

# Efficiency through Feedback-contingent Fees and Rewards in Auction Marketplaces with Adverse Selection and Moral Hazard

Chrysanthos Dellarocas  
Sloan School of Management  
Massachusetts Institute of Technology  
Cambridge, MA 02139, USA  
dell@mit.edu

## ABSTRACT

This paper proposes a novel mechanism for inducing cooperation in online auction settings with noisy monitoring of quality and adverse selection. The mechanism combines the ability of electronic markets to solicit feedback from buyers with the more traditional ability to levy listing fees from sellers. Each period the mechanism charges a listing fee contingent on a seller's announced expected quality. It subsequently pays the seller a reward contingent on both his announced quality and the rating posted for that seller by that period's winning bidder. I show that, in the presence of a continuum of seller types with different cost functions, imperfect private monitoring of a seller's effort level and a simple "binary" feedback mechanism that asks buyers to rate a transaction as "good" or "bad", it is possible to derive a schedule of fees and rewards that induces all seller types to produce at their respective first-best quality levels and to truthfully announce their intended quality levels to buyers. The mechanism maximizes average social welfare for the entire community and is robust to a number of contingencies of particular concern in online environments, such as easy name changes and the existence of inept sellers. On the other hand, the mechanism distorts the resulting payoffs of *individual* sellers relative to the complete information case, transferring part of the payoffs of more efficient sellers to less efficient sellers. The magnitude of this distortion is proportional to the amount of noise associated with observing and reporting the quality of a good.

## Categories and Subject Descriptors

K.4.4 [Computers and Society]: Electronic Commerce

## General Terms

Economics

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

EC'03, June 9–12, 2003, San Diego, California, USA.

Copyright 2003 ACM 1-58113-679-X/03/0006 ...\$5.00.

## Keywords

online auctions, trust, listing fees, feedback, moral hazard, adverse selection, contract theory

## 1. INTRODUCTION

Online feedback mechanisms (Resnick, et. al., 2000; Dellarocas, 2003) are emerging as a promising approach for building trust and inducing cooperation in online trading environments where more established methods of social control (such as state regulation or the threat of litigation) are often difficult or too costly to implement.

Most current online feedback mechanisms act as "passive" repositories of feedback information. For example, eBay's well-known "feedback forum" encourages buyers and sellers to rate one another by designating the outcome of a transaction as "positive", "negative" or "neutral". eBay then aggregates past ratings and makes the history of a member's past ratings as well as summary statistics of his recent ratings available to the entire community. eBay does not directly penalize sellers for accumulating bad ratings<sup>1</sup>. It is up to the buyers to decide how to factor that information into their subsequent interaction with sellers.

Since the seminal papers of Kreps, Milgrom, Roberts and Wilson (1982), Kreps and Wilson (1982), and Milgrom and Roberts (1982) it is well known that, in repeated bilateral exchange settings where a long-run player (seller) faces a sequence of short-run opponents (buyers), making the history of past stage game outcomes (or imperfect signals thereof) available to subsequent buyers provides incentives to the seller to cooperate with buyers in order to build a reputation for being "honest". A striking recent result, however, shows that, in environments with noisy observations of outcomes, such a setup makes it *impossible* to sustain long-term reputations (Cripps, Mailath and Samuelson, 2002). The basic intuition behind this result is that, once a seller has established a reputation for producing good quality, in environments with noisy observations of quality he will then be

<sup>1</sup>eBay will expel members whose feedback score (total number of positive minus the total number of negative ratings received from unique members) falls below a negative threshold. However, given the current setup of eBay's mechanism, such occurrences are exceedingly rare (Resnick and Zeckhauser, 2001). Furthermore, it is relatively easy for expelled members to re-enter the community under a new online identity.

tempted to occasionally cheat because buyers will assume that occasional bad ratings are simply due to noise. Cripps, Mailath and Samuelson then prove that, in the long run, such occasional deviations eventually reveal the “strategic” nature of the seller and the game then reverts to the inefficient Nash equilibrium where the seller always cheats and, expecting that, buyers don’t buy or pay very low prices.

A corollary of this result is that simple reputation repositories that publish histories of past ratings cannot sustain long-run social efficiency. The result motivates the need for further research in the design of feedback mechanisms that can sustain more stable cooperation in noisy environments.

This paper proposes a novel mechanism for inducing cooperation in online auction settings. It considers not only the standard moral hazard problem but also allows for adverse selection, that is, for the (realistic) possibility of heterogeneous sellers with different cost functions that are privately known to them. The mechanism combines the ability of electronic markets to solicit feedback from buyers with the more traditional ability to levy listing or participation fees from sellers. Each period the mechanism charges a listing fee contingent on a seller’s announced expected quality. It subsequently pays the seller a reward contingent on both his announced quality and the rating posted for that seller by that period’s winning bidder. I show that, in the presence of a continuum of seller types with different cost functions, imperfect private monitoring of a seller’s effort level and a simple “binary” feedback mechanism that asks buyers to rate a transaction as “good” or “bad”, it is possible to derive a schedule of fees and rewards that induces all seller types to produce at their respective first-best quality levels (the quality levels that maximize their respective profits under complete information) and to truthfully announce their intended quality levels to buyers.

The flavor of the mechanism relates to the literature on procurement (and, more generally, optimal contracting) with adverse selection and moral hazard (Laffont and Tirole, 1993). One important difference between that literature and the setting of this paper is that, in the former case, the procurement agency not only designs the mechanism but is also the buyer. It thus aims to set the terms of a contract so as to leave no rent to the seller. As a consequence, if the agency is unsure about the cost function of the seller it faces, it turns out that optimal contracting leads to underperformance: all sellers (except the most efficient ones) are induced to produce qualities that are inferior to their first-best quality levels under complete information.

In the current setting the mechanism designer and the buyers are different parties. Furthermore, the auction mechanism leaves all seller types with positive rents. If these rents are sufficiently high, the center can design an individually rational schedule of listing fees and rewards that (a) induce both truth-telling and first-best effort levels by all seller types and (b) maximize average social welfare for the entire community. Even in this case, however, adverse selection does result in an (inevitable) distortion. The optimal schedule of fees and rewards distorts the resulting payoffs of *individual* sellers relative to the complete information case, essentially transferring part of the fees paid to the center by more efficient sellers to less efficient sellers. Less efficient sellers therefore end up better off while more efficient sellers end up worse off (relative to the complete information case). The magnitude of this payoff distortion increases with the

amount of noise associated with observing and reporting the quality of a good.

From a practical perspective, the mechanism readily extends to repeated game settings and is easy to implement in environments such as eBay. In contrast to more traditional feedback mechanisms in noisy environments, it induces full cooperation that remains stable over time. Furthermore, it is robust to a number of contingencies of particular concern in online environments, such as easy name changes (that tempt sellers to cheat, disappear and re-enter a community under a new online identity) and inept sellers (sellers who cannot produce at any quality and, hidden behind the relative anonymity of online environments, enter the market for the explicit purpose of cheating). Finally, this mechanism demonstrates that simple binary ratings (i.e. a simple designation of a transaction as “good” or “bad”) are sufficient for sustaining full cooperation and efficient effort levels in environments of multiple qualities.

The rest of the paper is organized as follows. Section 2 presents the setting. Section 3 describes the mechanism and shows how the optimal schedule of fees can be calculated. Section 4 illustrates the properties of the mechanism through a concrete example. Section 5 discusses the mechanism’s robustness and practical implementation. Finally, Section 6 concludes and discusses possibilities for future research.

## 2. THE SETTING

The setting involves a monopolist seller who offers for sale a single unit of a product or service (“good”) to one of multiple identical buyers. The good’s quality can range from  $q$  to  $\bar{q}$ , where  $\underline{q} < \bar{q}$ . Buyers value goods of quality  $q \in [\underline{q}, \bar{q}]$  at  $w(q)$ , where  $w'(q) > 0$  and  $w''(q) < 0^2$ . The seller is characterized by an efficiency parameter (“type”)  $\theta \in [\underline{\theta}, \bar{\theta}]$ . The seller’s type is unknown to buyers and determines the seller’s per period cost function  $c(\theta, z)$  where  $z$  indicates the seller’s level of effort. I assume that higher efficiency implies lower cost for the same level of effort, i.e.  $c_1(\theta, z) < 0$ . Furthermore, for all  $\theta$ ,  $c_2(\theta, z) > 0$  and  $c_{22}(\theta, z) > 0$ .

At the beginning of the game the seller announces the expected quality  $y$  of his product. A mechanism is then used to allocate the good among the buyers by determining the buyer that receives the good and the price she pays to the seller. We assume that a second price Vickrey auction is used to award the good. In a setting where buyer valuations are identical this results in a winning bid  $G(y)$  equal to the buyers’ expected valuation of the promised good, given the seller’s announcement.

Following receipt of payment, the seller exerts effort  $z$  at cost  $c(\theta, z)$ . Seller effort stochastically determines the quality of the resulting good. More specifically, if a seller of type  $\theta$  exerts effort  $z$ , the quality  $q$  of the resulting good is characterized by a cumulative probability distribution  $\Phi(\theta, \kappa, z) = \Pr[q \leq \kappa | \theta, z]$  with mean  $\mu(\theta, z)$ , where  $\mu_2(\theta, z) > 0$  and  $\mu_{22}(\theta, z) < 0$ . With a slight abuse of notation, in the rest of the paper I will therefore rewrite the cost function as a function of the expected quality  $q$  associated with the seller’s effort  $c(\theta, q) \equiv c(\theta, \mu^{-1}(\theta, q))$  where  $c_1(\theta, q) < 0$ ,

<sup>2</sup>In this paper I use the notation  $f'(\cdot), f''(\cdot), \dots$  to designate derivatives of functions of one variable and  $f_1(\cdot, \cdot), f_2(\cdot, \cdot), f_{11}(\cdot, \cdot), f_{12}(\cdot, \cdot), \dots$  to designate partial derivatives of functions of two variables.

$c_2(\theta, q) > 0$  and  $c_{22}(\theta, q) > 0$ . Likewise, I will rewrite the probability distribution of resulting quality as  $\Phi(\theta, \kappa, q) \equiv \Phi(\theta, \kappa, \mu^{-1}(\theta, q))$ . Finally, I will *assume* that, for all types  $\theta$ ,  $\Phi(\theta, \kappa, q) \equiv \phi(\kappa - q)$ , where  $\phi'(\cdot) > 0$  and  $\phi''(\cdot) \geq 0$ <sup>3</sup>.

The winning buyer *privately* 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 exerting lower effort, providing the buyer with a good of lower expected quality. Adverse selection is introduced because the type of the seller is unknown to buyers and therefore the seller can attempt to increase his gains by misrepresenting his true type.

At equilibrium, buyers play a best response to the seller's expected strategy, given their (correct) beliefs about each type's strategy and their posterior beliefs about the current seller's type. In environments where the seller can credibly commit to producing at a certain quality, the dominant strategy for a seller of type  $\theta$  would be to produce at the *efficient quality level*  $q^*(\theta)$  that maximizes his payoff  $G(q(\theta)) - c(\theta, q(\theta))$ , where  $G(q) = \int_{\underline{q}}^{\bar{q}} w(\kappa)\phi'(\kappa - q)d\kappa$  is a concave function. That quality solves  $G'(q^*(\theta)) = c_2(\theta, q^*(\theta))$ , resulting in payoffs  $V^*(\theta) = G(q^*(\theta)) - c(\theta, q^*(\theta))$ . I assume that the mapping  $q^*(\theta)$  is one-to-one and onto (that is, for each possible quality level, there is exactly one type that finds it optimal to produce at that level) and that  $\partial q^*(\theta)/\partial \theta > 0$ .

In most electronic marketplaces sellers cannot credibly pre-commit to a particular effort level. Sellers then have incentives to exert less effort than what they promise to buyers. Knowing this, buyers will place lower bids. The only possible equilibrium in such cases is one where buyers bid amounts equal to  $G(q)$  and sellers produce at the lowest possible quality  $\underline{q}$ . This equilibrium results in seller payoffs  $G(\underline{q}) - c(\theta, \underline{q})$  that are inefficient for all except the lowest quality type.

### 3. THE MECHANISM

In this section I will show how a simple binary feedback mechanism coupled with rating-contingent listing fees and rewards can restore average social efficiency in the setting of the previous section. More specifically, my objective is to design a mechanism that achieves the following two objectives:

- Induce sellers of all types to truthfully announce their production quality
- Maximize the average social welfare given the probability distribution of seller types on the market

I consider a mechanism that charges sellers a listing fee  $f_0(y) \geq 0$  contingent on their announced expected quality  $y$ . Following completion of an auction, the mechanism asks the winning bidder to rate the seller based on the quality of the good she received. Buyers are given the option of reporting the outcome of a transaction as either "positive" (+)

<sup>3</sup>This simplifying assumption is made purely for the sake of notational simplicity. The mechanism presented in this paper can be designed in the presence of arbitrary probability distributions  $\Phi(\theta, \kappa, q)$ . However, in the general case the form of the resulting expressions is substantially complicated even though the basic intuitions remain the same.

or "negative" (-), with positive ratings indicating that the quality of the good was equal to or superior to the promised quality and negative ratings indicating that the quality of the good was inferior to that promised by the seller<sup>4</sup>.

I assume that buyers always submit ratings that truthfully reflect their quality observations. From a theoretical perspective this can be weakly justified if we make the assumption that buyers only transact with a given seller once (an assumption that is quite reasonable in large-scale electronic markets). Buyers are then indifferent between truthful reporting, untruthful reporting and no reporting. In reality, however, submission of online ratings incurs a small cost associated with the time required to log on to the feedback site and fill the necessary feedback forms. Fortunately, it is not difficult to devise a side payment mechanism that provides buyers with strict incentives to always rate truthfully (see Kandori and Matsushima, 1998; Miller, Resnick and Zeckhauser, 2002). Such a mechanism can be easily combined with the mechanism I present in this paper.

1. Seller announces his expected quality  $y \in [\underline{q}, \bar{q}]$
2. Seller pays fee  $f(y)$  to center; center publishes the seller's expected quality  $y$  to buyers
3. Buyers bid their expected valuations for the good in a second price Vickrey auction; the winning bidder pays  $G$ , equal to the second-highest bid
4. Seller decides to exert effort corresponding to expected quality  $q \in [\underline{q}, \bar{q}]$ , producing a good whose resulting quality  $\kappa$  follows the cumulative probability distribution  $\phi(\kappa - q)$
5. Buyer receives the good, experiences its quality, and realizes the corresponding valuation  $w(\kappa)$ ; buyer rates the seller, reporting "+" if  $\kappa \geq y$  and "-" otherwise
6. Center pays the seller  $f(x, y)$ , where  $x \in \{+, -\}$  is the buyer's report

**Table 1: Summary of bilateral exchange game studied in this paper**

Under the assumption of truthful reporting, if a seller of type  $\theta$  promises quality  $y$  but produces at an effort level associated with quality  $q$ , the probability of a negative rating is given by  $\phi(y - q)$ <sup>5</sup>.

Following the buyer's rating of the transaction outcome, the mechanism makes a payment  $f(x, y) \geq 0$  to the seller. The amount of the payment is contingent on both the an-

<sup>4</sup>Alternative interpretations of positive and negative ratings are also plausible. For example, one can assume that buyers post a negative rating if, after experiencing quality  $\kappa$ , their realized surplus  $w(\kappa) - G(y)$  is negative. Surplus-based rating schemes are, in fact, preferable in marketplaces where a product's quality cannot be easily described using a single scalar variable.

<sup>5</sup>The possibility of "irrational" buyers who make mistakes or fail to rate despite the side-payment mechanism can be easily factored into the function  $\phi(\cdot)$ .

nounced quality  $y$  and the rating  $x \in \{+, -\}$  posted on the seller. This setup is summarized in Table 1.

Given the above setting, the strategy of a seller of type  $\theta$  consists of a quality announcement  $y(\theta)$  and an effort level decision corresponding to a resulting expected quality level  $q(\theta)$ . The seller's objective is to select the strategy  $[y(\theta), q(\theta)] \in [\underline{q}, \bar{q}] \times [\underline{q}, \bar{q}]$  that maximizes:

$$V(\theta) = G(y(\theta)) - c(\theta, q(\theta)) - f_0(y(\theta)) + (1 - \phi(y(\theta) - q(\theta)))f(+, y(\theta)) + \phi(y(\theta) - q(\theta))f(-, y(\theta)) \quad (1)$$

The mechanism's goal is to derive a schedule of fees  $f_0(y)$ ,  $f(-, y)$ ,  $f(+, y) \geq 0$ ,  $y \in [\underline{q}, \bar{q}]$  that achieve the two objectives outlined at the beginning of this section:

**Objective 1:** For all types  $\theta$ , the mechanism induces credible announcements of expected production quality. In other words, given  $\theta$  and  $y(\theta)$ ,  $V(\theta)$  is maximized at  $q(\theta) = y(\theta)$ .

From (1) this requires that, for all types:

$$-c_2(\theta, y(\theta)) + \phi'(0) [f(+, y(\theta)) - f(-, y(\theta))] = 0 \quad (2)$$

and

$$-c_{22}(\theta, y(\theta)) - \phi''(0) [f(+, y(\theta)) - f(-, y(\theta))] < 0 \quad (3)$$

From (2) we get the following necessary condition:

$$f(+, y(\theta)) - f(-, y(\theta)) = \frac{c_2(\theta, y(\theta))}{\phi'(0)} \quad (4)$$

Given our assumptions about  $c_2(\cdot, \cdot) > 0$  and  $\phi'(\cdot) > 0$ , (4) implies that  $f(+, y(\theta)) - f(-, y(\theta)) > 0$ , that is, as expected, refunds are higher following a positive rating. Furthermore, if  $\phi''(0) \geq 0$ , (3) is satisfied for all seller types and (4) becomes a sufficient condition as well. In the rest of the analysis, without loss of generality I will set  $f(-, y(\theta)) = 0$ .

**Objective 2:** Assuming truthful quality announcements, the mechanism induces sellers to produce at qualities that maximize the expected social welfare, given the probability distribution of types in the market.

In the setting of this paper, at equilibrium the expected surplus of buyers is zero, therefore one plausible measure of expected social welfare is the function:

$$W = \int_{\underline{\theta}}^{\bar{\theta}} p(\theta) V(\theta) d\theta \quad (5)$$

where  $p(\theta)$  is the probability density function of types in the market. Assuming truthful quality announcements (and therefore (4)) and  $f(-, y(\theta)) = 0$  the payoffs of seller type  $\theta$  become:

$$V(\theta) = G(y(\theta)) - c(\theta, y(\theta)) - f_0(y(\theta)) + \frac{1 - \phi(0)}{\phi'(0)} c_2(\theta, y(\theta)) \quad (6)$$

By substituting (6) into (1) we get:

$$W = \int_{\underline{\theta}}^{\bar{\theta}} p(\theta) [G(y(\theta)) - c(\theta, y(\theta))] d\theta + \int_{\underline{\theta}}^{\bar{\theta}} p(\theta) \left[ -f_0(y(\theta)) + \frac{1 - \phi(0)}{\phi'(0)} c_2(\theta, y(\theta)) \right] d\theta \quad (7)$$

subject to the seller individual rationality constraints:

$$V(\theta) \geq 0 \Leftrightarrow f_0(y(\theta)) \leq [G(y(\theta)) - c(\theta, y(\theta))] + \frac{1 - \phi(0)}{\phi'(0)} c_2(\theta, y(\theta)) \quad (8)$$

The problem of maximizing (7) has a particularly simple solution in situations where we can assume upfront that seller individual rationality constraints are satisfied. This is the case if rents under perfect information are sufficiently high for all seller types. In the following analysis we will make that assumption. Any solution derived using this assumption must therefore be checked against (8) before it can be accepted.

The first term in the expression (7) is maximized when all types produce at their efficient quality levels, i.e. when for all  $\theta$ ,  $y(\theta) = y^*(\theta)$ , where  $y^*(\theta)$  solves:

$$G'(y^*(\theta)) - c_2(\theta, y^*(\theta)) = 0 \quad (9)$$

The second term represents the average net payment made to the seller by the center through the mechanism of listing fees and rating-contingent refunds. In order for the center to participate in the mechanism, the center must, on the average, not lose money. This requires that the average net payment be less than or equal to zero. Given this constraint, social welfare of buyers and sellers is maximized when the fees are chosen so that:

For all  $\theta$ ,  $V(\theta)$  is maximized at  $y(\theta) = y^*(\theta)$   
The center breaks even<sup>6</sup>:

$$\int_{\underline{\theta}}^{\bar{\theta}} p(\theta) \left[ -f_0(y^*(\theta)) + \frac{1 - \phi(0)}{\phi'(0)} c_2(\theta, y^*(\theta)) \right] d\theta = 0 \quad (10)$$

In order for  $V(\theta)$  to be maximized at  $y(\theta) = y^*(\theta)$  it must be:

$$V_2(\theta, y^*(\theta)) = 0 \quad (11)$$

$$V_{22}(\theta, y^*(\theta)) < 0 \quad (12)$$

The first-order condition requires that:

$$G'(y^*(\theta)) - c_2(\theta, y^*(\theta)) - f'_0(y^*(\theta)) + \frac{1 - \phi(0)}{\phi'(0)} c_{22}(\theta, y^*(\theta)) = 0 \quad (13)$$

From (9) and (13):

<sup>6</sup>It is, of course, perfectly plausible that the marketplace operator may want to realize some profits for operating the market. In this case, equation (10) can easily be modified so that the schedule of fees results in a net profit for the marketplace. If we include the marketplace operator in the community, this modification still results in maximum community social welfare.

$$f'_0(y(\theta)) = \frac{1 - \phi(0)}{\phi'(0)} c_{22}(\theta, y^*(\theta)) \quad (14)$$

Since we have assumed that the map  $y^*(\theta)$  is bijective, (14) can be rewritten as:

$$f'_0(y) = \frac{1 - \phi(0)}{\phi'(0)} c_{22}(\theta^*(y), y) \quad (15)$$

or equivalently:

$$f_0(y) = \frac{1 - \phi(0)}{\phi'(0)} \int c_{22}(\theta^*(y), y) dy + K \quad (16)$$

where  $\theta^*(y)$  solves  $G'(y) - c_2(\theta^*(y), y) = 0$ . The constant  $K$  is calculated from (10).

The second-order condition requires that:

$$G''(y) - c_{22}(\theta^*(y), y) - f''_0(y) + \frac{1 - \phi(0)}{\phi'(0)} c_{222}(\theta^*(y), y) < 0 \quad (17)$$

By substituting (9) and (15), this condition is simplified to:

$$c_{21}(\theta^*(y), y) - \frac{1 - \phi(0)}{\phi'(0)} c_{221}(\theta^*(y), y) < 0 \quad (18)$$

If (18) holds and individual rationality constraints (8) are satisfied for all  $\theta$  then the optimal schedule of fees is given from (4) and (16):

$$f_0(y) = \frac{1 - \phi(0)}{\phi'(0)} \int c_{22}(\theta^*(y), y) dy + K$$

$$f(+, y) = \frac{c_2(\theta^*(y), y)}{\phi'(0)} \quad (19)$$

$$f(-, y) = 0$$

The above mechanism succeeds in inducing all sellers to produce at their respective first-best qualities and to truthfully announce those qualities to buyers. Knowing this, buyers are expected to bid amounts equal to those corresponding to the complete information case. Finally, the mechanism fees can be calculated so that the center breaks even; this achieves full average social efficiency for the community of buyers and sellers.

On the other hand the mechanism introduces distortions to *individual* seller payoffs (relative to the complete information case). More specifically, the mechanism transfers some payoffs from more efficient to less efficient sellers. To see this, from (6), this distortion is equal to:

$$\Delta(\theta) = -f_0(y(\theta)) + \frac{1 - \phi(0)}{\phi'(0)} c_2(\theta, y(\theta)) \quad (20)$$

Differentiating (20) and taking into account (9) and (15) we get:

$$\frac{d\Delta(\theta)}{d\theta} = \frac{1 - \phi(0)}{\phi'(0)} c_{21}(\theta, y(\theta)) < 0 \quad (21)$$

The above equation shows that, as efficiency grows, the net payments received by the seller from the center decrease. If, in addition, (10) holds, this can only mean that less efficient types are net receivers of center payments (and thus increase their payoffs) while more efficient types are net payers (and thus decrease their payoffs).

The intuition behind this cross-subsidization is simple: in order for the center to remove incentives for less efficient types to try to masquerade as more efficient types, it has to make it relatively more attractive to produce at lower qualities and less attractive to produce at higher qualities. It achieves this by transferring part of the fees paid by sellers that announce high qualities to sellers that announce low qualities.

## 4. A CONCRETE EXAMPLE

This section presents an example of how the mechanism described in the previous section applies in a concrete setting and discusses the properties of the resulting outcomes. I consider a marketplace where expected bid prices and costs are given by:

$$G(y) = A \log y + B$$

$$c(\theta, y) = D(2 - \theta)^2 y^2$$

Types range from  $\underline{\theta} = 0$  to  $\bar{\theta} = 1$  and qualities from  $\underline{q} = 0.5\sqrt{A/2D}$  to  $\bar{q} = \sqrt{A/2D}$ . The efficient quality level  $y^*(\theta)$  for type  $\theta$  is given by the solution of  $G'(y^*(\theta)) = c_2(\theta, y^*(\theta))$ :

$$y^*(\theta) = \sqrt{\frac{A}{2D}} \frac{1}{2 - \theta}$$

Equivalently,

$$\theta^*(y) = 2 - \sqrt{\frac{A}{2D}} \frac{1}{y}$$

From equations (19), in order to induce sellers to announce and produce at their respective efficient quality levels the mechanism must set:

$$f(+, y) = \frac{c_2(\theta^*(y), y)}{\phi'(0)} = \frac{Ay^{-1}}{\phi'(0)} \quad (22)$$

$$f_0(y) = \frac{1 - \phi(0)}{\phi'(0)} \int c_{22}(\theta^*(y), y) dy + K = -\frac{1 - \phi(0)}{\phi'(0)} Ay^{-1} + K \quad (23)$$

The second-order condition (18) further requires that

$$\frac{1 - \phi(0)}{\phi'(0)} < \underline{q} = \frac{1}{2} \sqrt{\frac{A}{2D}}$$

Average social welfare of buyers and sellers is maximized if, in addition (10) is satisfied. For the sake of this example, let us assume that all types  $\theta \in [0, 1]$  are equally likely, that is  $p(\theta) = 1$ . This gives:

$$\int_{\underline{\theta}=0}^{\bar{\theta}=1} p(\theta) \left[ 2 \frac{1 - \phi(0)}{\phi'(0)} \sqrt{2AD} (2 - \theta) - K \right] d\theta = 0$$

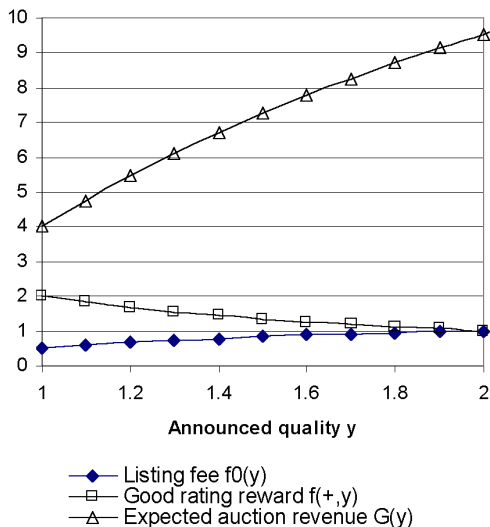
and therefore

$$K = 3 \frac{1 - \phi(0)}{\phi'(0)} \sqrt{2AD} \quad (24)$$

From equations (22), (23) and (24) we can now calculate the exact expressions for the fees  $f_0(y)$  and rewards  $f(+, y)$ . These are depicted in Figure 1 as a function of a seller's announced quality  $y$  for a representative set of parameters (revenue parameters  $A = 8, B = 4$ , cost parameter  $D = 1$ , and  $\phi(0) = 0.5, \phi'(0) = (0.1\sqrt{2\pi})^{-1}$ , corresponding to a Gaussian cdf  $\phi(\cdot)$  with mean 0 and standard deviation  $\sigma = 0.1$ ). For this set of parameters, qualities can range from  $\underline{q} = 1$  to  $\bar{q} = 2$ . The corresponding expected auction revenue is also shown for comparison. Figure 2 shows the fees as a percentage of the corresponding expected revenue. We see that, even though in our model the probability of a positive rating following a seller's truthful announcement of his intended production quality is independent of the exact level of that quality, refunds following a positive rating are highest when a seller announces the lowest possible quality and drop steeply as this quality grows.

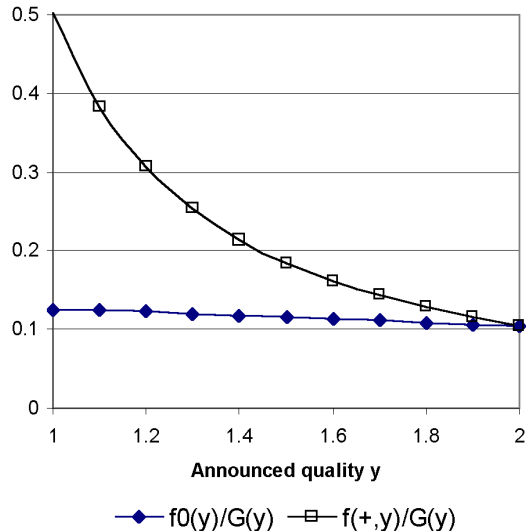
Finally, from equation (6) the resulting payoffs for a seller of type  $\theta$  are equal to:

$$V(\theta) = \left[ A \log \left( \sqrt{\frac{A}{2D}} \frac{1}{2-\theta} \right) + B - \frac{A}{2} \right] + \left[ \frac{1-\phi(0)}{\phi'(0)} \sqrt{2AD} (1-2\theta) \right] \quad (25)$$



**Figure 1: Schedule of mechanism fees and expected auction revenue as a function of announced quality**

The first term of equation (25) gives the seller payoffs under complete information. The second term corresponds to the distortion on the payoffs introduced by the mechanism. We see that, in order to induce all seller types to truthfully announce and produce at their respective first-best qualities, the mechanism increases the payoffs of less efficient types and reduces the payoffs of more efficient types relative to perfect information. Such distortions constitute a general property of contracting mechanisms under asymmetric



**Figure 2: Mechanism fees as a fraction of expected auction revenue**

information, discussed in great detail in (Laffont and Tirole, 1993).

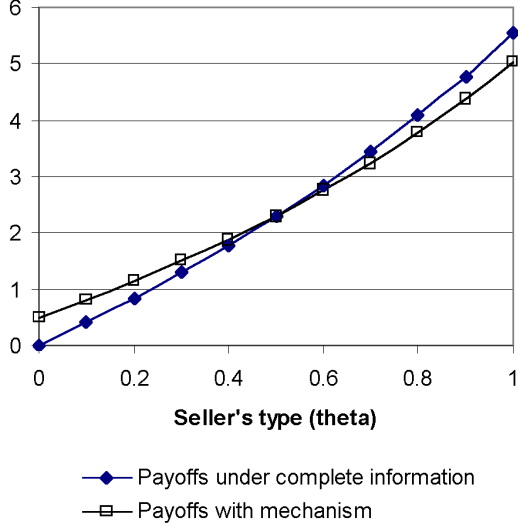
Figure 3 shows the seller payoffs  $V(\theta)$  as a function of a seller's type and compares them against the corresponding payoffs under complete information for the same parameters used to derive Figures 1 and 2. Since this type of distortion gives incentives to efficient sellers to migrate to other markets with better informational structures, in the presence of competing marketplaces it is in the interest of a marketplace operator to keep it as small as possible. From equations (16) and (20) we see that the distortion is proportional to  $\frac{1-\phi(0)}{\phi'(0)}$ , which is a measure of the noise introduced by imperfect private monitoring and reporting. In our model  $\phi'(0)$  represents the rate of increase of the probability of a negative rating if a seller deviates from truth telling. The larger that rate, the smaller the distortion introduced by the mechanism.

## 5. ROBUSTNESS AND IMPLEMENTATION CONSIDERATIONS

An important consideration in mechanism design is the robustness of a mechanism to the assumed features of the underlying environment. The sheer number and heterogeneity of participants in large-scale electronic markets makes such robustness considerations particularly important.

*Easy name changes.* An attractive property of the mechanism presented in this paper is that it induces cooperation and achieves efficiency in a one-shot setting. This is a particularly desirable property in electronic markets where the relative ease with which players can change online identities reduces the efficiency of mechanisms that rely on repeated interaction (Friedman and Resnick, 2001).

*Inept types.* I define an inept type as a seller who has no production facilities and enters an electronic market with the explicit intention to cheat, i.e. advertise products, receive money but never send anything back to buyers, for as long as he can. Given the relative ease with which players can



**Figure 3: Distortion of seller payoffs as a function of the seller’s efficiency (type)**

anonymously enter (and exit) most online markets, inept types are plausible and mechanisms ought to be robust to their existence. Robustness in this context requires that inept types cannot realize any profits by advertising at any quality and thus will not find it profitable to enter.

I assume that inept types incur zero production costs and receive negative ratings with probability  $1 - \varepsilon$ , where  $\varepsilon$  would normally be very close to zero. Then, from (8), their expected payoffs if they announce quality  $y$  are given by:

$$V_{inept}(y) = G(y) - f_0(y) + \varepsilon \frac{c_2(\theta^*(y), y)}{\phi'(0)} \quad (26)$$

In order for inept types to stay out of the market, it must be:

$$f_0(y) > G(y) + \varepsilon \frac{c_2(\theta^*(y), y)}{\phi'(0)} \quad (27)$$

for all  $y$ . In most environments this would require that the listing fees paid to the center be *increased* relative to those calculated by (10) so that (27) is satisfied for all  $y$ . Keeping inept sellers out of the market thus requires the payment of positive average rents to the center. From the perspective of “normal” sellers this can be viewed as inefficiency. From the perspective of marketplace operators, however, this is an interesting business opportunity.

If the fees required to keep inept sellers out of the market are too high, the center can periodically redistribute part of those fees back to sellers who maintain *average* ratings at acceptable levels. If the ratings profile of inept sellers is substantially distinct from that of “normal” sellers this can still discourage inept sellers from entering while maintaining the payoffs of “normal” sellers at acceptable levels.

*Honest types.* “Irrational” honest types who always perform what they promise are, perhaps, just as plausible in large-scale environments as inept types. Given that our mechanism induces strategic types to behave as if they were honest, assumption of honest types does not change the out-

comes induced by the mechanism.

*Implementation considerations.* The mechanism is straightforward to implement in online auction marketplaces provided that (a) there is a simple and relatively unambiguous way to communicate intended product quality  $y$  to the center and buyers (more ambiguity increases noise), (b) the center has reliable estimates of the set of possible seller cost functions  $c(\theta, y)$  and ratings distribution  $\phi$  (reasonable, if the center has some knowledge of the economics of the respective industry segment) (c) the center can estimate the expected auction revenue function  $G(y)$  (easy if the center keeps track of past auction results) and (d) for each product quality  $y$ , there is exactly one type  $\theta$  for which  $y$  is the first-best quality. Under the above assumptions, the basic idea of the mechanism applies even to more complex marketplaces where types  $\theta$  are multi-dimensional,  $y$  is communicated via a vector of attributes, and buyers supply separate ratings for each quality attribute.

## 6. CONCLUDING REMARKS

The establishment of incentives for cooperative behavior in online environments where state enforcement of contractual promises is usually difficult or too costly has emerged as one of the most important problems associated with the design of electronic markets. Feedback mechanisms constitute a promising approach in this context.

Recent results from economics show that “passive” feedback mechanisms that simply accumulate and publish past ratings are fragile in the long run or incur efficiency losses. This paper has shown how the combination of a simple feedback mechanism with the ability of a marketplace operator to levy listing fees from sellers can restore full average social efficiency in environments with both moral hazard and adverse selection.

The desirable properties of the presented mechanism include the ability to induce production at first-best quality levels, truthful announcement of qualities and expected auction revenues equal to those in environments with complete information. This desirable equilibrium behavior remains stable over time. Finally, the mechanism is robust (or can be extended to become robust) to a number of contingencies of particular concern in online environments, such as easy name changes and inept sellers.

The principal drawback of the mechanism is that it results in some distortion to individual seller payoffs, essentially transferring part of the payoffs of more efficient sellers to less efficient sellers. From contract theory it is known that, in the presence of asymmetric information, such distortions are, in principle, inevitable. In the presence of multiple competing marketplaces, however, such distortions are undesirable because they may cause the most efficient sellers to migrate to other markets with better information structures. It would therefore be interesting to investigate whether the magnitude of distortions can be minimized through more elaborate feedback mechanism design. For example, one remarkable feature of the current mechanism is that it is capable of inducing full cooperation and efficient production by relying on simple binary ratings. An interesting open question is whether the use of more elaborate rating schemes (such as multi-valued ratings or better targeted questions that induce more favorable rating distributions) can result in mechanisms that introduce less distortion.

## 7. ACKNOWLEDGMENTS

This research was supported by NSF CAREER GRANT IIS-9984147 and an MIT eBusiness Vision Fund Award.

## 8. REFERENCES

- [1] Cripps, M., Mailath, G. and Samuelson., L. (2002) Imperfect Monitoring and Impermanent Reputations. Penn Institute for Economic Research Working Paper 02-021<sup>7</sup>.
- [2] Dellarocas, C. (2003) The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms. *Management Science* (forthcoming).
- [3] Friedman, E. and Resnick, P. (2001) The Social Cost of Cheap Pseudonyms. *Journal of Economics and Management Strategy* 10 (1), pp. 173-199.
- [4] Kandori, M. and Matsushima, H. (1998) Private Observation, Communication and Collusion. *Econometrica*, 27 (3), pp. 627-652.
- [5] Kreps, D., Milgrom, P., Roberts, J and Wilson, R. (1982). Rational Cooperation in the Finitely Repeated Prisoners' Dilemma. *Journal of Economic Theory*, 27, pp. 245-52.
- [6] Kreps, D. and Wilson, R. (1982) Reputation and Imperfect Information. *Journal of Economic Theory*, 27, pp. 253-79.
- [7] Laffont, J.-J. and Tirole, J. (1993) *A Theory of Incentives in Procurement and Regulation*. MIT Press, Cambridge, MA.
- [8] Milgrom, P., and Roberts, J. (1982) Predation, Reputation and Entry Deterrence, *Journal of Economic Theory* 27, pp. 280-312.
- [9] Miller, N., Resnick, P., and Zeckhauser, R. (2002) Eliciting Honest Feedback in Electronic Markets. Harvard Kennedy School Research Working Paper RWP02-039<sup>8</sup>.
- [10] Resnick, P., Zeckhauser, R., Friedman, E., Kuwabara, K. (2000) Reputation Systems. *Communications of the ACM*, 43 (12), pp. 45-48.
- [11] Resnick, P., and Zeckhauser, R. (2002) Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. In *The Economics of the Internet and E-Commerce*. Michael R. Baye, editor. Volume 11 of Advances in Applied Microeconomics, JAI Press.

---

<sup>7</sup> Available at: <http://www.econ.upenn.edu/Centers/pier/Archive/02-021.pdf>

<sup>8</sup> Available at: <http://www.si.umich.edu/~presnick/papers/elicite/index.html>