort one's ratings.

In summary, although the elicitation of complete and truthful feedback is a crucial prerequisite for the efficient functioning of reputation mechanisms, the above discussion plus Miller, Resnick and Zeckhauser's result show that this issue can be effectively addressed with minimal impact on costs and efficiency. Since the objective of this paper is to assess the potential of online reputation rather than the details of any specific mechanism used in practice today, we have abstracted away this issue and assumed complete participation and truthful reporting.

Characterization of Equilibrium Play

Let $s(x,h_t)\in[0,1]$ denote the seller's strategy in period *t*, equal to the probability the seller will cooperate (i.e., exert high effort) in period *t* if his current reputation profile contains $x\in\{0,1\}$ negative ratings at the beginning of the period and the past history of play is h_t. We will restrict the seller to stationary strategies, where $s(x,h_t)$ does not depend on *t*, or the history of play.[2] Let $\mathbf{s}=[s(0),s(1)]$ denote the seller's strategy vector.

Since buyers are short-run, they will always play a best response to the seller's strategy $s(x)$. Furthermore, since they compete with each other on a Vickrey auction, each buyer's optimal bid in each period will be equal to her expected valuation

$$G_i(x,\mathbf{s})=\{s(x)\cdot(1-\alpha)+[1-s(x)]\cdot(1-\beta)\}\cdot w_i=[s(x)\cdot(\beta-\alpha)+(1-\beta)]\cdot w_i \tag{1}$$

resulting in expected auction revenue for that period

$$G(x,\mathbf{s})=[s(x)\cdot(\beta-\alpha)+(1-\beta)]\cdot w_2 \tag{2}$$

---

[2] Dellarocas (2002b) shows that, in a general class of repeated games that includes the current setting, the maximum efficiency achievable through stationary strategies is equal to the maximum efficiency achievable in any sequential equilibrium of the game.

where w_2 is the second highest bidder's valuation of a high quality good. The expected surplus for the winning bidder is

$$V_b(x,\mathbf{s}) = \left[s(x)\cdot(\beta-\alpha)+(1-\beta)\right]\cdot(w_1-w_2) \qquad (3)$$

where w_1 is the winning bidder's valuation of high quality. The seller's corresponding current period payoff is:

$$V_s(x,\mathbf{s}) = G(x,\mathbf{s}) - s(x)\cdot c = [s(x)\cdot(\beta-\alpha)+(1-\beta)]\cdot w_2 - s(x)\cdot c \qquad (4)$$

Since a seller's choice of effort takes place after payment for the current period has been received, the seller's objective is to select $\mathbf{s}$ so as to maximize the present value of his payoff in the remaining game. Let $U(x,\mathbf{s})$ denote the seller's expected *future* payoff, i.e. excluding the payment *G* from the current auction, when the seller's current reputation profile contains x negatives.

If the seller exerts high effort in the current period, he incurs an immediate cost c; the resulting quality of the good will be high with probability $1-\pi$ and low with probability π. Since we have assumed that all buyers leave truthful feedback to the system, if quality is high the reputation profile in the next round will contain $x=0$ negative ratings, otherwise it will contain $x=1$ negative rating. The expected future payoff is

$$U_{coop}(x,\mathbf{s}) = -c + \delta\cdot\left\{(1-\alpha)\cdot[G(0,\mathbf{s})+U(0,\mathbf{s})]+\alpha\cdot[G(1,\mathbf{s})+U(1,\mathbf{s})]\right\} \qquad (5)$$

If the seller exerts low effort, he avoids the cost *c* in the current period. However, the resulting quality of his product will be high with probability $1-\beta$ and low with probability $\beta>\alpha$, which increases the probability that the seller will enter the next period with a "bad" reputation ($x=1$). In this case the expected future payoff is

$$U_{cheat}(x,\mathbf{s}) = \delta\cdot\left\{(1-\beta)\cdot[G(0,\mathbf{s})+U(0,\mathbf{s})]+\beta\cdot[G(1,\mathbf{s})+U(1,\mathbf{s})]\right\} \qquad (6)$$

Note that, because our reputation mechanism discards past ratings, the seller's future payoff is independent of the current state *x* of his reputation profile. Therefore, $U_{coop}(0,\mathbf{s}) = U_{coop}(1,\mathbf{s}) \equiv U_{coop}(\mathbf{s})$ and $U_{cheat}(0,\mathbf{s}) = U_{cheat}(1,\mathbf{s}) \equiv U_{cheat}(\mathbf{s})$.

In the above setting, a strategy **s** is an equilibrium strategy if and only if it satisfies the incentive compatibility constraints:

$$\begin{aligned} &U_{coop}(\mathbf{s}) > U_{cheat}(\mathbf{s}) \Rightarrow s(0) = s(1) = 1 \\ &U_{coop}(\mathbf{s}) = U_{cheat}(\mathbf{s}) \Rightarrow 0 \leq s(0), s(1) \leq 1 \\ &U_{coop}(\mathbf{s}) < U_{cheat}(\mathbf{s}) \Rightarrow s(0) = s(1) = 0 \end{aligned} \tag{7}$$

Not surprisingly, this game has multiple equilibrium strategies (see Proof of Proposition 1 for the full characterization). In the rest of the paper, we will focus our attention on the equilibrium strategy $\mathbf{s}^*$ that maximizes the seller's expected discounted lifetime payoff $W(\mathbf{s}) = E[\sum_{t=0}^{\infty} \delta^t \cdot V_s(x(t), \mathbf{s})] = G(x_0, \mathbf{s}) + U(\mathbf{s})$, where $x_0 \in \{0,1\}$ is the initial state of the reputation profile of new sellers. For discount factors close to one, this strategy also maximizes the average single stage total surplus[3]. This optimal strategy $\mathbf{s}^*$ is the solution of the maximization problem: $W(\mathbf{s}^*) \geq W(\mathbf{s})$ for all $\mathbf{s} \in [0,1]^2$ subject to the incentive compatibility constraints (7). Let $\rho = w_2 / c$; ρ provides a measure for the ratio of the valuation of a high quality good to the cost of high effort and is also a rough measure of the profit margin of a fully cooperating seller. The following proposition summarizes the seller's optimal strategy:

Proposition 1:

(a) If $\rho < \dfrac{1}{\delta \cdot (\beta - \alpha)^2}$ then

I. the seller's optimal strategy is $\mathbf{s}^* = [0,0]$: always exert low effort

II. the expected stage-game auction revenue is equal to $G = (1 - \beta) \cdot w_2$

---

[3] To see this, from (4), if $(\beta - \alpha) \cdot w_2 > c$, a seller's stage game payoff is a linearly increasing function of his probability of cooperation. Therefore, for discount factors close to one the maximum seller discounted lifetime payoff corresponds to higher average levels of cooperation. Since from (3) the winning bidder's surplus is also a linearly increasing function of the seller's probability of cooperation, as the discount factor tends to one, the strategy profile that maximizes the seller's lifetime payoff also maximizes the buyer's one-shot average expected surplus and therefore the average total surplus.

(b) If $\rho \geq \dfrac{1}{\delta \cdot (\beta - \alpha)^2}$ then:

I. the seller's optimal strategy is $\mathbf{s}^* = [1, 1 - \dfrac{1}{\rho\delta \cdot (\beta - \alpha)^2}]$: always exert high effort if the most recent rating was positive and follow a mixed strategy with probability of cooperation $1 - \dfrac{1}{\rho\delta \cdot (\beta - \alpha)^2}$ if the most recent rating was negative;

II. the expected stage-game auction revenue is equal to $G(x) = (1-\alpha)w_2 - x \cdot \dfrac{c}{\delta \cdot (\beta - \alpha)}$.

Proof: See Appendix.

The intuition behind Proposition 1 is the following: From equation (4) it is easy to see that, if $(\beta - \alpha) \cdot w_2 > c$ (a condition that always holds if $\rho \geq 1/\delta \cdot (\beta - \alpha)^2$) a seller's profit from a single transaction is an increasing function of $s(x)$, where $s(x)$ is the probability that the seller will cooperate during periods when his reputation profile has x negatives. From equation (3) we see that buyer surplus also increases with $s(x)$. It is thus to everyone's benefit to cooperate as much as possible. Unfortunately sellers decide whether to cooperate *after* they receive payment and then they always have a short-term gain equal to c if they cheat. Therefore, the only way that a seller will credibly cooperate following receipt of payment is if there is a longer-term loss for him associated with cheating. The only consequence of cheating in this game is a higher probability of transitioning to state $x = 1$ (by receiving a negative rating) and the only way that a seller can have a lower payoff when $x = 1$ is by cooperating less during periods when $x = 1$ (because, expecting this, buyers will then place lower bids during those periods). Therefore, a seller can give himself incentives to cooperate during periods when his reputation is "good" by "promising" to cooperate less during periods when his reputation is "bad" (effectively "punishing himself" by doing so). Proposition 1 shows that, if $\rho \geq 1/\delta \cdot (\beta - \alpha)^2$, cooperating with probability $1 - \dfrac{1}{\rho\delta \cdot (\beta - \alpha)^2}$ during periods when $x = 1$ makes it optimal for a seller to cooperate always during periods when $x = 0$: his remaining game payoffs from cheating and cooperation then become equal and the seller becomes indifferent between these two actions and has no incentive to deviate from the overall strategy $\mathbf{s}^*$. This strategy maximizes both the seller's payoffs during periods of good reputation plus the probability of maintaining his good reputation.

Because of noise, it is inevitable that even cooperating sellers will occasionally produce bad quality products and will then receive negative ratings. During periods when their reputation profile is bad, they will receive lower revenues and their optimal strategy is to randomize between exerting high and low effort. It is remarkable that the way that our reputation mechanism succeeds in inducing full cooperation "most of the time" (when $x = 0$) is by making it optimal for sellers to "cheat a little" (and be penalized for it because buyers expect them to do so) when $x = 1$. This is the main source of inefficiency of this type of mechanism.

The condition $\rho \geq 1/\delta \cdot (\beta - \alpha)^2$ expresses the fact that, in order for our reputation mechanism to succeed in inducing cooperation, the buyers' valuation of high quality must be high enough (relative to cost of exerting high effort) so that discounted future payoffs from sustained cooperation are greater than short-term wealth increases obtained from cheating. This seems to be a general property of reputation mechanisms, first pointed out by Klein and Leffler (1981) and explored more formally by Shapiro (1983). Both Klein and Leffler and Shapiro focused on the implications of this property for profit margins: they concluded that the effectiveness of reputation as a mechanism for inducing cooperation depends on the profit margins of cooperating sellers being sufficiently high so that the promise of future gains is persuasive enough to offset the short-term temptation to cheat.

An alternative way to look at the above condition is by focusing on its implications for the discount factor δ. In order for reputation to be effective in inducing cooperation, it must be $\delta \geq 1\big/\rho(\beta - \alpha)^2$. This, in turn, requires that sellers transact with sufficient frequency within the community that operates the reputation mechanism so that the stream of future payoffs from that community is large enough to offset the short-term benefits of cheating. In Section 5 we expand on this argument and we show how it can be interpreted as requiring a minimum degree of participation before reputation mechanisms become effective in inducing *any* amount of cooperation.

Total Surplus

From (3) and (4) the single stage total surplus $V(x,\mathbf{s}) = V_b(x,\mathbf{s}) + V_s(x,\mathbf{s})$ is equal to:

$$V(x,\mathbf{s}) = \left[s(x)\cdot(\beta-\alpha) + (1-\beta)\right]\cdot w_1 - s(x)\cdot c \qquad (8)$$

The average single stage total surplus is given by:

$$V(\mathbf{s}) = p_0(\mathbf{s})\cdot V(0,\mathbf{s}) + p_1(\mathbf{s})\cdot V(1,\mathbf{s}) \qquad (9)$$

where $p_0(\mathbf{s}), p_1(\mathbf{s})$ are the stationary probabilities that a seller who follows strategy $\mathbf{s}$ will find himself in states $x=0, x=1$ respectively. Proposition 2 shows the total surplus corresponding to the seller strategy of Proposition 1.

Proposition 2:

(a) If $\rho < 1/\delta\cdot(\beta-\alpha)^2$ then the average total surplus per period is $V(\mathbf{s}) = (1-\beta)w_1$;

(b) If $\rho \geq 1/\delta\cdot(\beta-\alpha)^2$ then the average total surplus per period is

$$V(s^*) = [(1-\alpha)w_1 - c] - \left(\frac{\alpha c}{\beta-\alpha}\right)\frac{w_1(\beta-\alpha)-c}{\delta w_2(\beta-\alpha)-c}$$

Proof: See Appendix.

The term $[(1-\alpha)\cdot w_1 - c]$ equals the total surplus when the seller always cooperates, which is the first best outcome if $c < w_1(\beta-\alpha)$. The second term represents the loss in total surplus due to the less than perfect ability of the reputation mechanism to induce cooperation during periods when $x=1$.

4. Enforcement-based Framework ("litigation")

In this section we present a model for a simple litigation mechanism. Instead of reporting on the quality of the good received, the buyer may sue the seller for failing to deliver a high quality good. Since buyers experience the quality of the received product or service directly while courts usually rely on indirect expert testimony, we assume that court decisions are subject to noise. The court will find for the buyer with probability a if the quality of the good received is high, and with probability b if the quality of the good is low ($a < b$). If the court finds for the buyer, the seller must pay the buyer damages D. Whatever the decision of the court, each party incurs litigation costs L. Litigation costs include legal fees, trial fees, the opportunity cost of time spent by each party on this case and the amortized cost of sustaining the enforcement apparatus. The corresponding stage game is summarized in Figure 3 and is shown in extensive form in Figure 4.

1. Seller offers a single unit of a good, promising to deliver a high quality good (as there is no demand for a low quality good).
2. Buyers bid their expected valuations for the good in a second price Vickrey auction; the winning bidder pays G, which is the second-highest bid; the valuations for a high quality good of the winning bidder and the second-highest bidder are w_1 and w_2 respectively.
3. Seller decides on whether to exert high effort at cost c, or low effort at cost 0, with corresponding probabilities that the resulting good is low quality being α and β ($\alpha < \beta$).
4. Buyer receives the good, experiences its quality, and realizes the corresponding valuation w_1 for a high quality good or 0 for a low quality good.
5. Buyer decides whether or not to sue seller. If buyer does not sue, the stage game ends.
6. If the buyer sues, the court finds for the buyer with probability a if the good received was high quality, and with higher probability b if the good received was low quality. Independent of the decision of the court, each party incurs litigation costs L.
7. If the court finds for the buyer, then the seller has to pay to the buyer damages D.

Figure 3. Stage game for litigation mechanism.

As each period is independent, analysis of this game consists of analyzing the stage game. Proposition 3 shows the resulting outcomes of this game:

Proposition 3: In the litigation game,

(a) if $L > bD$, then the buyer will never sue and the seller will always exert low effort.

(b) if $L < aD$, then the buyer will always sue and the seller will exert high effort if and only if

$$D > \frac{c}{(\beta - \alpha)(b - a)}.$$

(c) if $aD < L < bD$, then the buyer will sue if and only if a good of low quality is received and the seller will exert high effort if and only if $c < (\beta - \alpha)(L + bD)$.

Proof: Proposition 3 follows directly from the analysis of cases L1, L2 and L3 in the Appendix.

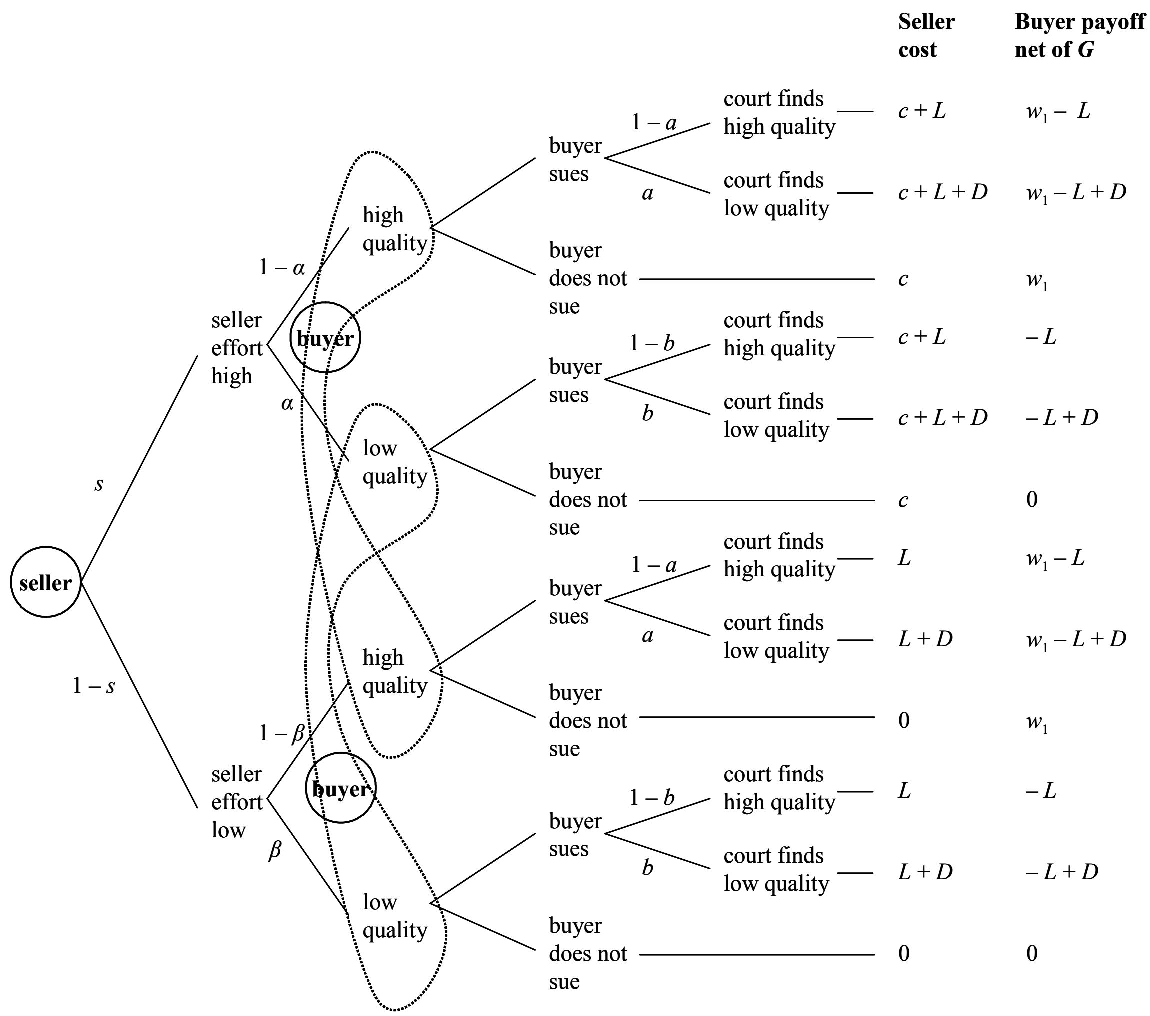

Figure 4. Stage game for litigation mechanism.

Proposition 3 results because in the last move of the stage game, the buyer will sue if the expected payment for damages (aD if the good is high quality and $bD > aD$ if the good is low quality) is higher than the litigation cost L. If $L > bD$, it is never optimal for the buyer to sue, even if a low quality good has been received. In this case, the seller will never exert high effort, as he receives no benefit offsetting the cost of high effort c. If $L < aD$, then the buyer will always sue, even if the quality of the good received is high. If $D > \frac{c}{(\beta-\alpha)(b-a)}$, i.e. if damages are sufficiently high, then the seller will always exert high effort, as its cost is outweighed by the reduction in expected damages, otherwise the seller will never exert high effort. The seller will stay in the market if his total payoff is positive, however this outcome is never optimal because of the reduction in total surplus resulting from the excessive litigation costs. Finally, if $aD < L < bD$, then the buyer will find it optimal to sue only when a low quality good is received. If $c < (\beta-\alpha)(L+bD)$, then the seller will always exert high effort, as its cost is outweighed by the resulting reduction in expected damages and litigation costs. Otherwise, the seller will always exert low effort.

Proposition 4 shows the implications of the above outcomes for maximizing total surplus:

Proposition 4:

(a) If $c < (\beta-\alpha)(L+bD)$ and $L < \frac{(\beta-\alpha)w_1 - c}{2\alpha}$, then social surplus in the litigation game is maximized by setting the level of damages to satisfy $\frac{L}{b} < D < \frac{L}{a}$. This leads to an outcome where the seller always exerts high effort, the buyer sues if and only if she receives a low quality good, and the average total surplus for the stage game is $V = [(1-\alpha)w_1 - c] - 2\alpha L$.

(b) Otherwise, social surplus is maximized by setting damages $D < \frac{L}{b}$, leading to an outcome where the seller always exerts low effort, the buyer never sues and the average total surplus for the stage game is $V = (1-\beta)w_1$.

Proof: Proposition 4 follows directly from the analysis of cases L1, L2 and L3 in the Appendix.

What is driving Proposition 4 is that if the three conditions $aD < L < bD$, $c < (\beta-\alpha)(L+bD)$ and $L < \frac{(\beta-\alpha)w_1 - c}{2\alpha}$ are simultaneously satisfied, then the buyer will sue only when a low quality good is received; the seller will always exert high effort because its cost is less than the expected reduction in legal costs and damages; and inducing the seller to exert high effort through the threat of litigation increases the total surplus. If the level of damages D cannot be chosen so that the above three conditions are simultaneously satisfied, for instance because the cost of litigation is too high, or because the cost of high effort is too high compared to the resulting increase in expected value for consumers, then the threat of litigation cannot efficiently induce the seller to stay in the market and exert high effort. The total surplus will be maximized by avoiding all litigation, even though the seller will always exert low effort as a result. This outcome can be achieved by setting damages so that $D < \frac{L}{b}$. In this case, a high quality good will be produced with probability β, and the expected valuation for consumer i will be βw_i.

5. Discussion

As we mentioned in the Introduction, both reputation and litigation mechanisms have been previously studied in the economics literature. The novelty of our work stems from its comparative focus: ours is the first work we are aware of that studies these two types of institutional mechanisms in the same setting, comparing their economic efficiency in inducing cooperative behavior, and thus allowing us to assess the likely relative impact of information technology.

Impact of Information Technology

Proposition 1 shows that our stylized reputation mechanism can only induce cooperative behavior when $\rho \geq \left[1/\delta(\beta-\alpha)^2\right]$, or, equivalently $\delta \geq \left[1/\rho(\beta-\alpha)^2\right]$.

Before the advent of the Internet, word-of-mouth regarding professionals and merchants took place within relatively small and (almost) mutually disjoint groups of neighbors, friends, co-workers, etc. From a modeling perspective, the situation is equivalent to a setting where each group operates an independent reputation mechanism that only receives and disseminates feedback from members of that group. If a seller operates over a large, fragmented territory, the number of such groups would be very large. The seller's decision problem for each group would then be identical to that captured by the analysis in Section 3, except that the seller's discount factor for future payoffs from any given group would be smaller: if the seller "normally" discounts the future by δ' per period and visits each group every n-th period on average, the appropriate discount factor in considering the seller's behavior for each group will be $\delta = (\delta')^n$. Since $\delta' < 1$, for large enough n it will be $(\delta')^n < \left[1/\rho(\beta-\alpha)^2\right]$. Therefore under the above interpretation of our setting, reputation mechanisms will fail to induce cooperation when feedback networks are sufficiently fragmented.

Internet-based online reputation mechanisms provide easily accessible, low cost focal points for previously disjoint groups to pool their experiences with service providers and merchants into a single feedback repository. As these feedback mechanisms cover more groups, the effect is equivalent to reducing the degree of fragmentation n of the feedback networks. This, in turn, increases the discount factor of the seller when facing participating buyers. At the limit, the spread of these mechanisms means that the outcome of any single transaction becomes immediately known to the entire population of prospective buyers. In our setting, this would result in the seller's discount factor getting closer to one.

Changing the discount factor δ in our analysis is significant, in terms of both the applicability as well as the efficiency of reputation as a mechanism for inducing cooperation. Proposition 1 shows that a minimum discount factor $\delta \geq \dfrac{1}{\rho \cdot (\beta-\alpha)^2}$ is necessary to induce *any* cooperation in our stylized reputation mechanism. Once this threshold is reached, however, significant levels of cooperation can be supported. Furthermore, Proposition 2 shows that as δ increases, the average total surplus achieved by the reputation mechanism increases as well. This is shown in Figure 5, for illustrative values of α, β, c, w and L. Figure 5 also shows that as δ increases, reputation becomes more efficient relative to litigation, since the efficiency of litigation is not affected by δ: while reputation may be less efficient than regulation for lower values of δ (many fragmented reputation networks), it may become more efficient than litigation as δ increases and approaches 1 (single large reputation network).

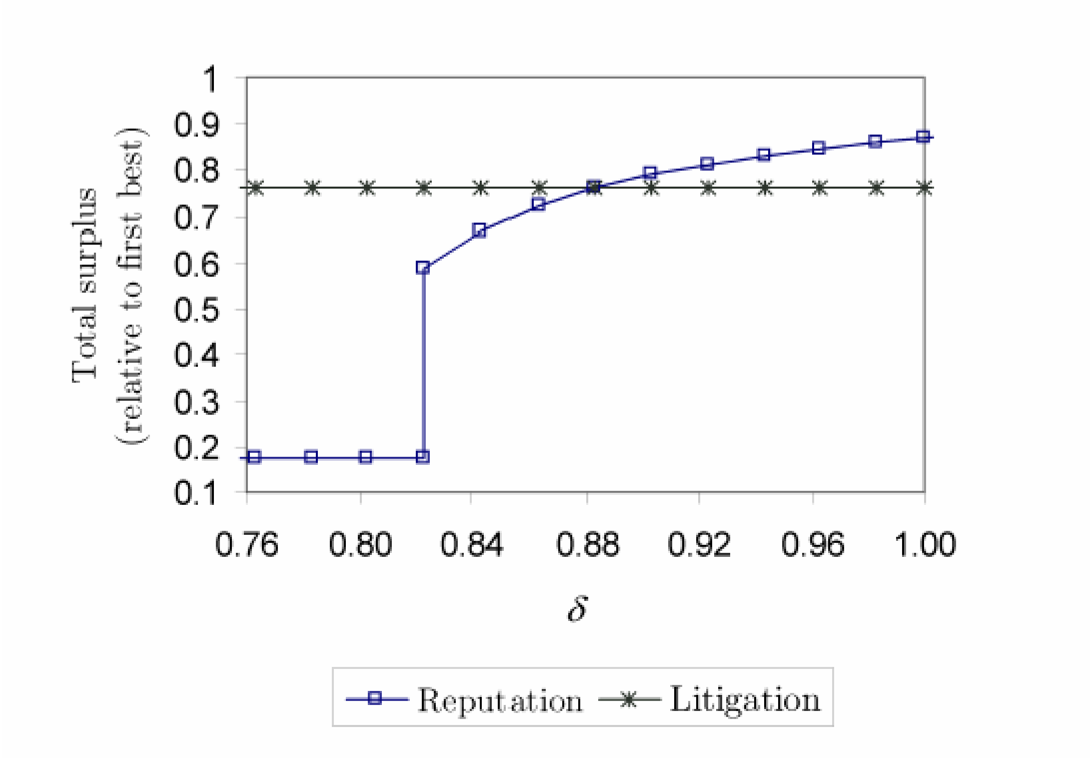

Figure 5. Impact of the discount factor δ on the efficiency of the reputation mechanism
(for $\alpha = 0.05$, $\beta = 0.95$, $c = \$1000$, $w_1 \approx w_2 = \$1500$ and $L = \$1000$)

The above discussion demonstrates how Information Technology enables institutional mechanisms based on online reputation systems to become a feasible alternative to litigation in promoting cooperative behavior in markets. As a result, the design of online reputation mechanisms and the comparative analysis of reputation vs. litigation mechanisms is a promising area of study for IS researchers, in the vein of the earlier research on the institutional implications of IT, such as the Markets vs. Hierarchies stream of research (Malone et al. 1987).

Comparing the efficiency of reputation vs. litigation mechanisms

This section explores the conditions under which reputation mechanisms may achieve higher social surplus than litigation.

The total surplus generated in each period is $(1-\beta)w_1$ when the seller exerts low effort, and $(1-\alpha)w_1 - c$ when the seller exerts high effort. Thus in the first best outcome the seller would exert high effort when $(1-\beta)w_1 < (1-\alpha)w_1 - c$, or $c < (\beta-\alpha)w_1$.

In the absence of a mechanism to induce cooperation, the seller's dominant strategy is to always exert low effort. This is the socially optimal (first-best) outcome when $c > (\beta-\alpha)w_1$, as exerting high effort is too costly compared to the increase in expected valuation, but reduces the total surplus by $(\beta-\alpha)w_1 - c$ when $c < (\beta-\alpha)w_1$. Our analysis shows that both the reputation and the litigation mechanisms under certain conditions can improve on this outcome by inducing the seller to exert high effort most or all of the time.

As shown in Proposition 1, the simple reputation mechanism analyzed in Section 3 induces the seller to exert high effort *most of the time* provided that $c < w_2\delta(\beta-\alpha)^2$, or, equivalently, $\rho > 1/\left[\delta(\beta-\alpha)^2\right]$. Specifically, a seller with "good" reputation ($x=0$) will always cooperate, while a seller with "bad" reputation ($x=1$) will cooperate with probability $1-\dfrac{c/w_2}{\delta\cdot(\beta-\alpha)^2}$.

According to Proposition 2 the resulting average per period total surplus is $V = [(1-\alpha)w_1 - c] - \left(\dfrac{\alpha c}{\beta-\alpha}\right)\dfrac{w_1(\beta-\alpha)-c}{\delta w_2(\beta-\alpha)-c}$, i.e., the reputation mechanism reduces the total surplus by $\left(\dfrac{\alpha c}{\beta-\alpha}\right)\dfrac{w_1(\beta-\alpha)-c}{\delta w_2(\beta-\alpha)-c}$ compared to the high-effort first best outcome.

The efficiency implications of the litigation mechanism we analyzed depend on the litigation costs *L*. If $L < \dfrac{(\beta-\alpha)w_1 - c}{2\alpha}$, then for a properly selected level of damages *D* (i.e., $L/b < D < L/a$), the litigation mechanism will induce the seller to always cooperate. The resulting surplus in this case is $[(1-\alpha)w_1 - c] - 2\alpha L$ (see Proposition 4), i.e. the reputation mechanism reduces total surplus by $2\alpha L$ compared to the high effort first best outcome, a reduction equal to the expected litigation costs of the two parties.

The relative efficiency of the reputation and litigation mechanisms depends on the relative magnitude of these reductions in total surplus. If $\delta \approx 1$ (a reasonable assumption if seller transacts frequently), and if $w_1 \approx w_2$ (a reasonable assumption if the number of buyers in each period is large, so that the valuations of the highest two bidders are approximately equal), then the reduction in surplus for the reputation mechanism simplifies to $\frac{\alpha c}{\beta - \alpha}$. In that case, the reputation mechanism is more efficient than litigation in terms of the total surplus generated if and only if:

$$2\alpha L > \frac{\pi c}{\rho - \pi}, \text{ i.e., } L > \frac{c}{2(\beta - \alpha)} \tag{10}$$

The crucial determinant of the relative efficiency of the two mechanisms is the magnitude of litigation costs L relative to the incremental cost of high effort c. The higher this ratio L/c, the more attractive the use of reputation mechanisms relative to litigation. When $\delta \approx 1$ and $w_1 \approx w_2$, and for most reasonable values of α, β, a and b, reputation will be more efficient than litigation when litigation costs are higher than 50% to 100% of the incremental cost of high effort. Finally, if litigation costs are very high relative to the cost of high effort (specifically if $L > \frac{(1-\alpha)w_1 - c}{2\alpha}$) then, whereas by Proposition 4 the threat of litigation fails to induce any cooperation from sellers, reputation mechanisms succeed to induce cooperation for a high enough ρ.

A numerical example

Consider a setting where the incremental cost of exerting high effort is equal to c=\$1,000, the probability of producing low quality if a seller exerts high effort is α=0.05 and low seller effort always results in low quality, that is, β=1. The frequency of seller transactions is high, and thus $\delta \approx 1$. Finally, the number of buyers is large ($w_1 \approx w_2 \equiv w$). The following points summarize the predictions of our theoretical framework regarding the effectiveness and relative efficiency of reputation and litigation mechanisms in such a setting:

- In order for reputation mechanisms to become effective in inducing *any* amount of cooperation, Proposition 1 requires that $\rho \geq 1.108$. This means that the highest bidder's valuation of high quality must satisfy $w=\rho c \geq \$1,108$. Equivalently, this condition requires that the expected auction revenue for sellers with good reputation must be greater than or equal to $G=(1-\alpha)w=0.95\times\$1,108=\$1,053$, that is, that the profit margin of reputable sellers be at least 5.3%.
- Assuming that the profit margin of reputable sellers satisfies the above condition, equation (10) predicts that reputation mechanisms will outperform litigation in terms of the resulting average total surplus as soon as legal costs rise above $L=c/2(\beta-\alpha)=\$526$.
- Finally, our model predicts that litigation will fail to both induce cooperation and sustain the market once litigation costs rise above $L=[(1-\alpha)w-c]/2\alpha$. The exact threshold depends on the ratio $\rho=w/c$. For example, if $\rho=1.108$ (the minimum ρ for which reputation mechanisms become effective) litigation fails as soon as $L>\$526$. Higher valuations of high quality make markets more tolerant of high legal costs. If $\rho=2$ litigation fails only when $L>\$9,000$. Irrespective of ρ, however, the social efficiency of litigation is surpassed by that of reputation mechanisms as soon as $L>\$526$.

Reputation Mechanisms and Markets

A central function of markets (electronic or otherwise) is the provision of an institutional infrastructure, such as a legal and regulatory framework; this infrastructure is especially important when market participants may behave opportunistically, and without it markets may fail to function efficiently, or break down completely. Consequently, reputation mechanisms may enable the emergence of new markets. For instance, when the three conditions $aD < L < bD$, $c < (\beta-\alpha)(L+bD)$ and $L < \dfrac{(\beta-\alpha)w_1 - c}{2\alpha}$ identified in Proposition 4 cannot be simultaneously satisfied in our setting, e.g., because the cost of litigation is too high compared to the value of a high quality product to consumers, then the litigation mechanism cannot induce the seller to stay in the market and exert high effort.[4]

A reputation mechanism may thus enable a new market to emerge, to the extent that such a mechanism may be able to induce sellers to exert high effort, and the resulting surplus may be sufficient to sustain the market. In other words, under certain conditions a reputation mechanism may succeed in providing trust to a market where a litigation mechanism will fail to do so. It has been previously argued (e.g., Bakos 1997, 1998) that intermediaries like eBay enable new markets to emerge by lowering search costs, when otherwise it would be too costly for potential buyers and sellers to find each other. Our analysis in this article shows that in the case of eBay, the provision of a reputation mechanism may play an equally important role in enabling the emergence of new markets.

The role of reputation mechanisms is likely to be particularly important in markets for professional services, such as legal, medical, accounting, home improvement, etc. In these cases legal costs are likely to be high compared to the cost of high effort, it may be costly for a court to verify the quality of the service provided, and the outcome of the court's evaluation may be noisy. All of these factors favor the relative attractiveness of reputation mechanisms for providing trust in these markets. This is particularly significant in view of predictions that information technology will increase the role of markets for professional services. For instance, Malone and Laubacher (1998) have argued that we are moving towards an "e-Lance economy" with professional services auctioned off on an ad-hoc basis. Our analysis suggests that reputation mechanisms would play a central role in enabling this type of markets.

[4] Proposition 4(b) shows that in such cases the best policy is to set damages to low levels so that litigation is avoided even though sellers always exert low effort. The resulting surplus of $(1-\beta)w_1$ may not be adequate to sustain the market, especially if β is high. For example, buyers and sellers may incur certain transaction and search costs in order to participate in the market, or sellers may incur a certain cost even if they exert low effort, in which case the market will break down.

6. Concluding Remarks

Recent advances in information technology are causing us to rethink many institutions that shape relationships in our everyday life. One important area where information technology can have a profound impact are the institutions that promote trust and cooperation among economic agents. The emergence of online communities has enabled the creation of low cost reputation networks of global reach. On the other hand, technology is having only a moderate impact on the costs of traditional mechanisms that depend on contract enforcement through litigation. As a result, online reputation mechanisms are likely to emerge as the preferred institutions to promote cooperation among economic agents in a large number of settings, augmenting or substituting for traditional litigation-based mechanisms, or enabling a more efficient outcome in markets where cooperative behavior was heretofore unsustainable.

The comparative analysis in this paper was based on a rudimentary binary reputation mechanism. Future research should explore more sophisticated mechanisms (e.g., with reputation profiles based on multi-valued ratings that can differentiate among multiple different qualities); such mechanisms may perform better, and thus will strengthen our results. Similarly, more sophisticated litigation models can be used in the comparison, for example ones that allow for a settlement offer before resorting to the court. Furthermore, we ignored the fixed costs of setting up the legal system and the fixed and variable cost of setting up and running the reputation mechanism. Since the variable cost of online reputation mechanisms is close to zero, it should not significantly affect the outcomes we derived. As the fixed costs of the legal system are sunk, the efficiency improvement introduced by a reputation mechanism provides an estimate of the maximum socially desirable investment in developing such a mechanism. Finally, our analysis can be extended to more general settings where the price and allocation of the good are determined by mechanisms other than a per-period auction.

References

Baird, D.G., Gertner, R.H. and Picker, R.C. (1995) *Game Theory and the Law*. Harvard University Press.

Bajari, P. and A. Hortascu (2000). Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights From eBay Auctions. Working paper, December 2000.

Bakos, Y. (1997) Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science*, 43, 12.

Bakos, Y (1998). Towards Friction-Free Markets: The Emerging Role of Electronic Marketplaces on the Internet, *Communications of the ACM*, Volume 41, Number 8, pp. 35-42.

Bebchuk, L. (1984) Litigation and Settlement under Imperfect Information. *Rand Journal of Economics* 15, pp. 404-415.

Benson, Bruce (1989). The Spontaneous Evolution of Commercial Law. *Southern Economic Journal* 55, Jan.: 644-61.

Cooter, R.D. and Rubinfeld, D. L. (1989) Economic Analysis of Legal Disputes and Their Resolution. *Journal of Economic Studies*, 27 (3), pp. 1067-1097.

Dellarocas, C. (2002a) The Digitization of Word of Mouth: Promise and Challenges of Online Reputation Systems. Sloan School of Management, Massachusetts Institute of Technology, Working Paper Series.

Dellarocas, C. (2002b) Efficiency and Robustness of Mediated Online Feedback Mechanisms: The Case of eBay. Working Paper.

Dewan, S. and Hsu, V. (2001) Trust in Electronic Markets: Price Discovery in Generalist Versus Specialty Online Auctions. Working Paper. January 31, 2001.

Diamond, D. (1989). Reputation Acquisition in Debt Markets. *Journal of Political Economy* 97, 4, pp. 828-862.

Fudenberg, D. and Levine, D.K. (1992) Maintaining a Reputation when Strategies are Imperfectly Observed. *Review of Economic Studies* 59, pp. 561-579.

Holmstrom, B. (1982). Managerial Incentive Problems: A Dynamic Perspective. In *Essays in Economics and Management in Honour of Lars Wahlbeck*. Helsinki: Swedish School of Economics, pp. 209-230.

Houser, D. and Wooders, J. (2000) Reputation in Auctions: Theory and Evidence from eBay. Working Paper, University of Arizona, 2000.

Klein, B. and Leffler, K. (1981) The Role of Market Forces in Assuring Contractual Performance. *Journal of Political Economy*. 89, 4, pp. 615-641.

Lucking-Reiley, D., Bryan, D., Prasad, N. and Reeves, D. (2000). Pennies from eBay: The Determinants of Price in Online Auctions, Working Paper, Vanderbilt University, 2000.

Mailath, G. J. and Samuelson, L. (1998) Your Reputation Is Who You're Not, Not Who You'd Like to Be. CARESS Working Paper 98-11.

Malone, T. W., Yates, J., and Benjamin, R. (1987) Electronic Markets and Electronic Hierarchies. *Communications of the ACM*, 30(6), pp. 484-497.

Malone, T. W. and Laubacher, Robert J. (1998) The Dawn of the E-Lance Economy. *Harvard Business Review*, September-October 1998, pp. 145-152.

Milgrom, P. R., D. C. North, and B. R. Weingast. (1990). The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs. *Economics and Politics* 2: 1-23.

Miller, N., Resnick, P. and Zeckhauser, R. (2002) Eliciting Honest Feedback in Electronic Markets. Woking Paper.

Nalebuff, B. (1987). Credible Pretrial Negotiation. *Rand Journal of Economics* 18, pp. 198-210.

P'ng, I. (1983) Strategic Behavior in Suit, Settlement and Trial. *Bell Journal of Economics* 14, pp. 539-550.

Reinganum, J. and Wilde, L. (1986) Settlement, Litigation and the Allocation of Litigation Costs. *Rand Journal of Economics* 17, pp. 557-566.

Resnick, P., Zeckhauser, R., Friedman, E., Kuwabara, K. (2000) Reputation Systems. *Communications of the ACM*, Vol. 43, (12), December 2000, pp. 45-48.

Resnick, P. and Zeckhauser, R. (2001) Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. Working Paper for the NBER workshop on empirical studies of electronic commerce. January 2001.

Shapiro, C. (1983). Premiums for High Quality Products as Returns to Reputations. *The Quarterly Journal of Economics*, November 1983, pp. 659-679.

Wilson, Robert (1985). Reputations in Games and Markets. In *Game-Theoretic Models of Bargaining*, edited by Alvin Roth, Cambridge University Press, pp. 27-62.

A. Appendix

A.1. Proof of Proposition 1

Because the seller decides his level of effort for the current period *after* receiving payment and because past ratings are discarded, the seller's decision problem is independent of the current state x of his reputation profile. Therefore we can write $U(0,\mathbf{s}) = U(1,\mathbf{s}) \equiv U(\mathbf{s})$, $U_{coop}(0,\mathbf{s}) = U_{coop}(1,\mathbf{s}) \equiv U_{coop}(\mathbf{s})$ and $U_{cheat}(0,\mathbf{s}) = U_{cheat}(1,\mathbf{s}) \equiv U_{cheat}(\mathbf{s})$. By substituting the above, plus the expressions for $G(x,\mathbf{s})$ from (2) into (4) and (5) we get:

$$U_{coop}(\mathbf{s}) = -c + \delta \cdot \{[((1-\alpha)\cdot s(0) + \alpha \cdot s(1))\cdot(\beta-\alpha) + (1-\beta)]\cdot w_2 + U(\mathbf{s})\} \tag{A.1}$$

$$U_{cheat}(\mathbf{s}) = \delta \cdot \{[((1-\beta)\cdot s(0) + \beta \cdot s(1))\cdot(\beta-\alpha) + (1-\beta)]\cdot w_2 + U(\mathbf{s})\} \tag{A.2}$$

There are three possible cases:

i) $U_{coop}(\mathbf{s}) > U_{cheat}(\mathbf{s})$. In this case, according to the incentive compatibility constraints (7) the seller would always find it preferable to cooperate. Therefore $\mathbf{s} = [1,1]$. By substituting $\mathbf{s}$ into (A.1) and (A.2) we get:

$$U_{coop}(\mathbf{s}) = -c + \delta \cdot \{(1-\alpha)\cdot w_2 + U(\mathbf{s})\} \tag{A.3}$$

$$U_{cheat}(\mathbf{s}) = \delta \cdot \{(1-\alpha)\cdot w_2 + U(\mathbf{s})\} \tag{A.4}$$

Thus, $U_{coop}(\mathbf{s}) < U_{cheat}(\mathbf{s})$, which contradicts the original assumption. Therefore, the strategy $\mathbf{s} = [1,1]$ is not an equilibrium.

ii) $U_{coop}(\mathbf{s}) < U_{cheat}(\mathbf{s})$. In this case, the seller will always prefer to cheat and thus $\mathbf{s} = [0,0]$. By substituting $\mathbf{s}$ into (A.1) and (A.2) we get:

$$U_{coop}(\mathbf{s}) = -c + \delta \cdot \{(1-\beta)\cdot w_2 + U(\mathbf{s})\} \tag{A.5}$$

$$U_{cheat}(\mathbf{s}) = \delta \cdot \{(1-\beta)\cdot w_2 + U(\mathbf{s})\} \tag{A.6}$$

It is easy to see that indeed $U_{coop}(\mathbf{s}) < U_{cheat}(\mathbf{s})$. Therefore the strategy $\mathbf{s_0} = [0,0]$ is an equilibrium of this game. In this equilibrium $U(\mathbf{s_0}) \equiv U_{cheat}(\mathbf{s_0})$, which, by substitution into (A.6) gives $U(\mathbf{s_0}) = \dfrac{\delta \cdot (1-\beta)\cdot w_2}{1-\delta}$. From (2) the expected auction revenue then becomes $G(\mathbf{s_0}) = (1-\beta)\cdot w_2$ irrespective of the current profile state. Finally, the seller's discounted lifetime payoff corresponding to strategy $\mathbf{s_0}$ is equal to:

$$W(\mathbf{s_0}) = G(\mathbf{s_0}) + U(\mathbf{s_0}) = \frac{(1-\beta)\cdot w_2}{1-\delta} \tag{A.7}$$

iii) $U_{coop}(\mathbf{s}) = U_{cheat}(\mathbf{s})$. In this case the seller would be indifferent between cheating and cooperation. From (A.1) and (A.2), equality of the payoffs implies:

$$s(0) - s(1) = \frac{c / w_2}{\delta \cdot (\beta-\alpha)^2} \tag{A.8}$$

or, equivalently, $\mathbf{s} = \mathbf{s}(s(0)) = [s(0), s(0) - \frac{c / w_2}{\delta \cdot (\beta - \alpha)^2}]$. Therefore, any strategy of this form where $\frac{c / w_2}{\delta \cdot (\beta - \alpha)^2} \le s(0) \le 1$ is an equilibrium strategy. Such mixed equilibria will exist only if $\frac{c / w_2}{\delta \cdot (\beta - \alpha)^2} \le 1$. Let $\rho = w_2 / c$. Then the above mixed equilibrium existence inequality becomes

$$\rho \ge \frac{1}{\delta \cdot (\beta - \alpha)^2} \tag{A.9}$$

The corresponding payoff function is given by substitution of $\mathbf{s}(s(0))$ into either (A.1) or (A.2):

$$U(\mathbf{s}(s(0))) = \frac{\delta \cdot [(\beta - \alpha) \cdot s(0) + (1 - \beta)] \cdot w_2}{1 - \delta} - \frac{\beta \cdot c}{(1 - \delta) \cdot (\beta - \alpha)} \tag{A.10}$$

Equation (A.10) is linear in $s(0)$ and is maximized for $s(0) = 1$. Therefore, the mixed equilibrium strategy that maximizes the seller's payoff is $\mathbf{s}_1 = \mathbf{s}(1) = [1, 1 - \frac{c / w_2}{\delta \cdot (\beta - \alpha)^2}]$. The seller's remaining payoff then becomes $U(\mathbf{s}_1) = \frac{\delta \cdot (1 - \alpha) \cdot w_2}{1 - \delta} - \frac{\beta \cdot c}{(1 - \delta) \cdot (\beta - \alpha)}$. From (2) the expected auction revenue becomes a function of the seller's current profile state and equal to $G(x, \mathbf{s}_1) = (1 - \alpha) w_2 - x \cdot \frac{c}{\delta \cdot (\beta - \alpha)}$. Finally, the seller's discounted lifetime payoff corresponding to strategy $\mathbf{s}_1$ is maximized when new sellers begin the game with a "good" reputation ($x_0 = 0$). Their lifetime discounted payoff then becomes equal to:

$$W(\mathbf{s}_1) = G(0, \mathbf{s}_1) + U(\mathbf{s}_1) = \frac{\delta \cdot (1 - \alpha) \cdot w_2}{1 - \delta} - \frac{\beta \cdot c}{(1 - \delta) \cdot (\beta - \alpha)} \tag{A.11}$$

By comparing (A.7) and (A.11), we see that $W(\mathbf{s}_1) > W(\mathbf{s_0}) \Leftrightarrow \rho > \frac{\beta}{\delta \cdot (\beta - \alpha)^2}$. Furthermore, in order for strategy $\mathbf{s}_1$ to be an equilibrium strategy, (A.9) must hold. However, since $\frac{\beta}{\delta \cdot (\beta - \alpha)^2} < \frac{1}{\delta \cdot (\beta - \alpha)^2}$, if (A.9) is satisfied, then strategy $\mathbf{s}_1$ results in higher payoff relative to strategy $\mathbf{s_0}$ (therefore $\mathbf{s}^* = \mathbf{s_1}$), whereas if (A.9) is not satisfied then strategy $\mathbf{s_0}$ is the only equilibrium strategy (therefore trivially $\mathbf{s}^* = \mathbf{s_0}$). This completes the proof.

A.2. Proof of Proposition 2

The proof of case (a) is trivial. The proof of case (b) follows.

The reputation game described in Section 2 can be viewed as a Markov process with state x and transition probabilities $\tau_{ij}(\mathbf{s}) = \Pr[x_{t+1} = j \mid x_t = i, \mathbf{s}]$ given by:

$$\tau_{ij}(\mathbf{s}) = \Pr[x_{t+1} = j \mid x_t = i] = s(i) \cdot \Pr[x_{t+1} = j \mid cooperate] + (1 - s(i)) \cdot \Pr[x_{t+1} = j \mid cheat] \tag{A.13}$$

where

$\Pr[x_{t+1} = 0 \mid cooperate] = 1 - \alpha$ $\qquad\qquad$ $\Pr[x_{t+1} = 0 \mid cheat] = 1 - \beta$

$\Pr[x_{t+1} = 1 \mid cooperate] = \alpha$ $\qquad\qquad$ $\Pr[x_{t+1} = 1 \mid cheat] = \beta$

Since we have assumed that $\rho \geq \frac{1}{\delta \cdot (\beta - \alpha)^2}$, by Proposition 1 and Footnote 3 the seller strategy that results in maximum average stage game total surplus is $\mathbf{s}_1 = [1, 1 - \frac{c / w_2}{\delta \cdot (\beta - \alpha)^2}]$. By substituting $\mathbf{s}_1$ into (A.13) we get the transition probability matrix:

$$\boldsymbol{\tau}(\mathbf{s}_1) = \begin{bmatrix} \tau_{00}(\mathbf{s}_1) & \tau_{01}(\mathbf{s}_1) \\ \tau_{10}(\mathbf{s}_1) & \tau_{11}(\mathbf{s}_1) \end{bmatrix} = \begin{bmatrix} 1-\alpha & \alpha \\ 1 - \alpha - \frac{c}{\delta \cdot (\beta - \alpha) \cdot w_2} & \pi + \frac{c}{\delta \cdot (\beta - \alpha) \cdot w_2} \end{bmatrix} \tag{A.14}$$

It is known from the theory of Markov processes that the stationary probabilities $[p_0(\mathbf{s}), p_1(\mathbf{s})]$ are equal to the normalized eigenvector corresponding to the unit eigenvalue of matrix $\boldsymbol{\tau}(\mathbf{s})$. After some algebraic manipulation we get:

$$[p_0(\mathbf{s}_1), p_1(\mathbf{s}_1)] = \left[\frac{\delta \cdot w_2 \cdot (\beta - \alpha) \cdot (1 - \alpha) - c}{\delta \cdot w_2 \cdot (\beta - \alpha) - c}, \frac{\delta \cdot w_2 \cdot (\beta - \alpha) \cdot \alpha}{\delta \cdot w_2 \cdot (\beta - \alpha) - c} \right] \tag{A.15}$$

If we now substitute $p_0(\mathbf{s}_1), p_1(\mathbf{s}_1)$ into (9) we get the final expression for the average total surplus.

A.3. Analysis of the litigation game

<u>Case L1</u>: $L > bD$ and the buyer never sues.

In this case the dominant strategy for the seller is to choose $s = 0$, i.e, the seller will never exert high effort. The buyer will realize a payoff $(1 - \beta) w_1 - G$, while the seller's payoff is G. The total surplus is $(1 - \beta) w_1$. Both the buyer and the seller will participate in the game, as both payoffs will be positive.

<u>Case L2</u>: $L < aD$ and the buyer always sues.

In this case the seller will choose his optimal strategy s^* to minimize its total cost K, where

$$\begin{aligned} K = {} & s[(1-\alpha)(1-\beta) + \pi(1-b)](c+L) + s[(1-\alpha)a + \alpha b](c+L+D)s + \\ & (1-s)[(1-\beta)(1-a) + \beta(1-b)]L + (1-s)[(1-\beta)a + \beta b](L+D) \end{aligned} \tag{A.16}$$

Differentiating (A.16) we obtain $\frac{\partial K}{\partial s} = c - (\beta - \alpha)(b - a) D$.

The optimal strategy for the seller is $s^* = 0$ if $\frac{\partial K}{\partial s} > 0$ and $s^* = 1$ if $\frac{\partial K}{\partial s} < 0$. We thus distinguish the following two subcases:

Case L2a

$\frac{\partial K}{\partial s} > 0$ implies $c - (\beta - \alpha)(b - a)D > 0$, or $D < \frac{c}{(\beta - \alpha)(b - a)}$ (A.17)

If (A.17) holds, then $s^* = 0$, and the seller always exerts low effort. The correspondng buyer and seller payoffs are $(1-\beta)w_1 - L + [\beta b + (1-\beta)a]D - w_1$ and $w_1 - L - [\beta b + (1-\beta)a]D$, and their total surplus is $(1-\beta)w_1 - 2L$; this surplus must be nonnegative for the buyer and the seller to participate in the market.

It is easy to see that L2a is dominated by case L1, where the seller also exerts low effort, but the litigation costs are avoided because the buyer never sues.

Case L2b

If $\frac{\partial K}{\partial s} < 0$, i.e., $D > \frac{c}{(\beta - \alpha)(b - a)}$, then $s^* = 1$, and the seller always exerts high effort. The buyer realizes payoff $(1-\alpha)w_1 - L + [\alpha b + (1-\alpha)a]D - G$, the seller realizes payoff $G - L - [\alpha b + (1-\alpha)a]D$, and the total surplus is $(1-\alpha)w_1 - 2L - c$. This outcome is dominated by case L3b below, where the buyer sues only if quality is low.

Cases L2a and L2b show that setting $D > L/a$ leads to inefficient outcomes due to excessive litigation costs.

Case L3: $aD < L < bD$; the buyer will sue only if quality is low.

In this case the seller faces cost $K = s(1-\alpha)c + s\alpha(c + L + bD) + (1-s)\beta(L + bD)$. (A.18)

Thus $\frac{\partial K}{\partial s} = (1-\alpha)c + \alpha(c + L + bD) - \beta(L + bD) = c - (\beta - \alpha)(L + bD)$. We distinguish two cases:

Case L3a: If $c > (\beta - \alpha)(L + bD)$, then $\frac{\partial K}{\partial s} > 0$, and seller will minimize his cost by setting s*=0.

This results in seller payoff $G - \beta(L + bD)$, buyer payoff $(1-\beta)w_1 + \beta bD - \beta L - G$, and total surplus $(1-\beta)w_1 - 2\beta L$. Not surprisingly, this outcome is dominated by case L1, where the effort of the seller is also low, but litigation costs are avoided.

Case L3b: If $c < (\beta - \alpha)(L + bD)$, then $\frac{\partial K}{\partial s} < 0$, and the seller will minimize his cost by setting s*=1.

This results in buyer payoff $(1-\alpha)w_1 - G - \alpha L + \alpha bD$, seller payoff $G - \alpha L - \alpha bD - c$, and total surplus $(1-\alpha)w_1 - 2\alpha L - c$.

The participation constraints are $G - \alpha L - \alpha bD - c > 0$ for the seller, and $(1-\alpha)w_1 - G - \alpha L + \alpha bD > 0$ for the buyer.

Comparing the total surplus in cases L1 and L3b, case L3b will result in a higher surplus if $L < \frac{(\beta - \alpha)w_1 - c}{2\alpha}$. In other words, if $L < \frac{(\beta - \alpha)w_1 - c}{2\alpha}$ then D should be chosen to satisfy $\frac{L}{b} < D < \frac{L}{a}$. Otherwise, the damages should satisfy $D < \frac{L}{b}$, leading to an equilibrium where the seller never exerts high effort and there is no litigation.