Simple Reputation Systems*

John Kennes
CAM, Institute of Economics
University of Copenhagen
john.robert.kennes@econ.ku.dk

Aaron Schiff
Department of Economics
University of Auckland
a.schiff@auckland.ac.nz

November 11, 2004

Abstract

This paper develops a model of simple 'reputation systems' that monitor and publish information about the behavior of sellers in a market with search frictions and asymmetric information. The reputations created by these systems influence the equilibrium search investments of buyers and thus provide for market-based 'punishment' of bad behavior. Our model allows us to determine the effects of the introduction of a reputation system on the behavior and welfare of buyers and sellers in such a market. We show that a simple reputation system that rewards honesty can enhance welfare by allowing good sellers to truthfully signal their type. However, we also show that the same reputation system can delay the creation of such information if bad sellers minic the signalling strategy of good sellers. In this case, we show that an alternative simple reputation system that screens for type can be superior.

1 Introduction

Many markets are characterized by both search frictions and asymmetric information between buyers and sellers. For example, tourists visiting a new city have only a limited amount of time to visit tourist attractions, and lack complete information about which attractions are the best. Other examples include diners looking for a restaurant and online auctions where buyers have very little

* **Acknowledgements:** This paper was presented at the 2004 Society of Economic Dynamics general meeting. The authors thank Heiko Gerlach, John Hillas, Ian King, Klaus Kultti, Matthew Ryan, Torben Tranæs, Jean Robert Tyran, Andrew Wait, and Julian Wright for very helpful comments and suggestions. Aaron Schiff thanks the Centre for Research in Network Economics and Communications and John Kennes thanks the Centre for Applied Microeconometrics (CAM) for financial support.

first hand information about sellers' past actions. In such markets, 'reputation systems' often exist to guide buyers in their search.

A reputation system is typically operated by a third party who does not transact in the market. For example, guidebooks exist to help tourists find the best activities and diners to find the best restaurants. In online auctions, the market operator (such as eBay) usually operates a system that collects information on the behavior of sellers (and buyers) and publishes this information to all market participants. Such third-party reputation systems can play an important role in ameliorating the effects of asymmetric information in markets where there are few or no repeated interactions between buyers and sellers.

In this paper we investigate the effects of simple reputation systems on a search market in the presence of asymmetric information. Our first goal is to understand how the behavior of buyers and sellers in this market is affected, and the consequent welfare effects. Our second goal is to compare different simple reputation systems and determine their relative efficiency in different types of markets.

Many common reputation systems choose to advertise the quality of sellers by a 1,0 metric, because it is both easy to produce and easy to use. Such a simple assignment of reputation occurs whenever a professional association accepts a new member or a guidebook gives a destination a simple thumbs-up. These simple reputation systems reduce market failures by ensuring that buyers are better informed about the quality of sellers and their products. However, all share the basic limitation of presenting only an incomplete description of sellers' histories.

When designing a simple reputation system, we have to take a stand on what we believe is important for the determination of the 1s and 0s. There are several alternative algorithms that could be used. In this paper we consider two relatively simple possibilities. First, in a market with many anonymous buyers and sellers, one type of reputation system could track whether a seller was honest in the past. Therefore, dishonestly would count as a strike against the seller yielding a 0 while honesty would yield a 1. This is analogous to the kind of reputation system operated by eBay and other online auction sites.

Alternatively, a reputation system might report something about the quality of the past transactions. For example, a bad transaction would count as a strike against the seller yielding a 0 while the alternative would yield a 1. This is similar to how guidebooks for tourists and restaurants work.

This paper provides a framework for evaluating the costs and benefits of simple reputation systems and determining whether there are any particular grounds for choosing one simple reputation system over another. In this paper we focus on the two types of reputation system described above, but the framework we provide could be adapted to analyze more complicated systems.

How do these two simple reputation systems differ in practice? On the one hand, as we show in this paper, a reputation system that screens for honesty allows sellers who usually have good quality products for sale to do the right thing and signal their type in the relatively rare events that they have bad quality products to sell. Thus the credible information about seller quality is created at very early periods in a seller's tenure. However, the potential downside of a reputation system for honesty is that sellers who usually have bad quality products for sale may attempt to mimic this strategy in order to delay the public realization that they are low quality. If this excessive honesty strategy by low quality sellers is optimal, the social gains from signalling are reduced and, as we show in this paper, a reputation system that screens for honesty is potentially inefficient relative to a system that screens for type.

Our analysis is framed in a two period general equilibrium search model of a market where there is asymmetric information between buyers and heterogeneous sellers, and sellers use competing auctions to sell their goods. The strategy available to sellers is whether to advertise honestly and the strategy available to buyers is where to search and what to bid subsequent to the sunk search investment. The reputation system matters for both the search and advertisement decisions because it links actions in the first period with information and payoffs in the second period. The fact that the reputation system is simple (i.e. imperfect) implies that some sellers are potentially dishonest in equilibrium.

We caution that our model abstracts from two problems that are also pertinent to the performance of a reputation system. The first abstraction is that we do not allow the possibility of cheap pseudonyms where sellers can exit in one period and reappear in the next with an untarnished reputation (see, for example, Friedman and Resnick 2001). We can rule out this opportunistic behavior simply by assuming that the reputation system awards the cheap pseudonym strategy a 0 reputation in the second period. The second abstraction is that we do not allow sellers a choice over the set of products for sale. Therefore, there is no possibility that a seller might choose to honestly sell one set of products in period one in order to get a good reputation and then choose to sell another set of products in the second period to exploit this reputation.[1] Consequently, our analysis demonstrates that such choices are not essential to there being a trade-off between reputation systems.

We also seek to avoid some ambiguities that exist in standard off-the-shelf general equilibrium search models. In particular, these macro models often predict market inefficiencies related to inefficient price formation – deviations from the Hosios rule – in addition to the problem that we have posed of an potentially inefficient reputation system. We avoid this complication by focusing on a decentralized trading structure in which buyers use information efficiently

[1] McAfee (2004) discusses this problem in connection to the reputation system of eBay.

within each period. In particular, this outcome is the result of our assumption that the prices in the local market of each seller are determined by a simple auction (i.e. Bertrand competition between buyers).[2]

There has been much research on the problem of asymmetric information since the seminal contributions of Akerlof (1970), Spence (1973) and Rothschild and Stiglitz's (1976). We extend this research by deriving a model in which past indiscretions are revealed to traders only by a reputation system, that is, we explicitly rule out the possibility of repeated interactions. This assumption is contrasted by a number of formal models of reputation that have the assumption of asymmetric information (Kreps and Wilson, 1982 and Diamond, 1989) but do not address the problem of how information about past indiscretions is gathered and reported. Random matching models by Ellison (1994) and Tirole (1996) address some of the issues related to our paper. However, these models do not attempt to compare alternative reputation systems.

The Internet has seen an explosion of formal reputation system algorithms, which are used by commercial websites such as eBay.[3] This is likely due to the lower costs associated with collecting, analyzing, and disseminating information in electronic form. This development has led to a great deal of theoretical research to develop and compare reputation systems, not only by economists, but also by sociologists and computer scientists. However, much of this work rests on either behavior assumptions, or in the case of economists, is driven by experimental research, with the implicit assumption being one of bounded rationality.[4] Our simple finite horizon model has fully rational agents.

The paper is organized as follows. In the next section we present a simple two period model of seller advertisements, buyer search investments, and reputational assignment. We then solve for the equilibrium of this model sequentially, considering: (i) buyers' choice of seller to bid for, (ii) the partition of sellers into quality differentiated submarkets under each reputation system, and (iii) the choice of reputation system. The equilibrium properties of the model are analyzed, and it is used to compare the efficiency of the two reputation systems. The final section offers concluding remarks.

2 The Model

A market operates for two periods, denoted $t = 1, 2$. The number of buyers in the market equals the number of sellers and is normalized to 1. These numbers

[2] McAfee (1993) offers an early model of competing auctions. Kennes (2004) offers a survey of recent research.

[3] See Zacharia and Maes (2000) and Dellarocas (2003) for an overview.

[4] Most of these experiments evaluate the actions of agents in a finite event horizon and seek to find whether the agents actions are disciplined by a set of trigger strategies, which are theoretically optimal only in an infinite event horizon model.

are also sufficiently large that the set of each type of agent can be treated as a continuum distributed on the unit interval.

2.1 Sellers and products

Sellers offer products for sale that have one of two quality levels: *high* and *low*. Let $Q = \{\theta, 1\}$ denote the set of product quality levels, where $0 \leq \theta < 1$ denotes the quality level of the low quality product relative to the high quality product.

The sellers are divided into two equal groups, which are distinguished by types: *good* and *bad*. At the start of each period, every seller draws one product from a probability distribution defined over Q. The distribution is such that the bad sellers always have a low quality product for sale. Good sellers have a high quality product for sale with some strictly positive probability $0 < \gamma \leq 1$ and thus $\widehat{q} = \gamma + (1 - \gamma)\,\theta$ denotes the expected quality of a good seller's product. The average product quality of all sellers is

$$\widetilde{q} = \frac{\widehat{q} + \theta}{2}. \tag{1}$$

In each period the sellers advertise, possibly untruthfully, whether they have a high quality or low quality product. All advertisements are seen by all buyers. Following the advertisements, every seller sells their product using an ascending bid auction. For simplicity, the reserve price at every auction is assumed to be zero.

2.2 Buyers and bidding

Each buyer i seeks to buy one unit of the product in each period. Buyers are identical in their willingness to pay for quality and the net utility function of buyers over all outcomes of bidding at an auction of seller j in period t is given by

$$u_i\left(q_j^t, p_{ij}^t\right) = \begin{cases} q_j^t - p_{ij}^t & \text{if } p_{ij}^t \text{ is the winning bid} \\ 0 & \text{otherwise} \end{cases},$$

where q_j^t is the quality of seller j's product and p_{ij}^t is the bid of buyer i at seller j's auction.

A buyer can purchase the product only by going to a seller's location and participating in that seller's auction. Upon visiting a seller, the buyer becomes perfectly informed of the good's quality, before bidding commences. The bidding at seller j's auction depends on the number of buyers visiting seller j. Under an ascending bid auction, if m_j^t is the number of buyers choosing to visit seller j in period t, buyer i maximizes utility by the bidding strategy

$$p_{ij}^t\left(m_j^t\right) = \begin{cases} 0 & \text{if } m_j^t = 1 \\ q_j^t & \text{if } m_j^t > 1 \end{cases}.$$

Thus, a seller receives a non-zero price for his or her product if and only if more than one buyer turns up to the auction.

We make these relatively simplistic assumptions about the auctions in the model so as to focus on the effects of reputation systems that come from the aggregation and dissemination of information about past seller behavior.

We also assume that buyers can choose to visit only one seller each period and never purchase from the same seller twice, that is, we rule out long-term relationships.

2.3 Submarkets

In each period the sellers may be separated into two quality differentiated submarkets – by their advertisements in the first period and, in the presence of a reputation system, by their reputations in the second period.[5] Let q_l^t and q_h^t denote the expected quality levels of sellers in the two submarkets in period t, and let α^t denote the fraction of sellers that are allocated to the submarket with expected quality q_l^t. Without loss of generality we assume $q_h^t \geq q_l^t$. If sellers are separated in such a manner, the average quality of sellers across submarkets cannot change, thus

$$\alpha^t q_l^t + \left(1 - \alpha^t\right) q_h^t = \widetilde{q} \quad \text{for } t = 1, 2 \tag{2}$$

Observe that (2) implies that, for a given value of $\widetilde{q}$, one of α^t, q_l^t or q_h^t can be recovered from knowledge of the other two parameters. We thus define an an information partition as follows.

Definition 1 *An* information partition *is a pair* (α, q_l) *where* $0 \leq \alpha \leq 1$ *is the proportion of sellers allocated to the submarket with quality level* $0 \leq q_l < \widetilde{q}$.

Given an information partition (α^t, q_l^t) in period t, we can calculate the expected quality in the high quality submarket as

$$q_h^t\left(\alpha^t, q_l^t\right) = \frac{\widetilde{q} - \alpha^t q_l^t}{1 - \alpha^t} \tag{3}$$

Buyers are informed of q_l^t, q_h^t and α^t and simultaneously choose to enter one submarket in each period. Let ϕ denote the buyer-seller ratio or market tightness of a particular submarket. We use ϕ_l^1 and ϕ_h^1 to denote the measures of market tightness in the first-period submarkets defined by sellers who advertise low and high quality respectively. Similarly, we use ϕ_l^2 and ϕ_h^2 to denote the market tightnesses in the second-period submarkets defined by sellers who have bad and good reputations respectively. The number of buyers equals the number

[5] Although we allow sellers to advertise in the second period also, such advertisements will have no informational content, due to a lack of credibility caused by the finite horizon of the game. Thus, in the second period, only reputations can define submarkets.

of sellers, so market tightness for each submarket is related to overall market tightness as follows:

$$\alpha^t \phi_l^t + \left(1-\alpha^t\right) \phi_h^t = 1 \quad \text{for } t = 1, 2. \tag{4}$$

From (4) we have

$$\phi_h^t \left(\alpha^t, \phi_l^t\right) = \frac{1-\alpha^t \phi_l^t}{1-\alpha^t} \tag{5}$$

2.4 Search frictions and payoffs

Search frictions exist because the buyers in each submarket make uncoordinated search investments when they choose the location of a single capacity constrained seller. The search investment of each buyer is directed by the common set of submarkets, which may be created by the advertisements and reputations of sellers. In a large market the probability distribution that exactly b buyers turn up to a seller's auction in a (sub)market with tightness ϕ is given by:[6]

$$\Pr(x=b) = \begin{cases} e^{-\phi} & \text{for } b=0 \\ \phi e^{-\phi} & \text{for } b=1 \\ 1-e^{-\phi}-\phi e^{-\phi} & \text{for } b>1 \end{cases}.$$

The most important of these probabilities for the seller is that for $b > 1$, since only in this case does the seller receives a non-zero price from his or her auction. Accordingly, we define the function $p(\phi) = 1 - e^{-\phi} - \phi e^{-\phi}$. The expected profit of a seller with product quality q_j in a submarket of tightness ϕ is then given by

$$V_j = p(\phi)\, q_j.$$

From a buyer's point of view, what matters is whether or not they are alone at a seller's auction, since if they are alone they will get a strictly positive surplus from the auction while if they are not alone their surplus will be zero. The probability that a buyer is alone at any given seller in a submarket with market tightness ϕ is given by $e^{-\phi}$ and the expected utility of buyer i of visiting a seller in a submarket with tightness ϕ is given by

$$U_i = e^{-\phi} q_z,$$

where q_z is the expected quality of products in this submarket.

2.5 The reputation system

At the end of the first period, the reputation system collects information on seller behavior and assigns reputations to sellers according to some algorithm. At the beginning of the second period, buyers observe sellers' reputations, as well as their new advertisements,[7] before choosing which seller's auction to visit.

[6] See Kennes (2004) for details.

[7] Recall, however, that second-period advertisements have no informational content.

We model the reputation system by introducing a parameter k which represents the probability that a seller receives a bad reputation, if the seller's behavior meets the reputation system's criteria for assigning a bad reputation. Thus k represents the effectiveness of the reputation system.[8]

We define two different types of simple reputation system, as follows.

Definition 2 *A reputation system* screens for honesty *if sellers that lied about their product quality in period 1 are assigned a bad reputation with probability k in period 2.*

Definition 3 *A reputation system* screens for type *if sellers that sold low quality in period 1 are assigned a bad reputation with probability k in period 2.*

Note that the reputation system does not have access to any more information than buyers do, however it has the ability to aggregate and publish this information. That is, the reputation system can 'see' (imperfectly, if $k < 1$) the advertisements of sellers and the quality of the goods that were actually sold. However it cannot determine whether a seller is of the good type or the bad type.

For the purposes of this paper, we take k as given. If reputations were assigned on the basis of buyer reports, as occurs on eBay, then k would become endogenous. In particular, under such a 'customer-report reputation system', the value of k would be proportional to the probability that at least one buyer turns up to a seller's auction in the first-period submarket of sellers who advertise high quality, that is, $k \sim 1 - e^{-\phi_h^1}$. However, basing reputations on customer reports means that we much also consider potential strategic behavior and incentives of buyers when making these reports. For simplicity we exclude these effects from the present paper, but we believe they are a useful avenue for future research.

2.6 Timing

The timing of the game with a reputation system is as follows. At the start of period 1, each seller observes their type (good or bad) and their product quality (high or low). They then choose to advertise either high or low product quality. Buyers observe the sellers' advertisements and simultaneously and independently choose one seller to visit. Buyers then bid on the seller's product and the good is sold to the highest bidder. At the start of period 2, a reputation system assigns a reputation to each seller based on events in the previous period

[8] The parameter k can be interpreted as the probability that a randomly selected seller is 'audited' by the reputation system. The conditional probability that a seller's reputation is correct, given that they have been audited, is assumed to be 1. Thus we assume that any imperfection in the reputation system stems from a lack of auditing, rather than a lack of accuracy of auditing.

and the properties of the reputation system. Sellers draw a new product according to their type. Buyers choose a reputational submarket and simultaneously and independently choose one seller to visit. Bidding then takes place, the good is sold to the highest bidder, and the game ends.

3 Equilibrium

In this section we characterize the equilibrium of the model. For comparison, we first briefly characterize the equilibrium in the case where all sellers appear identical to all buyers, that is, where there are no submarkets.

We then turn to the case where the existence of a reputation system creates submarkets in one or both periods. The equilibrium of the model with a reputation system is found backwards induction. In particular, we first evaluate the buyers' choice of seller to bid for given the sellers advertisement/ reputation. We then consider the sellers' choice of advertisement given the equilibrium strategy of buyers and the reputation imposed by the particular reputation system under investigation.

For much of this section we drop the t superscripts where possible for notational simplicity.

3.1 No submarkets

In the absence of quality differentiated submarkets, buyers randomize over the locations of all sellers. As the buyer-seller ratio is assumed to be 1, any given seller receives at least one buyer with probability $1 - e^{-1}$. Over both periods, total welfare is therefore given by

$$W_0 = 2\left(1 - e^{-1}\right)\widetilde{q} \tag{6}$$

The welfare loss due to search frictions is $2e^{-1}\widetilde{q}$.

3.2 Equilibrium distribution of buyers

The division of sellers into quality differentiated submarkets in a period leads to two basic configurations for the equilibrium distribution of buyers across the submarkets in that period. In the first, sellers in both submarkets are visited by all buyers with strictly positive probability. In such a mixed strategy equilibrium, we must have

$$q_h e^{-\phi_h} = q_l e^{-\phi_l}. \tag{7}$$

That is, the expected utility to buyers must be the same from locating in either submarket.

However, suppose that all buyers locate in the high quality submarket. In this case, $\phi_h = 1/(1-\alpha)$ and $\phi_l = 0$. If what we call the *exclusion constraint*,

$$q_h e^{-\frac{1}{1-\alpha}} \geq q_l \tag{EC}$$

is satisfied then a buyer is better off to locate in the high quality submarket even though if he located in the low quality submarket he would not have to compete with any other buyers and could obtain a payoff of q_l with certainty. Thus if the partition of sellers into submarkets satisfies (EC) all buyers locating in the high quality submarket only is the equilibrium assignment of buyers.

It follows from (4), (7) and (EC) that, for any partition of sellers, the equilibrium assignment of buyers in a period is given by

$$\phi_l(\alpha, q_l) = \begin{cases} 1-(1-\alpha)(\ln q_h(\alpha, q_l) - \ln q_l) & \text{if } q_h(\alpha, q_l)e^{-\frac{1}{1-\alpha}} \leq q_l \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Suppose that sellers in the bad submarket have strictly lower expected quality products than sellers in the good submarket, that is, $0 \leq q_l < \widetilde{q} < q_h \leq 1$. In this case, equations (5) and (8) imply

$$0 \leq \phi_l < 1 < \phi_h$$

Therefore, buyers visit each seller in the bad submarket less frequently than they visit each seller in the good submarket. In other words, search is directed.[9]

3.3 Welfare effects

The decentralized actions of buyers in response to the creation of quality differentiated submarkets raises a question of whether submarket creation is socially efficient. For example, if the inequality (EC) is satisfied, then there is increased competition between buyers for the remaining high quality sellers, and low quality sellers do not get to trade at all. Likewise, if low quality sellers are included, then it is not clear that the distribution of buyers over submarkets will be optimal.

To address this issue, let W denote the total welfare in a period for a given information partition. The maximum social welfare of any possible assignment of buyers to the two submarkets defined by an information partition (α, q_l) is given by

$$W(\alpha, q_l) = \max_{\phi_l \geq 0} \left\{ \alpha\left(1-e^{-\phi_l}\right) q_l + (1-\alpha)\left(1-e^{-\phi_h(\alpha,\phi_l)}\right) q_h(\alpha, q_l) \right\} \tag{9}$$

Proposition 4 *The solution to (9) is given by (8). That is, the decentralized equilibrium is constrained efficient.*

[9] Julien, Kennes and King (2000), Coles and Eeckhout (2003) and Shimer (2004) also develop models of directed search.

Proof. Substituting for $\phi_h(\alpha,\phi_l)$ and differentiating yields the first-order condition:

$$\frac{\partial W}{\partial \phi_l} = e^{-\phi_l} q_l - e^{-\frac{1-\alpha\phi_l}{1-\alpha}} q_h(\alpha, q_l) = 0$$

Solving for ϕ_l yields $\phi_l = 1 + (1-\alpha)\left[\ln q_l - \ln q_h(\alpha, q_l)\right]$, which satisfies $\phi_l \geq 0$, as long as $q_h(\alpha, q_l) e^{-\frac{1}{1-\alpha}} \leq q_l$. Otherwise, $\phi_l = 0$. ■

The following proposition describes how a change in the partition of sellers affects social welfare.

Proposition 5 ***(Efficient Partitioning)*** *Welfare increases if, in the bad submarket, (i) the number of sellers with low quality products increases or (ii) the number of good sellers, prior to the realization of their product type, decreases*

Proof. Let $\{q_l, \alpha\} = \{(x\theta + y\widehat{q}) / (x+y) \leq \widetilde{q}, x+y\}$ where x is the quantity of sellers that have realized a low quality product and y is the number of good sellers that have yet to realize product quality, in the bad submarket. Comparative statics on (9) yield $\partial W/\partial x > 0, \partial W/y < 0$. ■

Note that the welfare level in a period in the absence of submarkets, given by (6), can also be written as $W_0 = 2W(1,\widetilde{q})$. From proposition 5 it is straightforward to see that the creation of submarkets always serves to increase welfare, that is, $W(\alpha, q_l) > W(1, \widetilde{q})$ for $\alpha < 1$ and $q_l < \widetilde{q}$.

3.4 The equilibrium under a reputation system that screens for type

If the reputation system screens for type, a seller that sells a low quality product in the first period is given a bad reputation in the second period with probability k. In the first period, advertising will not affect a seller's second period payoffs – thus advertisements will have no informational content. Therefore, in the first period there are no submarkets created and the total value of trade between buyers and sellers is given by

$$W(1,\widetilde{q}) = W_0/2.$$

In the second period, the quantity of sellers in the bad submarket is

$$\alpha^2 = \tfrac{1}{2}k + \tfrac{1}{2}k\,(1-\gamma)\,,$$

where $\frac{1}{2}k$ is the number of bad sellers caught selling low quality products in the first period and $\frac{1}{2}k\,(1-\gamma)$ is the number of good sellers similarly caught. The average quality of sellers with bad reputations is thus

$$q_l^2 = \frac{\frac{1}{2}k\theta + \frac{1}{2}k\,(1-\gamma)\,\widehat{q}}{\alpha^2}.$$

The ratio of bad sellers to good sellers with bad reputations is greater than the ratio of bad sellers to good sellers for the whole economy, i.e. $1/(1-\gamma) > 1$.

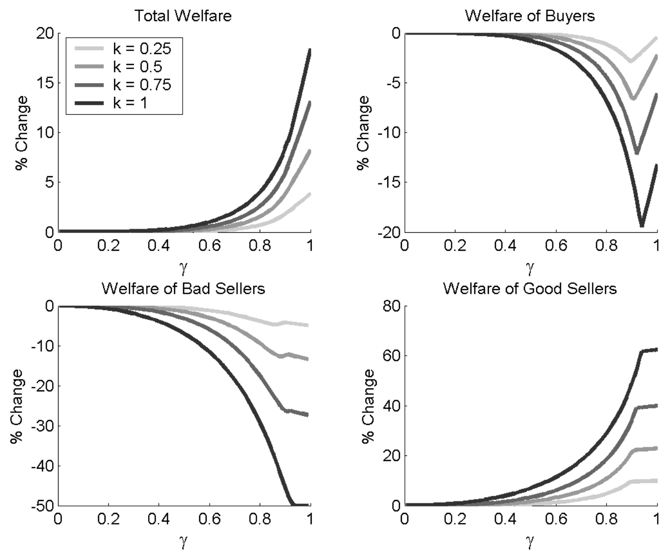

Figure 1: The value of a reputation system that screens for type, fixing $\widetilde{q} = \frac{1}{2}$.

Therefore, the average quality of products sold by sellers with good reputations is always greater that the average quality of products sold by sellers with bad reputations in period 2 under a reputation system that screens for type.

The total welfare of all transactions in periods 1 and 2 under this reputation system is given by

$$W^T = W_0/2 + W(\alpha^2, q_l^2) \tag{10}$$

Figure 1 compares the welfare of this reputation system, given by (10) with the welfare of unguided search, given by (6). It shows the percentage change in total welfare and welfare of buyers and sellers going from unguided search to a reputation system that screens for type. In this example we maintain a constant average quality of sellers and fix $\widetilde{q} = \frac{1}{2}$. On the horizontal axis we vary γ, and thus for each γ we set θ to satisfy (1).

If $\gamma = 0$, then good and bad sellers are identical and thus the reputation system cannot create any additional welfare, i.e. $W(\alpha^2, q_l^2) = W_0/2$. The greatest possible benefit of the reputation system that screens for type occurs if $\gamma = 1$, which corresponds to the largest possible difference in expected quality between good and bad sellers. The value of the reputation system that screens for type is also increasing in k.

From figure 1 we can see that all of the gains accrue to good sellers, while buyers and bad sellers are made worse off. Bad sellers are obviously worse off because they are now differentiated (with some probability) from good sellers in the second period. Buyers are worse off because although they have some information about seller types in the second period, they end up competing more intensely in the good submarket, which drives up the prices at these sellers.

3.5 The equilibrium under a reputation systems that screens for honesty

A seller who is *honest* truthfully advertises their product quality in period 1. If the reputation system screens for honesty, a first period liar gets a bad reputation in the second period with probability k. For obvious reasons, a good seller with a high quality product realization will always advertise high quality and thus always tells the truth. Therefore we only need to solve the decision problems for the bad type sellers and the good type sellers with a low quality realization in period 1.

3.5.1 Truthful sellers

We allow for symmetric mixed strategies by sellers and use ξ_b and ξ_g to respectively denote the probabilities that a bad seller and a good seller with a low quality realization are honest in period 1. We assume

- **(A1)** If $\xi_g = 0$ and $\xi_b = 0$, then $q_l^1 = \theta$

and

- **(A2)** If $\xi_g = 1$ and $\xi_b = 1$, then $q_l^2 = \theta$.

These two assumptions concern the cases where all sellers lie, or all sellers are honest. In both cases, there exists a period with only one 'submarket'. The first case leaves no sellers in the bad submarket in period 1. In this case, A1 implies that if a seller deviated and advertised bad quality, they would be believed to have quality level θ in the first period. The second case leaves no sellers in the bad submarket in period 2. In this case, A2 implies that if a seller deviated and was dishonest and obtained a bad reputation, they would also be believed to have low quality in period 2.

Other than the cases covered by A1 and A2 the probabilities ξ_g and ξ_b give two well-defined submarkets in each period. The probabilities ξ_g and ξ_b directly determine the total fraction of all sellers who advertise low quality in period one. Therefore,

$$\alpha^1 = \tfrac{1}{2}\left(\xi_b + (1-\gamma)\,\xi_g\right).$$

Only sellers with low quality products ever advertise low quality in the first period. Thus

$$q_l^1 = \theta.$$

All untruthful sellers with low quality products in period 1 are caught with probability k. Therefore, the total fraction of all sellers with bad reputations in period two is given by

$$\alpha^2 = \frac{1}{2}\left[(1-\xi_b) + (1-\gamma)(1-\xi_g)\right] k.$$

The expected quality of sellers with bad reputations in the second period depends on the relative quantity of bad and good sellers in this submarket. We have

$$q_l^2 = \frac{(1-\xi_b)\frac{k}{2}}{\alpha^2}\theta + \frac{(1-\xi_g)(1-\gamma)\frac{k}{2}}{\alpha^2}\widehat{q}$$

where $(1-\xi_b)\frac{k}{2}/\alpha^2$ is the fraction of sellers with bad reputation that are bad type and $(1-\xi_g)(1-\gamma)\frac{k}{2}/\alpha^2$ is the fraction of sellers with bad reputations that are good type.

3.5.2 Honesty valuations

We are particularly interested in whether some sellers (either good or bad type) who get a low quality realization in the first period will want to advertise this truthfully so as to gain a good reputation for the second period. For this purpose we define the 'honesty valuation', π_g, of a good seller with a bad realization to be the difference between the expected payoff to that seller from being honest and lying, that is,

$$\begin{aligned}\pi_g &= p\left(\phi_l^1\right)\theta + p\left(\phi_h^2\left(\alpha^2,\phi_l^2\right)\right)\widehat{q} \\ &\quad - \left[p\left(\phi_h^1\left(\alpha^1,\phi_l^1\right)\right)\theta + \left(kp\left(\phi_l^2\right) + (1-k)\,p\left(\phi_h^2\left(\alpha^2,\phi_l^2\right)\right)\right)\widehat{q}\right]\end{aligned}$$

which simplifies to

$$\pi_g = \left[p\left(\phi_l^1\right) - p\left(\phi_h^1\left(\alpha^1,\phi_l^1\right)\right)\right]\theta + k\left[p\left(\phi_h^2\left(\alpha^2,\phi_l^2\right)\right) - p\left(\phi_l^2\right)\right]\widehat{q}. \tag{11}$$

Similarly, the honesty valuation of a bad seller is

$$\pi_b = \left[p\left(\phi_l^1\right) - p(\phi_h^1(\alpha^1,\phi_l^1))\right]\theta + k\left[p\left(\phi_h^2(\alpha^2,\phi_l^2)\right) - p\left(\phi_l^2\right)\right]\theta. \tag{12}$$

3.5.3 Properties of the equilibrium

This subsection considers the properties of the equilibrium under a reputation system that screens for honesty. It is easy to establish that an equilibrium exists by the standard Nash argument, because there is a well defined mapping of the two mixed strategies ξ_b, ξ_g into payoffs π_g, π_b. We can also establish:

Proposition 6 ***(Quality fosters honesty)*** *Good sellers are always at least as honest as bad sellers, that is, in equilibrium,* $\xi_g \geq \xi_b$.

Case	Behavior of sellers with low quality products	π_g	π_b
1	All good sellers are honest, all bad sellers lie.	$+$	$-$
2	All good sellers are honest, some bad sellers honest.	$+$	0
3	Some good sellers are honest, all bad sellers lie.	0	$-$
4	All sellers lie.	$-$	$-$

Table 1: Potential equilibria under a reputation system that screens for honesty

Proof. From (11) and (12) we have

$$\pi_g - \pi_b = k\left[p\left(\phi_h^2\left(\alpha^2, \phi_l^2\right)\right) - p\left(\phi_l^2\right)\right](\widehat{q} - \theta) > 0$$

since $p\left(\phi_h^2\left(\alpha^2, \phi_l^2\right)\right) > p\left(\phi_l^2\right)$ and $\widehat{q} > \theta$. ■

Proposition 7 ***(Dishonesty)*** *Some bad sellers always lie, that is, in equilibrium,* $0 \leq \xi_b < 1$.

Proof. From proposition 6 we know that $\xi_b \leq \xi_g$. It remains to show that when $\xi_g = 1$, in equilibrium we cannot have $\xi_b = \xi_g$. To construct a contradiction, suppose that $\xi_b = \xi_g = 1$, i.e., all sellers are honest. This yields $\alpha^1 = \frac{1}{2}(2-\gamma)$, $q_l^1 = \theta$, $q_h^1 = 1$, $\alpha^2 = 0$, $q_l^2 = \theta$ (by A2) and $q_h^2 = \widetilde{q}$. From (5) and (8) we have

$$\phi_h^1 - \phi_l^1 = \begin{cases} -\ln\theta & \text{for } \theta \geq e^{-\frac{2}{\gamma}} \\ \frac{2}{\gamma} & \text{otherwise} \end{cases}$$

and

$$\phi_h^2 - \phi_l^2 = \begin{cases} \ln\widetilde{q} - \ln\theta & \text{for } \theta \geq \widetilde{q}e^{-1} \\ 1 & \text{otherwise} \end{cases}$$

Thus,

$$\Delta = \left(\phi_h^1 - \phi_l^1\right) - \left(\phi_h^2 - \phi_l^2\right) = \begin{cases} \frac{2}{\gamma} - 1 & \text{for } 0 \leq \theta \leq e^{-\frac{2}{\gamma}} \\ -\ln\theta - 1 & \text{for } e^{-\frac{2}{\gamma}} \leq \theta \leq \widetilde{q}e^{-1} \\ -\ln\widetilde{q} & \text{for } \widetilde{q}e^{-1} \leq \theta < 1 \end{cases}$$

The first and third cases are unambiguously positive regardless of θ. The second case is positive if $\theta \leq e^{-1}$, which is true if $\theta \leq \widetilde{q}e^{-1}$ since $\widetilde{q} < 1$. Thus Δ is always positive. Since $p(\cdot)$ is a strictly increasing function, from (12) we have $\pi_b > 0$ regardless of the value of k. Thus a bad seller can gain by deviating and being dishonest, so all sellers being honest cannot be an equilibrium. ■

Propositions 6 and 7 give four possible equilibrium configurations, which are characterized in Table 1.

With the above results and assumptions in mind, we can analyze how a reputation system that screens for honesty can foster equilibrium signalling by good sellers who occasionally have low quality products. The welfare of the reputation system that screens for honesty is given by

$$W^H = W\left(\alpha^1, q_l^1\right) + W\left(\alpha^2, q_l^2\right) \tag{13}$$

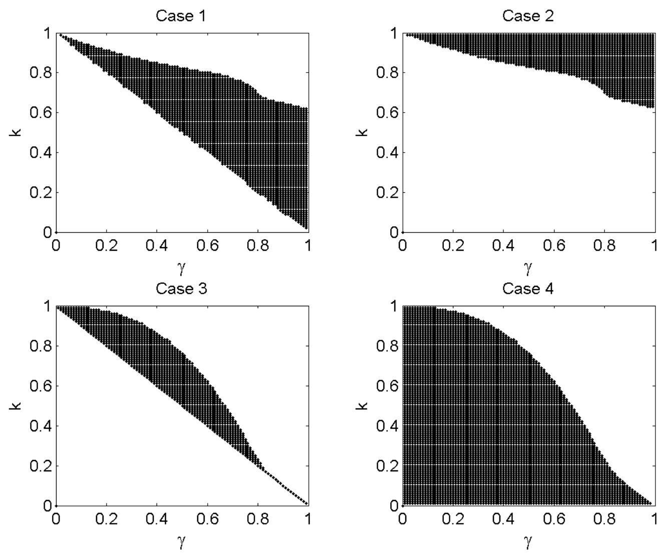

Figure 2: Equilibria if the reputation system screens for honesty, fixing $\widetilde{q} = \frac{1}{2}$.

where α^1, q_l^1, α^2 and q_l^2 are determined from the appropriate equations above.

To analyze each of the cases in Table 1, we used a numerical algorithm programmed in *Matlab* to test which type(s) of equilibrium occurred for any given set of parameter values.[10] As a benchmark, again fix $\widetilde{q} = \frac{1}{2}$ and vary γ while solving θ to satisfy (1). Figure 2 shows the existence of the different possible types of equilibria varying γ (the probability that a good seller has a high quality product) on the horizontal axis and k (the probability that a lying seller is assigned a bad reputation) on the vertical axis.

From figure 2 we can see that there is a large parameter region in which an equilibrium exists where all sellers lie (case 4). Part of this region overlaps with cases 1 and 3 in which there is full or partial honesty by good sellers, while all bad sellers continue to lie. If the probability of being assigned a bad reputation is high enough, then some bad sellers may also start to be honest (case 2). In general, the likelihood that the reputation system generates some honesty is increasing in both γ and k.

From figure 2 we can see that for some parameter values there are multiple equilibria. Figure 3 shows the equilibrium at each point that gives the highest overall social welfare out of all the possible equilibria at that point.

[10] Source code for the simulation program is available from the authors on request.

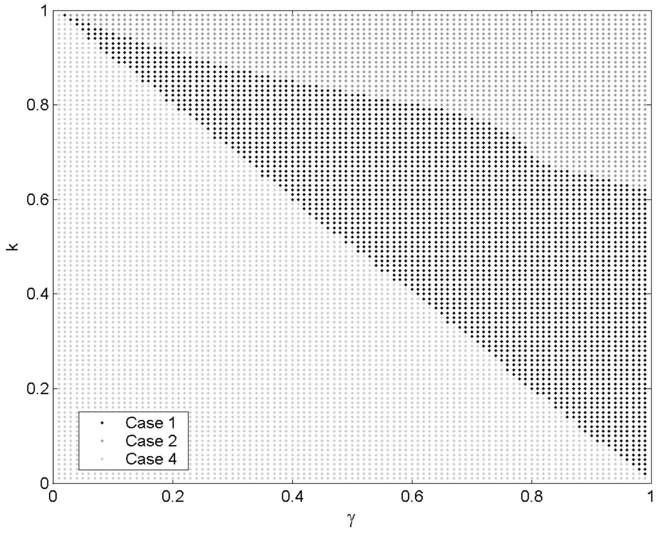

Figure 3: Equilibrium with the highest welfare level under a reputation system that screens for honesty, fixing $\widetilde{q} = \frac{1}{2}$.

4 Comparing Reputation Systems

When does a reputation system that screens for honesty perform better than a reputation system that screens for type? The welfare of the reputation system for honesty is given by (13) and the welfare of the reputation system for type is given by (10). The following proposition gives a sufficient condition for the dominance of a reputation system that screens for honesty.

Proposition 8 ***(Non-equivalence)*** *A sufficient condition for the superiority of a reputation system that screens for honesty ($W^H > W^T$) is some good sellers with low quality products advertise truthfully and all bad sellers lie about the quality of their products, that is, if in equilibrium $\xi_g > 0$ and $\xi_b = 0$.*

Proof. There is some separation of sellers into submarkets (i.e. $\alpha > 0, q_l < q_h$) in the first period only if the reputation system screens for honesty. Thus by proposition (5) the reputation system that screens for honesty outperforms the reputation system that screens for type in period 1. Furthermore, in period 2 there will be fewer good sellers with bad reputations and the same number of bad sellers with bad reputations if the reputation system screens for honesty. Therefore, proposition 5 also implies that welfare is higher in period 2 if the reputation system that screens for honesty. ■

Stated differently, a necessary condition for the superiority of a reputation system that screens for type ($W^T > W^H$) is that some bad sellers are truthful about the quality of their inferior product. Therefore a reputation system that screens for type is superior only if there is a problem of excessive honesty under a reputation system that screens for honesty.

When are the two reputation systems equivalent? The following proposition establishes a sufficient condition for the equivalence of the two reputation systems.

Proposition 9 ***(Equivalence)*** *A sufficient condition for the equivalence of the two reputation systems ($W^T = W^H$) is that all sellers with low quality products advertise untruthfully, that is, if in equilibrium $\xi_g = \xi_b = 0$.*

Proof. The first period in both cases is unguided with total welfare given by $W_0/2$. In the second period, the number of sellers with bad reputations is identical because the set of sellers that are untruthful is equivalent to the set of sellers with bad product. Thus $W(\alpha^2, q_l^2)$ is the same under both reputation systems. ■

Using the same parameters as figures 1 and 2, figure 4 gives the regions where each reputation system dominates. The numerical results confirm the result that the only case where the reputation system for type can dominate the reputation system for honesty is the region where some bad sellers are honest in period 1. The results can be related to the issue of type 1 and 2 errors in statistics. The basic idea is that a reputation system that screens for type reduces type 2 error – it avoids labelling someone as innocent when they are guilty – while the reputation system that screens for honesty reduces type 1 error – it can avoid labelling someone as guilty when they are innocent.

From figure 4 we can also see that for the honesty system to be superior, we must have k and/or γ high enough. The region in which the honesty system dominates corresponds to the region in which an equilibrium exists under this system where there is full or partial honesty by some sellers. This confirms our results that a necessary (but not sufficient) condition for the honesty system to dominate is that it generate some honesty among sellers.

5 Conclusion

In this paper we developed a simple model with search frictions and asymmetric information between buyers and sellers. In equilibrium, there exists a trade-off between two simple reputation systems. Therefore, this model can easily explain some observed differences about these institutions. For example, the model can explain why a professional association cares much about the codes of conduct of its members, thus always rewarding honesty, while a guidebook will often ignore what is said in the advertisements of sellers.

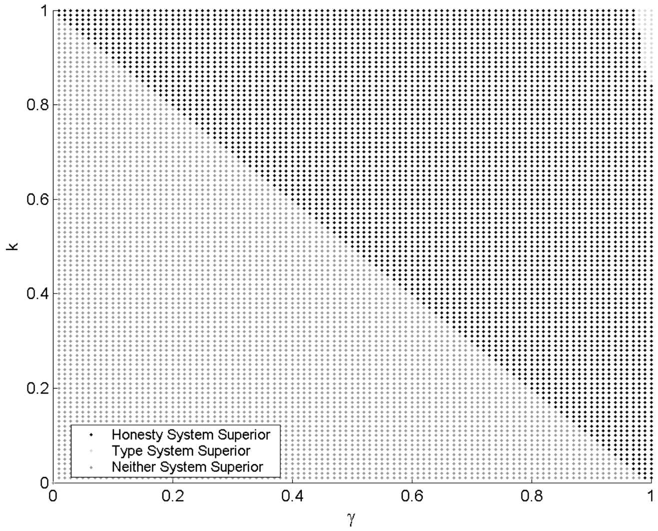

Figure 4: Choosing a reputation system, for $\widetilde{q} = \frac{1}{2}$.

Our model also gives some criteria by which to judge the performance of a reputation system. We found that the potential downside of a reputation system that screens for honesty is that bad sellers might overinvest in honesty to gain a good reputation in period two. The excessive honesty problem is exactly opposite to the problem of cheap pseudonyms in which a bad seller starts off with a good reputation and then chooses to start again when the reputation is reduced. Therefore, the solution to the cheap pseudonym problem – choosing a sufficiently low reputation starting point – has no bearing on the choice between the two simple reputation systems, because the problem of excessive honesty works in the opposite direction.

Another cheap way to establish a good reputation under a reputation system that screens for honesty is to acquire and truthfully sell a number of low cost products. We have ruled out this strategy by assumption. However, if we did allow both types of sellers to pursue this strategy, then we could expect to see reputation inflation, but it is unclear how these choices affect the relative performance of the two reputation systems. In either case, our model makes clear that it is a problem of excessive honesty, not excessive product choice, which is central to the trade-off between reputation systems.

There are a number of directions for further research. It might be interesting to extend our two period model to one with a longer planning horizon. For example, a multi-period model could illustrate factors that determine when a bad seller chooses to cash in his/her reputation. Another possible theoretical extension is to consider the sale of third party information services. One method to sell third information is an accreditation service that sells reputation services to its members (i.e. the sellers in our model). Another method of selling third party information is a guidebook that sells information about sellers directly to buyers. It would be of interest to discover whether there is a connection between the type of reputation system used and the method by which a third party sells its information.

Finally, there are several reasons why our model may find some use in experimental economics. One reason is that the assumptions of our model are somewhat more realistic than the assumptions of alternative models. Realism in our model is supported by our assumption of endogenous matching - buyers chose over which sellers to search - and by our assumption of endogenous price formation - buyers bid subject to the presence or absence of local market competitors. A second reason is that our model has a finite horizon and so it can be implemented in a lab where playing time is obviously a constraint. The final reason for using our model in experiments is its simple trading structure, which could be easily communicated to participants in a laboratory setting.

References

[1] Akerlof, G. (1970). The market for 'lemons': quality uncertainty and the market mechanism" *Quarterly Journal of Economics*, **84**: 488-500.

[2] Coles, M. and J. Eechout (2004). Indeterminacy and directed search. *Journal of Economic Theory*, forthcoming.

[3] Dellarocas, C. (2003). The digitalization of word of mouth: Promises and challenges of online reputation mechanisms. *Management Science*, **49**: 1407-25.

[4] Diamond, D. (1989). Reputation acquisition in debt markets. *Journal of Political Economy*, **97**: 828-62.

[5] Ellison, G. (1994). Cooperation in the prisoner's dilemma with anonymous random matching. *Review of Economic Studies*, **61**: 567-88.

[6] Friedman, E. and P. Resnick (2001). The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy*, **10**: 173-99.

[7] Julien, B., J. Kennes and I. King (2000). Bidding for labor. *Review of Economic Dynamics*, **3**: 619-49.

[8] Kennes, J. (2004) Competitive Auctions: Theory and Application. *University of Copenhagen discussion paper* 04-16.

[9] Kreps, D. and R. Wilson (1982). Reputation and imperfect information. *Journal of Economic Theory*, **27**: 253-79.

[10] McAfee, R. P. (1993). Mechanism design by competing sellers. *Econometrica*, **61**: 1281-1312.

[11] McAfee, R. P. (2004) The Real Lesson of Enron's Implosion: Market Makers Are In the Trust Business, *The Economists' Voice*, **1**, Issue 2.

[12] Rothschild M. and J. E. Stiglitz (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, **90**: 629-49.

[13] Shimer, R. (2004). The assignment of workers to jobs in an economy with coordination frictions. *Mimeo*, University of Chicago.

[14] Spence, M. (1973). Job market signalling. *Quarterly Journal of Economics*, **87**: 355-74.

[15] Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies*, **63**: 1-22.

[16] Zacharia, G. and P. Maes (2000). Trust management through reputation systems. *Applied Artificial Intelligence*, **14**: 881 – 907.