DISCUSSION PAPERS
Department of Economics
University of Copenhagen

Simple Reputation Systems

John Kennes
Aaron Schiff

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

This version: November 4, 2005
First version: August 1, 2002

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 patterns 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 in some cases the same reputation system is prone to strategic manipulation by sellers who always have low quality products. 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 markets where, due to the nature of

the trading environment, it is relatively easy for sellers to mislead buyers and difficult for buyers to assess the quality of products for sale. In such markets, what we call 'reputation systems' often exist to guide buyers in their search.

For the purposes of this paper, a reputation system is a third party that collects and publishes information about historic seller behavior in a 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.

The Internet in particular has lead to the emergence of a number of on-line reputation systems.[1] This is likely due to the lower costs associated with collecting, analyzing, and disseminating reputational information in electronic form. This development has led to theoretical and empirical research to examine the algorithms used to assign reputations, not only by economists, but also by sociologists and computer scientists. However, much of this work rests on either behavioral assumptions, or in the case of economists, is driven by experimental research,[2] with the implicit assumption being one of bounded rationality.[3] The simple finite horizon model that we use in this paper has fully rational agents.

Our first goal in this paper is to understand how the behavior of buyers and sellers in a search market with asymmetric information is affected by the presence of simple reputation systems, and the consequent welfare effects. Our second goal is to compare different simple algorithms for assigning reputations and determine their relative efficiency.

To achieve these goals we use a two-period general equilibrium directed search model where goods are sold through competing auctions. Sellers are heterogeneous and there is asymmetric information between buyers and sellers. We assume there are two types of sellers, 'good' and 'bad', and two product quality levels, 'high' and 'low'. Bad sellers always have low quality for sale, while good sellers sometimes have high quality and sometimes have low quality. Asymmetric information exacerbates the problem of search frictions. Our model shows how reputation systems reduce the additional frictions caused by asymmetric information.

The reputation system in our model is a third party that observes the market in the first period and publishes the information it gathers to all buyers in the second period in as seller reputations. In the first period sellers can advertise whether they have a high or low quality product for sale. The reputation system observes the advertisements and actual product qualities of all sellers, but does

---

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

[2] See, for example, Bolton *et al.* (2004).

[3] 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.

not observe whether any seller is of the good or the bad type.

We focus on reputational algorithms that use a simple 0,1 (good/bad) metric to represent a seller's reputation. We consider two simple algorithms for the assignment of reputations in the second period. The first algorithm assigns a seller's reputation by considering only the product quality that they had for sale in the first period. We call this *reputation for type* and under this system all sellers who had low quality for sale in the first period get a bad reputation, while sellers who had high quality get a good reputation. We call the second algorithm *reputation for honesty*. Under this system a seller's advertisement is compared with the actual product quality that they had for sale in the first period. All sellers who lied about their product quality get a bad reputation in the second period and all sellers who were honest get a good reputation.

In the context of this model we examine the effects on the equilibrium search patterns of buyers of the provision of such information about sellers. We show that this always increases welfare relative to the case where buyers have no information about sellers because buyer search is directed more accurately. However, there are distributional effects arising from the fact that in equilibrium buyers may choose not to trade with sellers identified as having lower quality, and that buyers themselves may be worse off if the provision of information results in intense competition for high quality products.

We also examine the relative efficiency of the two different reputation systems described above. There are two key differences between the systems. First, the honesty system links a seller's advertisement in the first period with their reputation in the second period. It therefore introduces a mechanism by which a seller's advertisement can be a credible signal of their product quality in the first period. If such signalling arises in equilibrium, the honesty system can create information in both periods. In contrast, the type system only creates information in the second period, as it links reputations to only the product quality that a seller had for sale in the first period, which is not a strategic variable for a seller.

The second key difference relates to the ability of the systems to sort sellers appropriately in the second period. Under the type system, a good seller will receive a bad reputation if they happened to have a low quality product in the first period, even though there is some probability that they will have high quality in the second period. On the other hand, the honesty system is prone to strategic manipulation by bad type sellers, as they can get a good reputation by simply being honest even though they always have low quality for sale.

The main implication of these differences between the two systems is that there is a tradeoff between reputation algorithms. For most of the parameter values in our model, the honesty system does a better job of solving the asymmetric information problem than the type system in terms of total wel-

fare. However, if the probability that a good seller has a high quality product is large enough, the equilibrium under the honesty system involves a high degree of manipulation by bad type sellers. In such situations we show that the type system is superior.

This paper extends existing research on the problem of asymmetric information. There has been much work on this topic since the seminal contributions of Akerlof (1970), Spence (1973) and Rothschild and Stiglitz (1976). We provide a model of how a third-party reputation system can ameliorate asymmetric information problems in search markets. Alternative solutions to similar problems are addressed by Biglaiser (1993), and Li (1998) who examine the role of middlemen, and Williamson and Wright (1994) who consider the role of money. In our model we explicitly rule out long-term relationships between buyers and sellers. 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 compare alternative reputation systems.

The paper is organized as follows. In the next section we present a simple two-period model of seller advertisements, buyer search, and reputational assignment. We then solve for the equilibrium of this model sequentially, considering: (i) buyers' search patterns, (ii) the partition of sellers under each reputation system, and (iii) the choice of reputation system. The equilibrium properties of the model are analyzed, and we compare the efficiency of the two reputation systems under. 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 are sufficiently large so the set of each type of agent can be treated as a continuum.

2.1   Sellers and products

Sellers offer products for sale that have one of two quality levels: *high* and *low*. We normalize the high quality level to 1 and let $0 \leq \theta < 1$ denote the relative level of the low quality product.

The sellers are divided into two equal groups, which are distinguished by types: *good* and *bad*. Bad sellers have a single unit of the low quality product for sale in each period. Good sellers have a single unit of the high quality product for sale with probability $0 < \gamma \leq 1$ in each period, otherwise they have

a low quality product. Let $\widehat{q} = \gamma + (1 - \gamma)\theta$ denote the expected quality of a good seller's product in each period, and the average quality of all products for sale in each period is

$$\widetilde{q} = \tfrac{1}{2}\left(\widehat{q} + \theta\right). \tag{1}$$

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 first-price auction.[4] 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. 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 product's quality, before bidding commences. Buyers are identical in their willingness to pay for quality and the net utility function of a buyer 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.

Buyer i maximizes utility by bidding zero if they are the only buyer at the auction and bidding q_j^t if there is at least one other bidder. Therefore, a seller receives a non-zero price if and only if more than one buyer turns up to the auction. Effectively we assume Bertrand competition among the buyers located at any given seller.

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 in each period and never purchase from the same seller twice, that is, we rule out long-term relationships.

2.3 Submarkets

For our purposes, a 'submarket' refers to a group of sellers that appear identical from a buyer's point of view. In each period the sellers may be separated into two quality differentiated submarkets – by their advertisements in the first

---

[4]McAfee (1993) offers an early model of competing auctions. Kennes (2004) offers a survey of recent research. Efficiency is a well-known property of competing auction models (Julien, Kennes and King, 2000).

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 across the submarkets cannot change, thus

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

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, the expected quality in the high quality submarket is

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

By observing advertisements in the first period and reputations in the second period, buyers are informed of q_l^t, q_h^t and α^t and simultaneously choose their search patterns in each period. 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 of sellers, so market tightness in the submarkets 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 also 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 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 set of submarkets potentially created by the advertisements and reputations of sellers. It can be shown that, in a

---

[5]Second period advertisements have no informational content due to a lack of credibility caused by the finite horizon of the game.

(sub)market with tightness ϕ, the probability that more than one buyer turns up to any given seller's auction is given by:[6]

$$p(\phi) = 1 - e^{-\phi} - \phi e^{-\phi}.$$

The expected profit of a seller with product quality q_j in a submarket of tightness ϕ is then given by $V_j = p(\phi) q_j$.

What matters to a buyer is whether or not they are alone at a seller's auction, since they only get a strictly positive surplus from the auction in that case. The probability that a buyer is alone at any given seller in a (sub)market with market tightness ϕ can be shown to be $e^{-\phi}$ and the expected payoff of buyer i of visiting a seller in a submarket with tightness ϕ and expected quality q_z is given by $U_i = e^{-\phi} q_z$.

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 before choosing which seller's auction to visit. We define two different simple algorithms for assigning reputations:

Definition 2 *A reputation system* screens for honesty *if sellers that lied about their product quality in period 1 are assigned a bad reputation in period 2, and all sellers who were honest get a good reputation.*

Definition 3 *A reputation system* screens for type *if sellers that had a low quality product for sale in period 1 are assigned a bad reputation in period 2, and all sellers that had a high quality product for sale get a good reputation.*

A reputation system that screens for honesty thus compares the advertisements of sellers in period 1 with their actual product quality in that period. Any seller who had a different quality from what they advertised receives a bad reputation in period 2, and the honest sellers get a good reputation. In contrast, a reputation system that screens for type only looks at the actual quality that a seller had in period 1, and ignores their advertisement.

Note that the reputation system does not have access to any more information than buyers do, but it does have the ability to aggregate and publish this information.[7] The reputation system can 'see' the advertisements of all sellers

---

[6] See Kennes (2004) for details.

[7] Monitoring can be done in many different ways and to keep the analysis as simple and transparent as possible we assume perfect monitoring. In an earlier version of this paper we have shown that imperfect monitoring or assigning reputations based on reports of buyers do not change the main results of the paper.

and the quality of the goods that were actually available but it cannot determine whether any given seller is of the good type or the bad type. In addition, in this paper we limit our attention to simple mechanisms that do not impose explicit pecuniary punishments.[8]

Another possible reputation system is a combination of the above two systems whereby a seller's reputation informs buyers of both whether the seller was honest in the first period and the product quality that they had for sale. Such a system would provide some additional information to buyers, but complicates the analysis as it would potentially create at least three submarkets in the second period. One of the main goals of this paper is to examine the conditions under which honesty of sellers is a more important characteristic than the quality of goods that they sold. To focus on this tradeoff, we have chosen only to model simple reputation systems that report a single characteristic of sellers.

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, the reputation system assigns a reputation to each seller. Sellers draw a new product according to their type. Buyers observe the reputations 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 in this case is found backwards induction. In particular, we first evaluate the buyers' choice of seller to bid for given the submarkets that are created. We then consider the sellers' choice of advertisement given the equilibrium strategy of buyers and the particular reputation system under investigation.

---

[8] An alternative to a third-party reputation system is a third-party system that imposes explicit punishments on sellers in the second period depending on their behavior in the first period (for example a system of consumer protection laws). A possible avenue for future research is to compare the efficiency of such systems with our reputation systems.

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

3.1 Equilibrium with no submarkets

In the absence of quality differentiated submarkets, buyers randomize over the locations of all sellers. We call this 'unguided search'. 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

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

3.2 Equilibrium search patterns with submarkets

The division of sellers into quality differentiated submarkets in a period leads to two possible configurations for the equilibrium search patterns of buyers in that period. One possibility is that 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 visiting any seller.

The other possible type of equilibrium is that buyers visit only sellers in the good 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 locate in the high quality submarket in equilibrium.

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

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

and ϕ_h is given by (5).

If sellers in the bad submarket have strictly lower expected quality than sellers in the good submarket, that is, $0 \leq q_l < \widetilde{q} < q_h \leq 1$, then (5) and (8) imply $0 \leq \phi_l < 1 < \phi_h$. Therefore, buyers visit each seller in the bad submarket with lower probability than they visit each seller in the good submarket. In other words, search is directed by the creation of quality-differentiated submarkets.[9]

---

[9] Julien, Kennes and King (2000), Coles and Eeckhout (2003) and Shimer (2005) also develop models of directed search.

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 (EC) is satisfied, then there is increased competition between buyers for the remaining high quality sellers, while low quality sellers are unable to trade at all. Likewise, if low quality sellers are included, then it is not clear that the distribution of buyers over submarkets is optimal.

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

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

Proposition 1 *The decentralized search equilibrium is constrained efficient. That is, the solution to (9) is given by (8).*

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 $\phi_l = 1 - (1 - \alpha) \ln(q_h(\alpha, q_l)/q_l)$, which satisfies $\phi_l \geq 0$ as long as $q_h(\alpha, q_l) e^{-\frac{1}{1-\alpha}} \leq q_l$. Otherwise, $\phi_l = 0$. ∎

Proposition 1 confirms that the equilibrium search patterns of buyers maximize welfare within our framework. The following proposition describes how equilibrium welfare is affected by the 'informativeness' of the partition of sellers, and is a useful result for the comparison of reputation systems.

Proposition 2 *(Efficient Partitioning)* *Equilibrium welfare increases when:*

1. *The number of sellers in the bad submarket with low quality products increases; or*

2. *The number of good sellers (prior to the realization of their product quality) in the bad submarket decreases.*

Proof. Let $(\alpha, q_l) = (x + y, (x\theta + y\widehat{q})/(x + y))$ 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 $\partial W/\partial x > 0$ and $\partial W/\partial y < 0$. ∎

Proposition 2 says that the more informative the information partition, the greater the level of equilibrium welfare. This is because greater information leads to buyer search being directed more accurately, which reduces search frictions. Note that the welfare level in a period in the absence of submarket creation, given by (6), can also be written as $W_0 = 2W(1, \widetilde{q})$. From proposition 2 it is straightforward to see the following corollary.

Corollary 1 *The creation of quality-differentiated submarkets always increases equilibrium welfare. That is, $W(\alpha, q_l) > W(1, \widetilde{q})$ for $\alpha < 1$ and $q_l < \widetilde{q}$.*

This result means that if a reputation system creates quality-differentiated submarkets in one or both periods it will always increase total welfare.

3.4 Equilibrium with full information

As another benchmark, we will compare the equilibrium welfare under a reputation system to a situation of 'full information' in which all buyers know exactly which sellers have high and low quality products for sale in each period. In this case we have $\alpha = \frac{1}{2} + \frac{1}{2}(1 - \gamma)$, $q_l = \theta$ and $q_h = 1$ in both periods. From (EC), sellers with low quality products are excluded if $e^{-\frac{2}{\gamma}} \geq \theta$. From (3), (8) and (9), equilibrium welfare (over both periods) is given by:

$$W_F = \begin{cases} 2\left(\widetilde{q} - e^{-1}\theta^{1-\frac{\gamma}{2}}\right) & \text{if } e^{-\frac{2}{\gamma}} \leq \theta \\ \gamma\left(1 - e^{-\frac{2}{\gamma}}\right) & \text{otherwise} \end{cases}. \tag{10}$$

In our model, the presence of asymmetric information exacerbates the welfare losses due to search frictions. Figure 1 compares the full information equilibrium welfare (given by (10)) with the equilibrium welfare when buyers have no information about seller qualities (given by (6)) and shows the fraction of full information equilibrium welfare that is lost due to asymmetric information.

3.5 Equilibrium with reputations for type

If the reputation system screens for type, a seller with a low quality product for sale in the first period gets a bad reputation in the second period. 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 $W(1, \widetilde{q}) = \frac{1}{2}W_0$.

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

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

which comprises all bad sellers plus the good sellers who had a low quality realization in the first period. The expected quality of sellers with bad reputations is thus

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

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

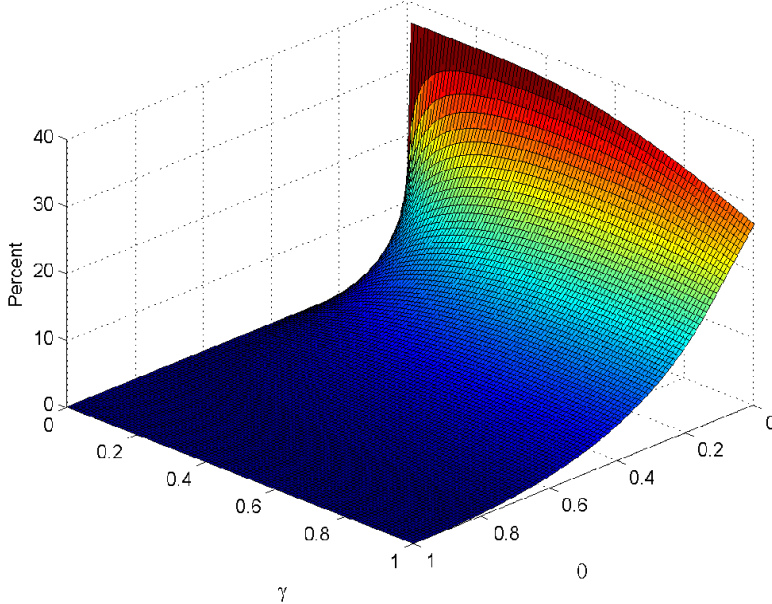$$W^T = \tfrac{1}{2}W_0 + W(\alpha^2, q_l^2). \tag{11}$$

Figure 1: Percentage welfare losses due to asymmetric information. Note the axes have been rotated.

To quantify the effects of introducing a reputation system that screens for type, Figure 2 compares the percentage change in welfare under this system, given by (11) with the equilibrium welfare when search is unguided, given by (6), as well as the effects on buyers and sellers. On the horizontal axes we vary the probability that a good seller has high quality in a period (γ) and show the results for three different low quality levels (θ).

Obviously, if $\gamma = 0$ then good and bad sellers are identical, and thus the reputation system cannot create any additional welfare. 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 is also increasing in the difference between high and low quality levels.

From Figure 2 we can also see that for the parameter values shown 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 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 in the second period, which
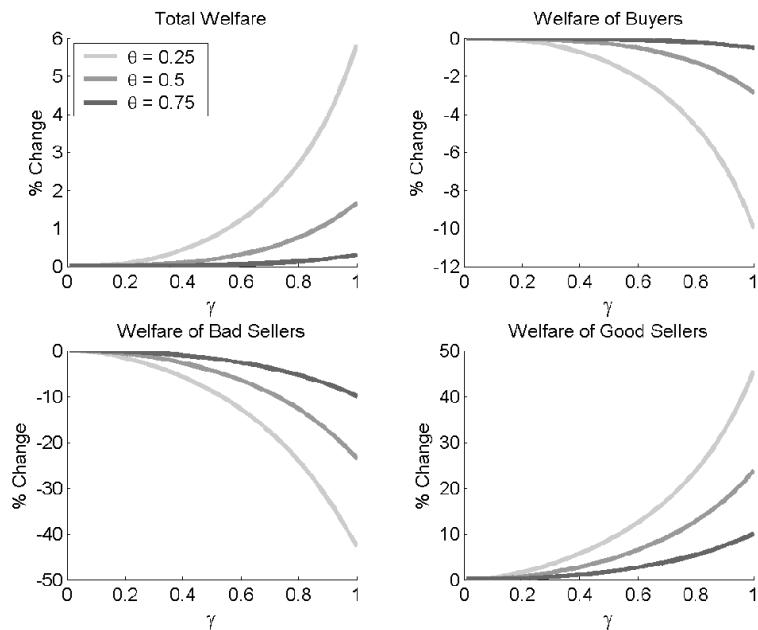
Figure 2: Equilibrium welfare effects of introducing a reputation system that screens for type.

drives up the prices at these sellers.[10]

3.6 Equilibrium with reputations for honesty

A reputation system for honesty links a seller's advertisement in the first period with their reputation in the second period. Unlike the reputation system for type, it offers a mechanism by which sellers may be able to reveal meaningful information in their advertisements, and thus has the ability to create submarkets in the first period as well as the second period.

A seller who is honest truthfully advertises their product quality in period 1. For obvious reasons, a good seller with a high quality realization will always advertise high quality and thus is always honest and always gets a good reputation. 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. These sellers must choose whether to truthfully advertise that they have a low

---

[10]In Kennes and Schiff (2004) we show that an arbitrary partition of sellers into two quality-differentiated submarkets makes buyers worse off relative to the equilibrium in the absence of such a partition if (i) the partition is not sufficiently informative so as to allow them to exclude sellers in the low quality submarket, or (ii) the overall ratio of buyers to sellers is sufficiently high.

quality product, or whether to untruthfully advertise high quality.

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. These are the bad seller and the good sellers with low quality realizations who have chosen to be honest. Thus,

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

Only sellers with low quality products advertise low quality in the first period, so

$$q_l^1 = \theta.$$

All untruthful sellers with low quality products in period 1 get bad reputations in period 2 and all truthful sellers get good reputations. Therefore, the total fraction of all sellers with bad reputations in period 2 is given by the bad sellers who lied in the first period plus the good sellers with low quality realizations who lied in the first period:

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

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:

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

We are particularly interested in whether some sellers with low quality for sale in the first period will want to advertise this truthfully so as to gain a

14

good reputation for the second period. For this purpose we define the 'honesty valuation', π_g, of a good seller with a low quality realization in the first period to be the difference between the expected payoff to that seller from being honest and lying, that is,

$$\pi_g = \left[p\left(\phi_l^1\right) - p\left(\phi_h^1\left(\alpha^1, \phi_l^1\right)\right) \right] \theta + \left[p\left(\phi_h^2\left(\alpha^2, \phi_l^2\right)\right) - p\left(\phi_l^2\right) \right] \widehat{q}. \qquad (12)$$

Similarly, the honesty valuation of a bad seller is

$$\pi_b = \left[p\left(\phi_l^1\right) - p(\phi_h^1(\alpha^1, \phi_l^1)) + p\left(\phi_h^2(\alpha^2, \phi_l^2)\right) - p\left(\phi_l^2\right) \right] \theta. \qquad (13)$$

We now consider the equilibrium behavior of sellers 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 the following results about the equilibrium behavior of sellers:

Proposition 3 (Quality fosters honesty) *Good sellers with low quality realizations in period 1 are always at least as honest as bad sellers, that is, in equilibrium, $\xi_g \geq \xi_b$.*

Proof. From (12) and (13) we have

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

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

Proof. From proposition 3 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,

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

Case	First-period 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 good sellers lie and all bad sellers lie.	$-$	$-$

Table 1: Potential equilibria under a reputation system that screens for honesty. Note this table only shows the first-period behavior of bad sellers and good sellers with low quality realizations. As mentioned above, good sellers with high quality realizations are always honest.

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 (13) we have $\pi_b > 0$. Thus a bad seller can gain by deviating and being dishonest, so all sellers being honest cannot be an equilibrium. ∎

Proposition 3 arises because good sellers have a higher expected product quality than bad sellers, which means that a good seller's gain from having a good reputation in the second period is always greater than that of a bad seller. Proposition 4 arises because if all sellers are honest in the first period, the difference in expected qualities of the first period submarkets is large, leading to large differences in first period market tightnesses, and a high payoff from being dishonest. In addition, in the second period the differences in market tightnesses are smaller, because all sellers have a good reputation and so reputations do not serve to distinguish sellers. This makes the 'punishment' that a bad seller suffers in the second period from being dishonest small relative to the gains in the first period.

Propositions 3 and 4 give four possible equilibrium configurations, which are characterized in Table 1. We used a numerical algorithm programmed in *Matlab* to test which type(s) of equilibrium as given in Table 1 occurred for any given set of parameter values.[11] The parameters of our model are γ and θ, which are both numbers between zero and one, thus it is possible to check the span of the parameter space numerically. We found that only equilibria in which all good sellers are honest and some bad sellers are honest (case 2) exist.[12] Thus in equilibrium under the honesty system we have $\xi_g = 1$ and $0 < \xi_b < 1$, with ξ_b being the solution to $\pi_b = 0$.

The equilibrium welfare of the reputation system that screens for honesty is

---

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

[12]In an earlier version of this paper we assumed that bad reputations were assigned only with some probability to sellers who lied in period 1. If this probability is strictly less than one, then equilibria of the other three types can also arise depending on the parameter values.
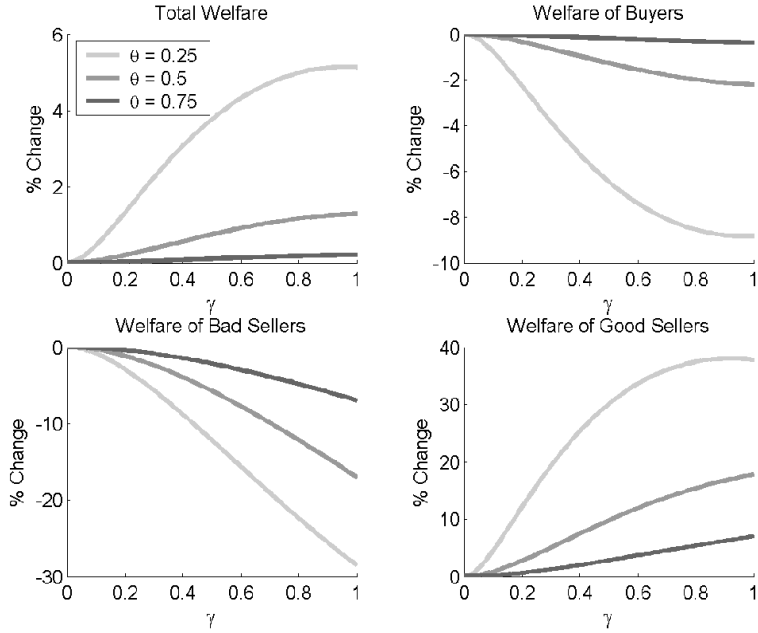
Figure 3: Equilibrium welfare effects of introducing a reputation system that screens for honesty.

given by

$$W^H = W\left(\alpha^1, q_l^1\right) + W\left(\alpha^2, q_l^2\right) \tag{14}$$

where α^1, q_l^1, α^2 and q_l^2 are determined from the appropriate equations above. Figure 3 shows the equilibrium effects on total welfare, buyers, and sellers of introducing a reputation system that screens for honesty relative to when buyers have no information about sellers. As in the case of a reputation system for type, for the parameter values shown the gains accrue to good sellers, with buyers and bad sellers being worse off.

4   Comparing Reputation Systems

When does a reputation system that screens for honesty perform better than a reputation system that screens for type? An honesty system is able to create information in both periods, but allows bad sellers to get a good reputation by being honest in the first period. A type system is immune to strategic manipulation by bad sellers, but creates information only in the second period, and also assigns bad reputations to good sellers who happened to have low quality for sale in the first period. In this section we examine the welfare implications of these tradeoffs.

The welfare of the reputation system for honesty is given by (14) and the welfare of the reputation system for type is given by (11). The following proposition gives a sufficient condition for the dominance of a reputation system that screens for honesty.

Proposition 5 *A sufficient condition for a reputation system that screens for honesty to be superior to a reputation system that screens for type ($W^H > W^T$) is if in the first period some good sellers with low quality products are honest and all bad sellers lie, that is, if $\xi_g > 0$ and $\xi_b = 0$ in equilibrium.*

Proof. If $\xi_g > 0$ and $\xi_b = 0$ then there is some separation of sellers into submarkets in the first period (i.e. $\alpha^1 > 0, q_l^1 < q_h^1$) under the honesty system, but there is never first period separation under the type system. Thus by proposition (2) the reputation system that screens for honesty outperforms the reputation system that screens for type in period 1.

In period 2, under the type system all sellers who sold low quality in period 1 will get a bad reputation. This will include all bad sellers and a fraction $1 - \gamma$ of good sellers. Under the honesty system, since $\xi_g > 0$ and $\xi_b = 0$, those with bad reputations will include all bad sellers and a fraction $(1 - \xi_g)(1 - \gamma)$ of good sellers. Therefore, proposition 2 also implies that welfare is higher in period 2 if the reputation system 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 having low quality. 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. As mentioned in section 3.6 above, the type of equilibrium that arises under the honesty system has partial honesty by bad sellers. Depending on the parameters, it is therefore possible that the type system could be superior to the honesty system.[13]

Figure 4 shows which reputation system is superior for the range of possible parameter values in our model. It can be seen that the type system can be superior to the honesty system if the probability that good sellers have high quality products is high enough. Under such parameter values, the equilibrium honesty level of bad sellers is so high that the honesty system performs poorly, while the type system performs relatively well. In such cases the strategic manipulation

---

[13]It is also straightforward to see that the two reputation systems will generate the same amount of equilibrium welfare if the honesty system does not induce any honesty in the first period. This is because if all sellers lie, there are no submarkets created in the first period under the honesty system. In addition, all sellers who sold low quality in the first period were liars, so all will receive a bad reputation under an honesty system. The honesty system therefore creates the same submarkets in the second period as the type system. In our model, an equilibrium in which all sellers lie (case 4 in table 1) never arises, so the two systems are never equivalent. As mentioned previously, if reputations are imperfectly assigned, then other equilibria, including case 4, are possible.
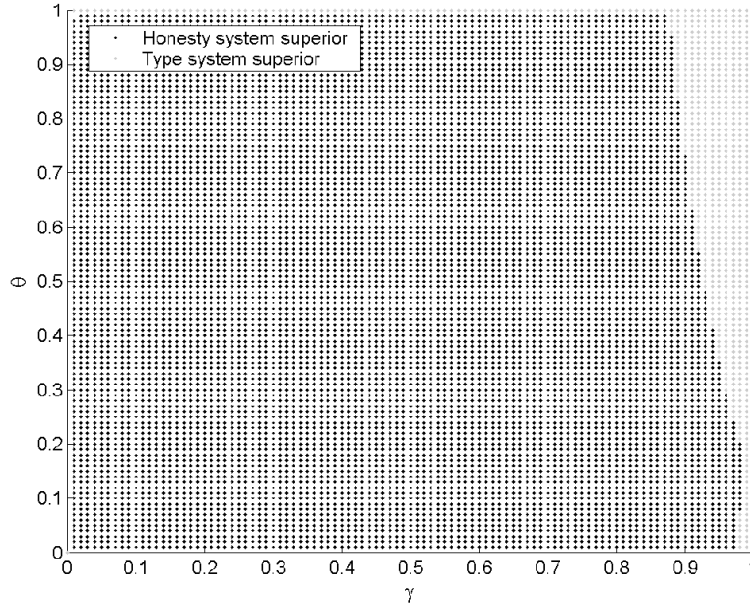
Figure 4: Welfare comparison of reputation systems.

of the honesty system by bad sellers causes negative effects that outweigh the value of the creation of submarkets in the first period, rendering the type system superior.

These results can also be related to the issue of type 1 and 2 errors in statistics. 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.

As a final comparison, figure 5 shows the fraction of the full information welfare level (given by (10)) that is achieved in equilibrium under both reputation systems.

5   Conclusion

In this paper we developed a simple model of reputation systems in markets with search frictions and asymmetric information between buyers and sellers. In equilibrium, there exists a trade-off between the two simple reputation systems that we considered. Therefore, this model can explain some observed differences about these institutions. For example, the model can explain why a professional association cares most about the conduct of its members, thus
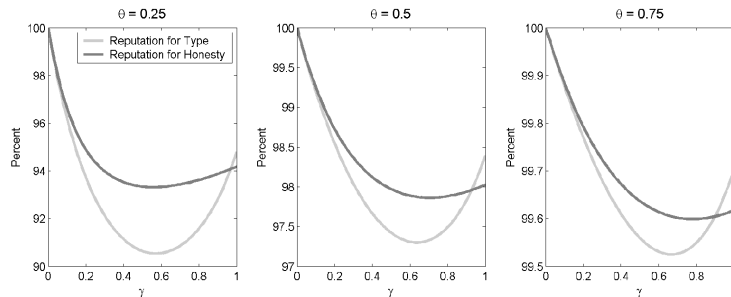
Figure 5: Percentage of full information welfare achieved by the reputation systems.

always rewarding honesty, while a restaurant guidebook will often ignore what is said in the advertisements of sellers. 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.

We caution that our model abstracts from two problems that are also pertinent to the performance of a reputation system in practice. 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). In the context of our model, we can rule out this opportunistic behavior simply by assuming that the reputation system awards the cheap pseudonym strategy a bad reputation in the second period. The excessive honesty problem in our model is exactly opposite to the problem of cheap pseudonyms. 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.

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.[14] If we did allow both types of sellers to pursue this strategy 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 potential directions for further research. It might be

---

[14] This is known as 'reputation milking' and McAfee (2004) discusses this problem in connection to the reputation system of eBay.

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 possibility is to consider the sale of third party information services. One method to sell third party information is an accreditation service that sells reputation services to its members (i.e. the sellers in our model). An alternative 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 choose which sellers to search over – 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] Biglaiser, G. (1993). Middlemen as experts. *The Rand Journal of Economics*, **24**: 212-23.

[3] Bolton, G. E., E. Katok and A. Ockenfels (2004). How effective are online reputation mechanisms? An experimental investigation. *Management Science*, **50**: 1587-1602.

[4] Coles, M. and J. Eechout (2004). Indeterminacy and directed search. *Journal of Economic Theory*, **111**: 265-76.

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

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

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

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

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

[10] Kennes, J. (2004) Competitive auctions: Theory and application. *University of Copenhagen Discussion Paper* 04-16.

[11] Kennes, J. and A. Schiff (2004). The curse of guided search. *Mimeo*, The University of Auckland.

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

[13] Li, Y. (1998). Middlemen and private information. *Journal of Monetary Economics*, **42**: 131-59.

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

[15] McAfee, R. P. (2004) The real lesson of Enron's implosion: Market makers are in the trust business, *The Economists' Voice*, **1**, Issue 2.

[16] 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.

[17] Shimer, R. (2005). The assignment of workers to jobs in an economy with coordination frictions. *Journal of Political Economy*, **113**: 996-1025.

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

[19] 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.

[20] Williamson, S. and R. Wright (1994). Barter and monetary exchange under private information. *The American Economic Review*, **84**: 104-23.

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