Improving the Lemons Market with a Reputation System:

An Experimental Study of Internet Auctioning

Toshio Yamagishi

Masafumi Matsuda

(Hokkaido University)

(May, 2002)

Improving the Lemons Market with a Reputation System:

An Experimental Study of Internet Auctioning

Abstract

Three experiments examined the role of reputation for alleviating the lemons problem in an online market, and produced the following findings. First, information asymmetry drives the experimental market into a lemons market. Second, reputation about other traders moderately alleviates the lemons problem. Third, the power of reputation as a solution to the problem of lemons is substantially reduced when traders can freely change their identities and cancel their reputations. Fourth, the negative reputation system is particularly vulnerable to identity changes. It was argued that the lack of a closed market among online traders, which appears, at first grant, to be a formidable problem, can actually be a blessing.

Improving the Lemons Market with a Reputation System:

An Experimental Study of Internet Auctioning

Introduction

The problem of lemons (also known as fraudulent commodities) is a potential threat to many traders who conduct trades without reliable institutional mechanisms for enforcing contracts. Maghribi traders of the 11th century faced this problem (Grief, 1989, 1993), as do contemporary users of Internet auction sites such as eBay and Yahoo. Both Maghribi traders who conducted business across the Mediterranean and contemporary online traders face highly uncertain environments in which they cannot directly observe the quality of the commodities they trade. Moreover, these traders cannot directly control the behavior of partners with whom they trade. In a now-classical paper, Akerof (1970) argues that such asymmetry of information drives honest traders and high quality goods out of the market, resulting in a market which only lemons or fraudulent commodities pervade. Akerof (1970) examines the used car market as an example of a market for lemons as most consumers are incapable of discerning lemons, and only sellers have information about hidden problems. Online traders are in a similar position as consumers in markets characterized by lemons are. Individuals who make a purchase can learn about the quality and the condition of the good only after they have paid for the good. In the worst scenario, the good never arrives and the buyer cannot track down the seller.

Thus, online traders face greater risk than traders in the traditional form of trading (Friedman and Resnick, 2001; Houser and Wooders, 2000; Kollock, 1999; Resnick and Zeckhauser, 2001; Resnick et al., 2000; Standifird, 2001).

Despite the fact that the high level of information asymmetry has created a situation ripe for frauds, embezzlements, and the lemons problem, surprisingly few frauds are observed in Internet trades (Kollock, 1999; Resnick and Zeckhauser, 2001; Resnick et al, 2000)[1], or in trades among Maghribi traders (Greif, 1989). The students of these traders have found that the reputation system endogenously created keeps the level of frauds low and the problem of lemons from developing. Although the online trading market and the Maghribi trading market is separated by a millennium, it is interesting to note that the same solution—the reputation system—is found to operate in both markets as the central mechanism addressing the problems arising from information asymmetry. Through exhaustive reading of archival records of trades made by Maghribi merchants, Greif (1989) concluded that a unique institution of *coalition* provided a solution to the problems arising from information asymmetry. Grief (1989) argues that the key to the success of the *coalition* in curtailing opportunistic behavior and promoting "trust" among traders was in the fact that reputations were shared within the *coalition*. Observers of online trades (Avery, Resnick and Zeckhauser, in press; Kollock, 1999; Resnick and Zackhauser, 2001) claim that the vast quantity of cheaply available reputation information in online trades offsets the lack of quality and reliability of reputation. In this paper, we examine the claim that the existence of reputations can solve the problem of lemons arising from information asymmetry using an Internet auctioning site set up in our laboratory. Before explaining the details of this experiment, let us further discuss the roles of reputation as a solution to the problem of lemons.

[1] Federal Trade Commission reports, however, that the number of consumer complaints about Internet actions are exploding from 107 in 1997 to 10,700 in 1999. Federal Trade Commission press release, February 14, 2000.

Commitment Formation and Reputation as Solutions to the Problem of Lemons

Does reputation always provide a solution to the problem of lemons? The answer to this question is definitely no. To understand why, let us first assume a market among anonymous traders in which only sellers have information about the quality of the commodities they sell. Buyers can discern the quality only after they have paid for the commodity, and there is no enforcement institution to track down and sanction the seller. "Trust" will not emerge in this market. Buyers cannot distinguish honest sellers from dishonest ones. Thus, sellers cannot be rewarded for their honest trading behavior. The possibility of honest behavior in the absence of third-party enforcement emerges when anonymity of the traders does not exist anymore. On the other hand, when traders are able to identify individuals, they can use future profits as a hostage to induce their trading partners to behave honestly. A good example of this is found in the traditional market for raw rubber in South East Asia. According to Kollock (1994), buying raw rubber involves information asymmetry with very serious consequences, since the quality of the material is known only it is processed. That is, by the time the buyer finds that the material is of low quality he would have invested not only the costs for acquiring the material but also for processing it. The means that rubber traders created to avoid getting lemons in trading rubber is to develop long-term commitment relations with specific producers, sometimes extending over generations. Kollock (1994) contrasts this market with another market in the same region: the market of rice. The quality of rice is apparent to seasoned traders upon simple inspection. Thus, the trading of rice involves no information asymmetry concerning the quality of the traded commodity. Such a situation reduces the need to conduct rice trades through networks of commitment relations, which require high opportunity costs. Rice is in fact traded in the open market.

Commitment formation provides a solution to the problem of lemons to the degree that trades in the market are conducted through commitment relations. In a market in

which most traders are conducted openly, a high premium is required to keep a particular trading partner from behaving dishonestly. This is because the dishonest partner can engage in trading in the open market even when he is refused further trades by the one whom he has exploited. The trader will behave honestly only when the total future premium (with discounting) expected from the partner exceeds the immediate profit of behaving dishonestly. Let us compare this situation with another market in which most trades are conducted through commitment relations. The market for raw rubbers discussed by Kollock (1994) is a typical example of such a market. In such a market, rubber brokers are not willing to buy from a producer with whom he has never traded before unless the producer offers an extremely low price. This is because the probability is high that the previous trading partner has rejected the producer. Otherwise, he would not be seeking a new trading partner. Seeking a new trading partner is a sign of some problem in the previous trades. The broker will pay only the lemons price to such a producer. The difference between the prices the broker pays to his long-lasting trading partner for such a commodity and the lemons price now becomes the basis for the hostage. What the buyer needs to pay to keep the seller from behaving dishonestly is much lower because the basis for the premium is the lemons price (or the price one is willing to pay for the lowest quality goods) *instead of* the current market price.

According to Greif (1989; 1993), sharing reputation about a particular agent with other coalition members reduces the premium required to keep him from behaving dishonestly. When information that an agent behaved dishonestly to one coalition member is shared with other coalition members, and other members who have received such information refuse to deal with the agent, the agent is in the same position as the producer of raw rubber who is rejected by the current broker and for whom other brokers are willing to pay only the lemons price. The sharing of reputation seriously lowers the ability of the dishonest agent to make profit in the future. Then, the premium required to keep the agent from behaving dishonestly is substantially reduced. When reputation is not

shared, “shadow of the future” (Axelrod, 1984) alone is hard to contain dishonest behavior.

Networks of commitment relations and the sharing of reputation are important ingredients for providing a solution to the lemons problem. Another important factor that addresses the lemons problem was pointed out by Greif (1989): the fact that the coalition has closed boundaries. If a dishonest agent can freely engage in trade with non-coalition members, then the basis of the premium is raised to the profit he earns in dealing with non-coalition members. Such a situation greatly hampers the power that shared reputation has to inhibit dishonesty. Finally, reputation will have no power to curtail dishonest behavior if the coalition members do not base their behavior on the reputation. In short, reputation can provide an effective solution to the problem of lemons when (1) it is shared by all or most of traders in the market, (2) traders in fact base their behavior on it, and (3) the market is closed such that the trader who is excluded from it cannot find an alternative market. Greif (1989; 1993) argues that the trades among Maghribi coalition members met all of these three conditions. Observers of online trades (Kollock, 1999; Resnick and Zeckhauser, 2001) are less certain as to the extent to which the Internet auction market meets these conditions. Nonetheless, they provide evidence that the lemons problem is mostly contained in the market. Before discussing why the lemons problem is contained in online trades, let us briefly introduce the terminology we use in the remainder of the paper.

Experience-based Information and Reputation

The above argument by Greif (1989) makes a distinction between shared and non-shared information important. An individual can learn about her trading partner through direct experience, or she can learn about her trading partner from someone else who has had direct experience with him. We may call the former “private reputation” or “direct reputation” and the latter “public reputation” or “shared reputation.” To make the terminology clear, we use “experience-based information” to refer to the former and

restrict the use of the term "reputation" to the latter. According to a definition of reputation by Wilson (1985: 27-8), a reputation is a "characteristic or attribute ascribed to one person … by another." Similarly, according to Standifird (2001: 281), "(r)eputation is defined as the current assessment of an entity's desirability as established by some external person or group of persons." According to our terminology, a reputation is a characteristic or attribute ascribed to one person (and is believed to be useful in predicting that person's future behavior) by a third-party.

Reputation System as a Solution to the Lemons Problem in Online Trades

The contemporary observers (Kollock, 1999; Resnick and Zeckhauser, 2001) of online trades, especially Internet auction markets such as eBay or Yahoo Auction, report surprisingly low levels of fraud. This finding implies that the problem of lemons is mostly contained to an acceptable level. Yet, contemporary online traders face a market that greatly differs from the one faced by Maghribi traders of the 11th century. Most importantly, the online market has no closed boundaries. Furthermore, online traders can assume many identities and/or create many pseudonyms. Below, we will discuss the factors that are expected to work against a reputation system as a solution to the lemons problem. Then, we will point out factors that are expected to offset the negative elements, and that are hypothesized to support the power of reputation as a solution to the lemons problem.

The first issue concerns the free-riding problem in providing reputation (Avery, Resnick and Zeckhauser, in press; Resnick et al., 2000; Resnick and Zeckhauser, 2001). There is no incentive for online traders to share one's own experience-based information with others. Shared reputation is a public good to which one can freely access. Contributing to its provision is not a condition for access. According to the "logic of collective action" (Olson, 1965), such a public good will eventually fail to be provided, or if provided at all, will be underprovided. Although Greif (1989; 1993) does not explicitly discuss the problem of insufficient incentives for the voluntary provision of reputation

information in his analysis of Maghribi traders, we can imagine reasons why this was not a major problem with Maghribi traders. The provision of sanctioning is logically identical with the provision of a public good, which involves a second-order social dilemma problem (Yamagishi, 1986). In short, people who benefit from a sanctioning system are not willing to bear the cost for providing and maintaining the system. According to Axelrod (1986; see also Yamagishi and Takahashi, 1994), this problem can be resolved when people sanction those who do not sanction non-contributors to a public good. Applying this logic to the provision of a reputation system, the second-order problem can be resolved when those who do not share their personal experiences with others receive a negative reputation for that and are then treated in a similar manner as dishonest traders are treated. This could have been the case, although Greif (1989) is not explicit on this issue, among Maghribi traders, but definitely not the case for online traders. It is not common to give negative evaluations to someone who has failed to provide reputation information on the Internet auction market. Thus, the public good problem is expected to be more serious among online traders than among Maghribi traders.

The second issue concerns the stability of identity. Online traders can have as many identities or handling names as they wish and can change their identities as often as they wish. An online trader who has accumulated negative reputations can shake them off by assuming a new identity with a new email address and thus pretending to be another person. As a result, the effectiveness of reputation as a solution to the problem of lemons can be greatly undermined among online traders. This problem is especially serious for negative reputation, which, according to research of trades on Ebay (Houser and Wooders, 2000; Lucking-Reiley et al., 2000; Resnick and Zeckhauser, 2001; Standifird, 2001) has a more powerful effect on price than positive reputation. Positive reputation, however, is free from this problem since online traders have every incentive to maintain the positive reputation that accompanies their identity. Positive reputation is a valuable asset, and traders who have acquired a high reputation will not voluntarily change their identities.

The above discussion leads to the third issue: the existence of both positive and negative reputations. The effectiveness of a negative reputation system to contain the lemons problem is compromised to the degree that those who behave dishonestly have alternative markets to move into without paying exit and entrance costs. And, the extreme example of such a loose market is the online market, or more specifically, the Internet auction market such as eBay and Yahoo. A dishonest trader does not even need to exit from the market in which he has exploited others and has acquired a negative reputation. He can simply change his handling name and re-enter the same market as a different person. In addition, fear of retaliation prevents online traders to provide a negative evaluation to their trading partners (cf., Resnick and Zeckhauser, 2001). Thus, providing a solution to the lemons problem in the online market with a reputation system seems to be a hopeless endeavor. The reality, however, is much brighter than it appears (Kollock, 1999; Resnick and Zeckhauser, 2001). Reported frauds are rare and the reason for the low rate of dishonest behavior has been attributed to the reputation system used in the online market. Many researchers (e.g., Kollock, 1999; Resnick and Zeckhauser, 2001; Standifird, 2001) assert that the surprising power of reputation to contain dishonest behavior stems from the speed and low cost for disseminating information in online markets. They claim that the disadvantages of online markets for reputation are well offset by the sheer quantity of cheaply and quickly disseminated reputation information.

Kollock (1999) proposes another possible reason for the theoretically unexpected strong effect of reputation in online trades. He argues that positive reputation rather than negative reputation should function as an effective solution to the lemons problem in the online market. He argues that the negative reputation system is doomed to failure in the online market due to the reasons summarized above. On the other hand, he continues, traders have incentives to maintain a positive reputation because a positive reputation is a valuable asset. The freedom to change identities will not affect the effectiveness of the positive reputation system since traders voluntarily keep their “brand name” once they

have acquired a positive reputation. We agree with Kollock's (1999) assessment of positive reputation. However, Kollock (1999) is missing an important theoretical issue. What is the mechanism with which the positive reputation keeps traders from dishonest behavior? In the case of negative reputation, it was the fear of losing trading partners who would provide opportunities for future profits. We will return to this important theoretical issue in the discussion section of this paper, after empirically examining if a positive reputation system is equally or more effective than a negative reputation system in the experimentally created online market.

Basic Design of the Experiment

We use a laboratory version of online trade to represent a market characterized with information asymmetry. In this experimental market, all players (i.e., voluntary participants) take on two roles: the buyer's role and the seller's role. As a seller, each player produces a "commodity" (abstract commodity without any substance) by expending some cost, ranging from 10 yen to 100 yen per commodity, in increments of 10 yen.[2] The quality level of the produced commodity is determined by the amount of money a seller has expended for its production. For simplicity, the quality level is expressed in terms of how much money (yen) the seller has spent for its production, ranging from 10 to 100 as an integer number. For example, a commodity with a quality level of 70 is a commodity that the seller spent 70 yen for its production. The commodity produced by the seller is then put on a market for sale with a price set by the seller. At the same time, the seller announces the quality of the commodity—i.e., how much she has spent for its production. The "advertised" quality level may or may not reflect its true quality. For example, a seller S_1 who has spent 30 yen for the production of a commodity may put it on the market with an announcement that its quality level is 80 (whereas its true quality

[2] The exchange rate for a dollar was roughly 100 yen at the time the experiment was conducted.

level is 30). Any other player, acting as a buyer, can purchase the commodity for that price. Thus, when a seller S_1 expends 30 yen on a commodity and advertises it for sale for 120 yen, and a buyer, B_2, purchases it from S_1, the seller S_1 makes a profit of 90 yen (she received 120 yen from the sale of a commodity in which she invested 30 yen for its production: 120 – 30 = 90). As a buyer, each player may purchase any of the commodities placed on the market. In purchasing, the information available to a buyer is the price and the advertised quality of the commodity. The purchased commodity is immediately and automatically resold to the experimenter. The experimenter adds a 50% premium on the commodity's true quality level (i.e., its production cost). That is, the experimenter pays 1.5 times the true quality of a commodity to the buyer. It is only when the commodity is resold to the experimenter that the buyer finds out the true quality of the commodity. The difference between the resale price (1.5 times the true quality level of the commodity) and the purchase price is the profit for the buyer. In the above example of a trade in which a buyer pays 120 yen for a commodity with a true quality level of 30, the buyer B_2 loses 75 yen since the commodity is resold to the experimenter for 1.5 times its true quality of 30 (45 yen). He has paid 120 yen for something which is worth only 45 yen.

As mentioned before, each player takes on the role of a seller as well as that of a buyer. He can act as a buyer at any time. On the other hand, he can act as a producer/seller only sporadically. Specifically, each player is provided with a "production opportunity" every 50 seconds with a random variation within the range of plus and minus 15 seconds. Once a production opportunity is provided, a player decides how much to invest in the production of a commodity (in increments of 10 yen). The amount of money she spends for the production of a commodity is immediately deducted from her account. A player is required to produce one and only one commodity per production opportunity.

Figure 1 is an example of how the market looks on a player's computer screen. The screen is divided into two sections: the upper section and the lower-section. The upper section of the screen displays commodities placed for sale by the other players, and the

lower section displays commodities the player herself has produced and placed for sale on the market. Each box shown in the upper section represents a commodity. In the box located on the lower-left corner of the screen, for example, the seller/producer *to-me-ko* sells the commodity for 40 yen, while advertising its quality level to be 30. Please note that the identity of the producer/seller is provided in this example, whereas this is not always the case in the experiments presented below. Also note that the original screen is written in Japanese, not in English. Each player's screen is unique, since the commodities the player has produced does not appear in the upper section of the screen for that player.

Insert Figure 1 About Here

The lower section of the display (Figure 1) indicates that this player *se-se-i* is currently selling two commodities represented by the two boxes shown in that section. To produce the first commodity displayed as the box on the far left side of the lower section, this player spent 10 yen. She advertises its quality level as 100, and is selling it for 140 yen. The box next to the right of the first box represents another commodity this player is currently selling in the market. The right corner of the lower section displays the identity of this player, *se-se-i*, and her current financial situation. Another feature of this experimental market, common to all experiments presented below, is that a commodity placed on the market remains on the market for only 5 minutes. A commodity that has not been sold within 5 minutes perishes. The commodity disappears from the market and the producer/seller cannot recover the cost he spent for its production. This option was introduced to the current experiment to avoid overcrowding the market with too many commodities.

Experiment 1: Emergence of a Market for Lemons

Purpose

We conducted the first experiment to demonstrate that the market for lemons will result from the market described above. In the control condition of the first experiment,

no additional information such as the seller's identity or reputation was provided. That is, boxes displayed in the upper section of each player's computer display had only two pieces of information: the price and the advertised quality level. The buyer was not informed even who was selling which commodity. In this market with complete anonymity, rational sellers are expected to spend only the minimum requirement of 10 yen to produce a commodity per opportunity[3]. Knowing this, rational buyers are expected not to spend more than 15 yen, or the resale price of a minimum quality good, to buy a commodity. Even when a producer/seller spends 100 yen to produce a commodity of which the level of quality is 100 and places it for sale on the market at the price of 110 yen, potential buyers cannot discriminate it from a lemon with the lowest quality level of 10 being sold for 110 yen with an advertised quality level of 100. The likelihood that a producer of a high quality good successfully sells it on this market is exactly the same as the likelihood that another producer of a lemon will successfully sell it for the same price. Knowing that rational producers will not invest more than 10 yen in the production of a commodity, or the smallest required amount, no rational buyer will pay more than 15 yen[4] for a commodity regardless of its advertised quality. This is the pure form of the market for lemons that we predict to result in the control condition. One purpose of the first experiment is to demonstrate that this theoretical prediction of the market for lemons does in fact occur in the control condition with complete anonymity.

The second purpose of the first experiment is to determine the magnitude of the effect of additional information such as the seller's identity or reputation. In the identity condition, each player was assigned a unique identity, one of 48 Japanese characters.[5]

---

[3] As stated earlier, the player was required to spend a minimum of 10 yen for the production of a commodity in each production chance.

[4] Note that the minimum increment in sales price is set at 10 yen to alleviate information overload on the player. A buyer who purchases a commodity of which true quality level is 10 makes 5 yen if she pays 10 yen for it. If he or she pays 15 yen, then she does not make a profit nor suffer a loss.

[5] The number of letters was increased to three in the second and the third experiments.

Each commodity on the market came with an ID letter denoting its producer/seller. The buyers in this condition would naturally avoid purchasing commodities from the seller with whom they had sour experiences. They would prefer purchasing commodities from the sellers who had sold them, in the past, profitable commodities. This discriminatory purchasing behavior of buyers based on their own experience is expected to provide producer/sellers to behave honestly and to be chosen as trading partners. If a producer/seller can persuade potential buyers that purchasing his product will not lead to a loss, he will be able to make a profit of 45 yen per production opportunity. This happens when he spends 100 yen on production and sells the commodity for 145 yen (leaving a minimum of 5 yen as profit for the buyer)[6]. If the producer/seller is not trusted, he or she will produce the lowest quality commodity for which buyers will not pay more than 15 yen. The seller who always sells a lemon will eventually be deserted by the other players as a trading partner. He cannot make any profit once everyone in the market has learned that he is a seller of lemons. One can thus expect a "premium of honesty" worth 45 yen per production chance, and this premium of honesty expected in the future trade is expected to work as a hostage to discourage players from behaving dishonestly. The players were not told how long the experimental market would last.

Does the above argument imply that the problem of lemons is resolved in the identity condition? "Not completely" is the answer to this question. This is because a producer/seller does not need to be trusted by all potential buyers. If the size of the market is large enough, a seller can successfully get away with occasional cheating while keeping the minimum circle of trustful buyers. Thus, it is predicted that giving players identity will improve the market for lemons to a certain degree, but the improvement will be less than perfect.

The last condition used in the first experiment is the reputation condition. As mentioned in the introduction, we use the term "reputation" as indirectly acquired

[6] Prices were set in increments of 5 yen. Therefore, the minimum profit was 5 yen.

information. To assess the effect of reputation independent of the effect of experience-based information (available in the identity condition), we made only reputation information available in the reputation condition. Specifically, unique identities were not attached to the players as in the control condition. Each buyer evaluated a seller after a purchasing deal was completed and the true quality level of the purchased commodity was revealed. Evaluation was made by assigning a number between -2 and +2 (-2 = very bad, -1 = bad, 0 = neutral, +1 = good, +2 = very good) to the seller from whom the buyer had just purchased a commodity.[7] Thus, buyers repeatedly evaluated each seller, and an overall evaluation score or the sum of evaluations the seller had acquired was calculated. Each commodity placed on the market for sale by a seller came with a label indicating the overall evaluation score of the seller at the time when she made the product. However, providing an exact evaluation score for each seller potentially makes it possible for the buyers to identity the seller by the evaluation score. If this happens, a unique evaluation score that substitutes a unique ID can make measuring the effect of reputation separately from the effect of identification very difficult. To avoid this, we decided not to provide the exact evaluation score for each seller. Instead, we provided an evaluation score for the seller by a color and the shade of the color. More specifically, each box representing a commodity produced by a particular producer/seller was colored either blue or red[8]. Blue indicates the total evaluation score of that producer/seller was positive, and red indicates that the total evaluation score was negative. In addition, the shade of green or red reflected the absolute value of the total evaluation score. That is, dark green revealed a highly negative score while dark red revealed a highly positive score. Light green revealed a slightly negative score and light red revealed a slightly positive score.

[7] Evaluation of the seller was made in all conditions, although feedback of the aggregated evaluation to the buyer was made only in the reputation condition.
[8] Those colors were changed to green and red from the second experiment on.

If buyers honestly reveal their evaluations of the producer/sellers from whom they have purchased a commodity, the "quality label" expressed as the color of the box will provide a solution to the lemons problem. Buyers will avoid commodities with a negative label as they would avoid commodities produced by producers/sellers with whom they have had sour experiences. That is, a quality label should fulfill the role that direct experiences play in the identity condition. Furthermore, the quality label can overcome the expected shortcomings of the experience-based information as the solution to the lemons problem. The quality label is available to any potential buyer, and thus the sellers who want to earn good quality labels have to behave consistently. Exploiting some buyers and reserving a few as repeated customers is not a viable option in the reputation condition. This situation is in fact the essence of Greif's (1989) discussion of the Maghribi traders' coalition. However, the "quality label" has its own shortcomings. There is no incentive for the buyers to honestly reveal true evaluation. Sharing one's own information about a seller—a private good—with others is providing a free gift to them. Furthermore, there is even an incentive to give a negative evaluation to all the other sellers regardless of the quality of the commodities sold, since negative evaluations of other sellers improve the relative standing of one's own evaluation. In the absence of selective incentives for honest revelation of evaluations, reputation or the quality label is expected to provide at best a partial solution to the problem of lemons.

Based on the above arguments, we conducted the first experiment to address the following research questions.

Question 1: Will the market for lemons result in the control condition? A lemons market occurs when the quality of the commodities produced and sold in the market (and thus the total profits generated in the market) is at the lowest possible level.

Question 2: Will providing the seller's identity alleviate the lemons problem?

Question 3: Is providing the seller's identity less than perfect in trying to solve the lemons problem?

Question 4: Will providing the seller's reputation alleviate the problem of the lemons?

Question 5: Is providing the seller's reputation less than perfect in trying to solve the lemon's problem?

Other questions: We will also address more detailed questions such as the following: Does honesty pay? How accurately does the evaluation score reflect the honesty of the seller?

Procedure

Participants in the first experiment were fourteen students (10 males and 4 females) who had been recruited from a participant pool of about 1,500 students enrolled in a major university in Japan. Monetary incentives were emphasized, and no class credit was involved in the recruitment of the participants. Participants were randomly assigned to two groups, each consisting of seven participants. Each group participated in six experimental sessions. The first three sessions each lasted 45 minutes, and the last three sessions each lasted 20 minutes. In total, the experiment lasted an entire afternoon (approximately 5 hours). Participants were paid the amount they earned in the experiment. The most any one participant earned was 8,965 yen (about $80) and the least any one participant earned was 910 yen (about $8). The average pay was 5,597 yen. Each group experienced the three conditions in the three 45-minute sessions; each of the 45-minute sessions were assigned to one of the three conditions. In addition, we added the 20-minute sessions to learn how the experience of the first three sessions would affect the participants' behavior. The three 20-minute sessions were assigned to the three conditions as well. The order of presentation of the three conditions was randomized for each group. A rest period of 10 minutes was provided between sessions.

The experiment was conducted in a laboratory consisting of 8 small rooms. Each room was equipped with a computer that was connected via LAN to a host computer controlled by the experimenter. Upon arrival at the laboratory, each participant was

escorted to his or her individual room. All instructions were presented on the computer screen. The production and purchasing decisions made by the participant were entered to the participant's computer with a mouse. In each session, participants were allowed to produce commodities as many times as possible for the first two minutes. That is, a new production opportunity was given to the participant as soon as he or she finished producing a commodity during the first two minutes of the session. The experiment was conducted in this way to provide the market with a sufficient number of commodities with which to start. After the first two minutes, a production opportunity was provided to the participant every 50 seconds with a random variation within the range of plus or minus 15 seconds.

Findings

Figure 2 presents the average level of quality during the 45-minute session and the 20-minute session. Each session is broken into time blocks lasting five minutes each. The main effect of the condition in a condition (a repeated factor) × time block (a repeated factor) ANOVA on the 45-minute session indicates a significant main effect of the conditions, $F(2, 39) = 16.73$, $p < .0001$. The main effect of the time block, $F(8, 312) = 2.05$, $p < .05$, and the conditions × time block interaction, $F(16, 312) = 4.34$, $p < .0001$, were also significant. As shown in the figure, both the main effect of the time block and the interaction effect reflect the downward trend of the average quality of the commodity that occurred only in the control condition. A post-hoc analysis of means with Tukey's studentized range test indicates that the mean quality level of the control condition (29.65) is significantly different from either of the other two conditions (70.76 in the identity condition and 71.07 in the reputation condition), of which quality levels were not significantly different from each other. A similar analysis was conducted on the 20-minute session that was held after the first round of three conditions were completed. Similar results were observed for the 20-minute session. As in the 45-minute session, the main effect of the conditions was significant, $F(2, 39) = 41.71$, $p < .0001$, and so was the

main effect of the time blocks, $F(3, 117) = 12.74$, $p < .0001$. The conditions × time blocks interaction effect, however, was not significant, $F(6, 117) = 1.60$, *ns*. As in the 45-minute session, the difference was found only between the control condition (16.52) and the other two conditions (67.51 in the identity condition and 76.89 in the reputation condition).

Insert Figures 2 About Here

The level of dishonesty—the difference between the advertised quality and the true quality—was much higher in the control condition than in the other two conditions. In the ANOVA of the level of dishonesty for the 45-minute session, the main effect of the conditions was significant, $F(2, 39) = 5.36$, $p < .01$. The main effect of the time block, $F(8, 312) = 1.93$, $p < .06$, and the conditions × time blocks interaction effect, $F(16, 312) = 1.56$, $p < .08$, were marginally significant. For the 20-minute session, the main effect of the condition was significant, $F(2, 39) = 3.61$, $p < .01$, and so were the main effect of the time blocks, $F(3, 117) = 2.93$, $p < .05$, and the interaction effect, $F(6, 117) = 4.74$, $p < .001$. Below, we will examine each of the three conditions more carefully.

Control condition. As seen in Figure 2, the lemons market was found in the control condition. The average quality of the commodities was reduced to close to the minimum level of 10 for the last 10 minutes of both the 45-minute session and the 20-minute session. The overall average quality level was 29.65 in the 45-minute session and 16.52 in the 20-minute session; the average quality during the last 10 minutes was 12.41 (45-minute session) and 10.59 (20-minute session). Furthermore, the level of dishonesty was quite high in this condition as shown in Figure 3. In addressing the first question: Will the market for lemons result in the control condition, these results clearly provide a positive answer.

Another interesting finding concerns the relationship between dishonesty and profit. To analyze the relationship, we performed a regression analysis in which the dependent variable was the profit participants earned by producing and selling commodities and the

independent variable was the participant's level of dishonesty. In order to control for the between-groups differences, a dummy variable for the groups (0=Group 1 and 1=Group 2) was added to the list of independent variables. The same analysis was conducted separately for the 45-minute session and the 20-minute session. The results indicated that the level of dishonesty was positively related to the profit both in the 45-minute session, $b = 30.03$, $t(11) = 4.36$, $p < .01$, and in the 20-minute session, $b = 4.38$, $t(11) = 1.23$, *ns.*, though the relationship was statistically significant only in the 45-minute session. In the 45-minute session, an average increase of one point in the average advertised quality level above the true quality level generated about 30 yen of total profit. Dishonestly claiming high quality was clearly a better strategy for making profits than honestly revealing the true quality. The reason why the effect of dishonesty on profit was greatly reduced in the 20-minute session may be attributed to the fact that hardly anyone could make profit from producing and selling commodities. The average profit from producing and selling commodities in the 20-minute session was only 1.43 yen, compared to 717.14 yen in the 45-minute session.

Identity condition. The overall quality of the commodities produced in the identity condition was 70.76 in the 45-minute session, and 67.51 in the 20-minute session. As shown in Figure 2, there was no consistent upward or downward trend over time in the quality of produced commodities in the 45-minute session. The 20-minute session started with a fairly high level of average quality, but the average quality level quickly fell and then remained around 55. These results show a modest success of availability of experience-based information as a means to resolve the problem of lemons, and provide a positive answer to the second question: Will providing the seller's identity alleviate the lemons problem? At the same time, the answer to the third question, Is providing the seller's identity less than perfect in trying to solve the lemons problem?, is clearly positive. The success was rather modest.

Did dishonesty pay as it does in the control condition? The result of a regression

analysis (similar to the one used in the control condition) indicates a weaker positive effect, $b = 11.87$, $t(11) = 1.23$, *ns.*, in the 45-minute session. In the 20-minute session, the regression coefficient was similar, but the effect was statistically significant, $b = 14.44$, $t(11) = 3.02$, $p < .05$. Given the fact that dishonesty paid in the identity condition, the limited success of the identity information in curtailing the lemons problem seems to be precarious.

Reputation condition. The overall quality of the commodities produced in the reputation condition was 71.07 in the 45-minute session, and 76.89 in the 20 minute session. There was not much difference in the average quality level of commodities produced in the identity condition and in the reputation condition in the 45-minute session. Although the average quality level was slightly higher in the reputation condition (76.89) than in the identity condition (67.51) in the 20 minute condition, the difference was not statistically significant. The trend over time shown in Figure 2 for the reputation condition was similar to the pattern for the identity condition. There was no consistent upward or downward trend over time in the 45-minute session, and an initial sharp decline followed by a steady trend was observed in the 20-minute session. These results show a modest success of reputation as a means to resolve the problem of lemons—positive answers to the fourth and the fifth questions.

Dishonesty did not pay in the reputation condition. The regression coefficient for dishonesty in the 45-minute session was negative, $b = -8.49$, $t(11) = .46$, *ns*. In the 20-minute session, the regression coefficient was positive but not significant, $b = 11.16$, $t(11) = 1.29$, *ns*. Thus, neither honesty nor dishonesty paid in the reputation condition. This result suggests that the reputation information seems to be more promising for resolving the lemons problem than the identity information or experience-based information. Finally, we added the average reputation score of the seller to the above equation. The regression coefficient for reputation in this analysis for the 45-minute session was positive but not statistically significant, $b = 19.87$, $t(10) = 1.60$, *ns*. In the

20-minute session, the coefficient was smaller, $b = 6.06$, $t(10) = 0.29$, *ns*. Earning a good reputation seemed to help the participants achieve profits, but the effect was not unequivocal.

The results of the first experiment emphasize several findings. Clearly, the market for lemons, predicted for the control condition in which neither identity or reputation was provided, does occur. Either providing identity information of the players or reputations of the players help alleviate the problem, but either identity or reputation was shown to provide only limited success. These findings provide a benchmark for evaluating the results of the second and the third experiments that will be presented below.

Experiment 2: Cancellation of Reputations

Purpose

In the first experiment, identity and reputation each provided at least a partial solution to the problem of lemons. However, such results might have critically depended on a unique feature of the experimental market used in the first experiment. That is, identity and reputation was permanent in the sense that the players themselves could not cancel them. This situation of permanent identity and reputation does not exist, however, when transactions take place in an open market among a large number of players. An extreme case of such an open market is found in internet auctioning in which one participant can assume a practically unlimited number of identities. The question we address in the second experiment is whether or not identity and reputation have a positive effect on resolving the problem of lemons when players can freely change their identities and thus, erase their reputation.

Letting players freely change their identities and erase their reputation seems to neutralize the positive effect of identity and reputation observed in the first experiment. A seller who cheats a buyer and then assumes a new identity is allowed to approach the same buyer, and the buyer will not be able to determine if he is the one who has cheated

her. This simple assessment of the situation, however, can be misleading. The effect of identity change is asymmetrical. On the one hand, players who cheat others and acquire a negative reputation want to assume a new identity as briefly suggested above. On the other hand, players who conduct honest deals and acquire a positive reputation want to keep the old identity and reputation. As a result, dishonest players will often change their identities and re-enter the market with no reputation, whereas honest players will maintain their established identity and the positive reputation associated with that identity. When this happens, a new identity and a neutral reputation might signal risks to buyers. Players in this market will come to disregard the distinction between a negative reputation and a neutral reputation, and focus on the distinction between a positive and a neutral reputation. Treating a neutral reputation and a negative reputation identically will create an obstacle for true new comers to enter the market (Friedman and Resnick, 2001). Besides this problem of market entry, identity and reputation will have positive effects in this market similar to those in the market with permanent player identities. We conducted the second experiment to examine whether identity and reputation, when they can be freely changed, will have positive effects in alleviating the problem of lemons, similar to those found in the first experiment.

Procedure

Only one condition was used in the second experiment. This condition was the one in which ***both*** identity and reputation of the seller were provided. That is, the commodity placed on the market by a seller came with an identity label of the seller and her reputation color. The seller could choose her identity as a combination of three Japanese characters, at any time during the experiment. As in the first round of the first experiment, the experiment lasted for 45 minutes. Thirty-four participants recruited from the same participant pool used in the first experiment participated in five groups. The experimental session was not repeated. Each participant experienced only one 45-minute session and was paid the amount he or she earned in that session. The average payment was 589 yen

with the minimum of -1,920 yen[9] and the maximum of 3,195 yen.

Findings

The average quality level of commodities during the 45 minutes of experiment is presented in Figure 3. The overall average quality level of the commodities was 44.09, which was much higher than the quality level in the control condition of the first experiment (29.64, $t(46) = 2.22$, $p < .05$). However, the overall average quality level was less than the theoretical maximum of 100 or even the average quality level in the identity condition (70.76, $t(46) = 3.89$, $p < .001$) or in the reputation condition (71.07, $t(46) = 4.41$ $p < .0001$) of the first experiment.[10] Thus, providing the identity and reputation of the sellers helped to alleviate the problem of lemons even when players can freely change their identities and cancel their reputations. However, the freedom of assuming new identities substantially reduces the positive effect of identity and reputation observed in the first experiment.

Insert Figure 3 About Here

The results of the second experiment also indicate that dishonesty paid. The regression coefficient for the level of dishonesty on the profit from producing and selling commodities was positive and significant, $b = 23.20$, $t(28) = 3.71$, $p < .001$. The regression coefficient for the average reputation score when that term was added to the list of independent variables was positive but not statistically significant, $b = 35.06$, $t(28) = 1.69$, *ns*. Acquiring a good reputation seemed to help the participants to achieve good profits, but the effect was not statistically reliable. Overall, the limited positive effect of identity and reputation observed in the first experiment was weakened by the option to freely change identities and cancel reputations in the second experiment, despite the fact

[9] Participants whose earnings did not reach 300 yen were paid a minimum payment of 300 yen.

[10] Only the 45-minute sessions in the first experiment were used in these comparisons.

that both identity and reputation were available there.

We then analyzed the relationship between the frequencies with which participants changed their identities and the levels of their reputation and profit. On average, participants changed their identities 7.00 times. While 38% of the participants did not change their identities at all, 23.5% of them changed their identities more than 10 times. As expected, the correlation between the frequency of identity changes and reputation was negative, $r = -.41$, $p < .05$. The participants whose overall average reputation scores were zero or less (n = 16) changed their identities 12.06 times on average, whereas those with a positive average reputation score (n = 18) changed their identities only 2.50 times on average. Among 9 participants whose overall average reputation score was 5 or more changed their identities only 0.33 times on average. The correlation between the level of dishonesty and the frequency of identity changes was more pronounced, $r = .88$, $p < .0001$. These figures confirm our expectation that it is mostly dishonest sellers with negative reputations who change their identities. In conclusion, selling fraudulent lemons and frequently changing identities constituted a "smart" strategy in the market adopted in this experiment where players could freely change identities.

Experiment 3: Positive and Negative Reputation Systems

Purpose

The results of the second experiment illustrate that the freedom of adopting new identities and canceling reputations does not completely nullify the positive effect of identity and reputation. The positive "residual" effect of the unstable identity and reputation, it was argued, is due to the asymmetrical nature of reputation. That is, players want to cancel negative reputations, on the one hand, and yet want to keep positive reputations, on the other. This asymmetry of reputation implies another asymmetry with respect to the effect of reputation. Suppose that the evaluations of sellers that buyers provide are only in the negative direction. Specifically, suppose that evaluation scores

used in the above experiments had only three levels, 0 for neutral, -1 for bad, and -2 for very bad. Such a scale would make reputation completely meaningless, since those who earn negative reputation scores soon change their identities and enter the market under new identities and a reputation score of zero. If this happens, practically everyone will have a reputation score of zero, which is equivalent to the situation in which no reputation is provided. And yet, even in this dismal situation, there must be an incentive to maintain an established "brand name." An established brand name with the reputation score of zero means that the seller has not cheated and received negative reputation in the past. Compared with a seller with a relatively new name with the same reputation score of zero, buyers are expected to seek trade with the seller who has an established brand name. It is thus expected, as in the second experiment, that even negative reputations alleviate the problem of lemons to some degree.

How about positive reputation? Suppose buyers evaluate the seller with three levels: 0 for neutral, 1 for good, and 2 for very good. The positive effect of reputation is expected to be the strongest here. Since even a slightly positive reputation is better than having none, anyone who has earned a positive evaluation score should be motivated to maintain it. This incentive to maintain a positive reputation will discourage a seller from changing identities. The players who have earned high reputation scores will be even more strongly motivated to maintain the current identity. Thus, the detrimental effect of the freedom to change identities will be minimized with a positive reputation system. We conducted the third experiment to examine if this predicted asymmetry of positive and negative reputation systems exists.

Procedure

The experimental procedure used in the first two experiments was employed in this third and final experiment. Participants were 32 volunteers recruited from the same participant pool. Twenty were assigned to the positive reputation condition in three groups, and the remaining 12 were assigned to the negative reputation condition in two

groups. As in the second experiment, each participant was allowed to assume a new identity, a new combination of three Japanese characters. When they did, they started over with a reputation score of zero. In the positive reputation condition, participants gave a positive evaluation score (0 for neutral, 1 for good, 2 for very good) to the seller every time they purchased a commodity. They assigned a negative evaluation score (0 for neutral, -1 for bad, -2 for very bad) in the negative reputation condition. The experimental session was not repeated. Each participant experienced only one 45-minute session and was paid the amount he or she earned in that session. The average payment was 366 yen with the maximum of 1,540 yen and the minimum of –1,185 yen.

Findings

The difference in the average quality levels of the commodities produced under the positive reputation system and the negative reputation system is apparent in Figure 4, especially toward the end of the 45-minute session. The overall quality level in the positive reputation system is 48.46, which is higher than that in the negative reputation system, 38.94. While the main effect of the condition was not statistically significant, $F(1, 29) = 1.33$, *ns.*, the condition × time block interaction was significant, $F(8, 232) = 7.48$, $p < .0001$. The main effect of time blocks was not significant, $F(8, 232) = 1.58$, *ns*. The difference between the two conditions in the average quality of the produced commodities reached the significance level in the last time block, $F(1, 29) = 9.06$, $p < .01$. Interestingly, the average quality level was higher in the negative reputation condition (55.21) than in the positive reputation condition (40.57) in the first time block, and the difference was significant, $F(1, 29) = 4.84$, $p < .05$. Negative reputations initially had a strong impact, but it quickly waned as sellers who had acquired negative reputations started to change their identities. Overall, the negative reputation system provided only a minor improvement in the average quality of produced commodities (29.65 vs. 38.94, $t(24) = 1.11$, *ns.*), compared to the control condition in the first experiment in which no identity or reputation information was provided. The positive reputation system was more

promising. The overall average quality of produced commodities was significantly better in the positive reputation system than in the control condition of the first experiment (29.65 vs. 48.46, $t(32) = 2.45$, $p < .05$). And, as shown in Figure 4, the difference increased toward the end of the 45-minute session. In the last time block, the difference was fairly large (11.96 vs. 60.28, $t(31) = 4.95$, $p < .0001$).[11] However, the average quality level in the positive reputation condition was lower than that in the reputation condition in the first experiment (48.46 vs. 71.06, $t(32) = 3.20$, $p < .01$). The difference in the last time block was smaller and not significant (60.28 vs. 77.31, $t(31) = 1.60$, *ns*.). These findings seem to suggest that the effect of the positive reputation system takes time to be realized by the participants.

Insert Figure 4 About Here

The level of dishonesty had a negative but non-significant effect on profits from producing and selling commodities in the positive reputation condition, $b = -8.82$, $t(16) = .1.14$, *ns*., when group differences were controlled. In the negative reputation condition, dishonesty had a positive but non-significant effect, $b = 3.36$, $t(9) = .49$, *ns*. The average reputation, when it was added to the regression equation, had a positive effect in the positive reputation condition, $b = 47.60$, $t(15) = 5.07$, $p < .0001$, but not in the negative reputation condition, $b = -7.42$, $t(8) = .10$, *ns*. As expected, acquiring positive reputations was instrumental for achieving good profits, but avoiding negative reputations was not useful in this respect.

The expected asymmetry in reputation clearly emerged, as shown in the differential frequencies of identity changes. In the positive reputation system, participants changed their identities 1.75 times on average, whereas the average frequency of identity changes in the negative reputation system was 10.17, and the difference was statistically significant, $t(30) = 3.78$, $p < .001$. Interestingly, in the negative reputation condition, the

[11] The *df*. in this t-test is 31, instead of 32, because one of the participants in the positive reputation system did not produce any commodities during the last time block.

correlation between reputation score and frequency of identity change was positive, although not significant, $r = .42$, *ns*. This seems to reflect the fact that frequent identity changers' reputation score stays close to zero, whereas infrequent identity changers accumulated negative reputations. The correlation was negative in the positive reputation condition, $r = -.52$, $p < .05$, reflecting the fact that those who have acquired high reputation scores maintained their brand names. The correlation between the level of dishonesty and frequency of identity changes was very high in the negative reputation condition, $r = .88$, $p < .001$, whereas the correlation was positive but modest in the positive reputation system, $r = .40$, $p < .08$. Finally, the correlation between frequency of identity changes and profit from selling commodities in the positive reputation condition was negative, $r = -.34$, *ns*., confirming our expectation that participants would not have incentives to change their identities. In the negative reputation condition, frequency of identity changes and profits from selling commodities was positive, though not significant, $r = .41$, *ns*.

Accuracy of Evaluations

Did participants honestly report their evaluation of the seller? This question was addressed by calculating the correlation between the level of dishonesty of the commodity the participant purchased and the evaluation score he or she attached to the seller of the commodity, for each participant. A negative within-participant correlation means honest evaluation. A correlation around zero means random or inaccurate evaluation. And, a positive correlation means an intentionally false evaluation. To avoid confusion, we reversed the sign of the correlation, so that high scores reflect honest evaluations. We call this reverse correlation the *index of honest evaluation*. Buyers in all three experiments evaluated the seller each time they made a purchasing decision, although evaluation scores were not displayed on the buyers' computer screens in the control condition and the identity condition of the first experiments. The first column of Table 1 reports the average index of honest evaluation. Except in the control condition of

the first experiment in which the market degenerated into a lemons market, the correlation was moderate to high. Participants reasonably honestly reported their evaluations of the honesty of the sellers from whom they purchased commodities.

Insert Table 1 About Here

Among the 156 participants[12] across three experiments, the index of honest evaluation exceeded .4 for 128 (82%) of the participants. More than 92% of them had index scores greater than .2. When the control condition and the identity condition are excluded, 94 of 100 participants had index scores greater than .2. These results clearly show that the overwhelming majority of the participants were honestly revealing their evaluations of the sellers.

The next question is whether or not honestly reporting evaluations affected the buyer/evaluators' earnings, and if it did, whether positively or negatively. All participants from the three experiments except the control and identity conditions of the first experiment were used for this analysis. The overall correlation between profit from purchasing and the honest evaluation index was negative (r = -.28, p < .01), and the correlation in each condition was also negative, suggesting that honest evaluators were disadvantaged in making profits from purchasing activities. Conversely, those who failed to make satisfactory profits in the purchasing deals were more strongly motivated to report their negative evaluations than those who were more successful. The overall correlation between honest evaluation and selling profit was positive, but non-significant (r = .10, *ns.*). Honest evaluators were not disadvantaged in the production/sales of commodities. The individual correlations varied widely across experiments and conditions. Finally, the overall correlation between honest evaluation and the total profit was slightly negative, but negligible (r = -.05, *ns.*). In addition, individual correlations varied across experiments and conditions. These correlations suggest that although

[12] In this analysis, each participant in the first experiment was counted twice, once in the 45-minute session and once in the 20-minute session.

honest evaluators did not make as much profit from purchasing deals as did less honest evaluators, by including profits from their selling, they made almost the same level of profit. It is interesting to note that in the positive reputation condition of the third experiment, honest evaluation was positively correlated with selling profit (r = .63, p < .05).

To assess how the two types of honesty, the first-order honesty of the seller for not cheating on the advertised quality of the commodity and the second-order honesty of the buyer for reporting true evaluations, relate to the profit the participant earns in the experiments, we run a multiple regression analysis in which the dependent variables are the three types of profit (total profit, profit from selling, and profit from purchasing), and the independent variables are the index of honest evaluation and the level of dishonesty. All participants except those in the control and identity conditions were used in this analysis. To control for the experiment and group differences, a set of dummy variables representing group differences were added to the list of independent variables. The results of this analysis show that the level of dishonesty has a ***negative*** effect on all three types of profit. The regression coefficient of the level of dishonesty was b = -42.72, $t(81)$ = 4.54, p < .0001 for total profit, b = -18.58, $t(81)$ = 5.50, p < .0001 for the profit from buying activities, and b = -24.14, $t(81)$ = 2.96, p < .001 for the profit from selling activities. While the effect of dishonesty reported earlier with regard to the analysis of each experiment was mixed, the overall picture emerging from this combined analysis is more promising. When reputation is shared, more-honest traders made *more* profit than less-honest traders. The results with respect to honest evaluation are not as rosy. The regression coefficient of the index of honest evaluation was -238.70, $t(81)$=1.02, *ns*. for total profit, -150.54, $t(81)$=1.81, p < .08 for the profit from buying activities, and -88.17, t(81) = .43, *ns*. for the profit from selling activities. However, as discussed earlier, the negative relationship between honest evaluation and profits can be a result of unsuccessful traders being more critical and resentful of dishonest sellers than successful

traders.

Table 1 also reports the correlation between the two types of honesty, the first-order honesty and the second-order honesty. The two kinds of honesty tend to be consistent. The overall correlation between the index of honest evaluation and the level of dishonesty was negative (r = -.31, p < .01), indicating that those who more honestly reported their evaluations tended to be honest sellers as well.

General Discussion

The messages we receive from the three experiments presented in this paper are clear. First, information asymmetry drives the experimental market among anonymous traders into a lemons market in which only the lowest quality goods are traded and thus opportunities to achieve better profits from trading high quality goods are forgone. Second, either experience-based information or reputation about other traders moderately alleviates the lemons problem when traders' identities are permanent. Third, the power of experience-based information or reputation as a solution to the problem of lemons is substantially reduced when traders can freely change their identities and cancel their reputations. Fourth, the negative reputation system that is designed to illuminate dishonest traders is particularly vulnerable to identity changes, whereas the positive reputation system designed to illuminate honest traders is not so vulnerable to identity changes. Overall, the results of the experiments point to the importance of properly designed reputation systems to resolve the lemons problem.

Despite the theoretical problem of free riding, the overall picture emerging from the experimental findings is that reputation systems provide decent success in curtailing the lemons problem. The first question to address is how the free rider problem was resolved. One answer points to a feature of the experiment. All participants were ***required*** to submit an evaluation score each time they made a purchase. That is, free riding was made impossible by the design of the experiment. However, participants did not need to report

accurate evaluations. Because of this, free riding was possible concerning submission of an ***accurate*** evaluation score. Keeping experience-based true information to oneself is more advantageous to the seller than sharing that information with other buyers. Furthermore, there were incentives to submit a false evaluation score. Lowering other traders' reputation by attaching them negative or low evaluation scores increases a trader's own relative standing in the market. The issue we addressed in the experiment was whether or not participants would honestly report their evaluations, and whether or not those who honestly report their experience-based information with others would be disadvantaged. The results of the experiments indicate, first, that most participants honestly reported their evaluations of the sellers from whom they purchased commodities. The correlation between honest evaluation and profit was not strong, and the interpretation is problematic. The negative correlation may indicate that honestly reporting evaluations of buyers made the player less successful in making profit, on the one hand. Yet the negative correlation may also indicate that those who were unsuccessful were more concerned with fairness and evaluated the buyers more seriously.

In an effort to design better reputation systems to resolve the lemons problem, our findings concerning the positive and the negative reputation systems provide valuable insights. Past research on Internet auction sites have pointed out stronger effects of negative reputation than positive reputation (e.g., Eaton, 2002; Houser and Wooders, 2000; Lucking-Reiley et al., 2000; Resnick and Zeckhauser, 2001; Standifird, 2001). At the same time, negative reputations are used less often than positive reputations (Resnick and Zeckhauser, 2001). Resnick and Zeckhauser (2001) suggest fear of retaliation as well as a courtesy help explain why Internet traders use more-effective negative reputations less often than less-effective positive reputations. That is, Internet traders provide positive evaluations to their trading partners to elicit similar positive reputations from them, and refrain from providing negative evaluations to avoid receiving retaliatory negative evaluations. Standifird (2001) attributes the stronger effect of negative

reputation to the gain-loss asymmetry in subjective utility (Kahneman and Tversky, 1979). The results of the third experiment, however, suggest a different reason why negative reputations are used less often. Negative reputations may be more effective than positive reputations in the short run, but can be less effective in the long run. In the third experiment, the average quality of commodities was higher in the negative reputation condition than in the positive reputation condition in the first time block. This initial strong effect of negative reputation may reflect the strong aversion of negative outcomes suggested by Standifird (2001). However, this advantage of the negative reputation system did not last long. Soon, the average quality in the negative reputation condition plummeted whereas the average quality in the positive reputation condition steadily improved. The effect of the negative reputation system was short-lived because it was neutralized by frequent identity changes. The effect of positive reputations, on the other hand, may take time to be realized, but it works in a cumulative manner. This occurs because identity changes are rare in the positive reputation system. The differential long-term effects of positive and negative reputation systems, in addition to the psychological factors discussed by Resnick and Zeckhauser (2001), may be behind the differential frequencies of the two types of reputations.

And finally, we are interested in the roles of positive reputation and negative reputation among Maghribi traders and Internet traders. The success of the Maghribi *coalition*, according to Greif's (1989) analysis, lies in the fear of exclusion. A Maghribi coalition member balances the immediate profit of behaving dishonestly with the risk of being excluded from the *coalition* (and from the future profit). Thus, it was predominantly ***negative reputation*** that was critical to the success of the *coalition*. Greif (1989) further argues that the success of negative reputation requires closure of the *coalition*. Exclusion from a market matters only when it is closed to non-members. This issue speaks to why negative reputation was not very effective in curtailing the lemons problem in the third experiment. The individuals who have acquired bad reputations can

freely “re-enter” the same market under a new identity. The central characteristic of Internet trading—that is, its openness—thus prevents negative reputation from exerting its power. The openness of online trading, on the other hand, promotes positive reputation as an effective means for curtailing lemons problem. To understand why positive reputation is useful in online trading, we need to realize the two functions of reputation: exclusion and inclusion. The power of negative reputations is based on the ***principle of exclusion***. Negative reputations exclude dishonest traders from the market. In contrast, the power of positive reputation is based on the ***principle of inclusion***. Positive reputations are not effective for excluding dishonest traders from the market. However, positive reputations are useful means to attract potential trading partners. Attracting new partners is not of central importance in a closed market since the membership of the market is limited. Establishing a good reputation in the Maghribi coalition does not help a Maghribi trader to expand his deals beyond the boundaries of the *coalition*. In sharp contrast to this, the number of potential trading partners is unlimited in an open market or in online trading. Therefore, the merit of obtaining good reputations in an open market is also unlimited. Established brand names are much more valuable, in this sense, in open markets than in closed markets.

In short, the closed nature of the coalition was a critical condition for the negative reputation system to control the lemons problem among Maghribi traders, as discussed by Greif (1989). The lack of a closed market among online traders, which appears, at first grant, to be a formidable problem, can actually be a blessing. To realize why positive reputations, rather than negative reputations, are effective in resolving the lemons problem in the experimental market we designed as well as in the real Internet auction sites (cf., Kollock, 1999), we have to turn to the principle of inclusion rather than the principle of exclusion. Put differently, although the principle of exclusion has been the central principle adhered to when examining the effect of reputations in markets, we find that the effect of reputations in ***open*** markets is best understood using the principle of

inclusion.

References

Akerof, G.
1970 “The market for lemons: Quality uncertainty and the market mechanisms.” The Quarterly Journal of Economics, 84: 488-500.

Avery, C., P. Resnick, and R. Zeckhauser
In press “The market for evaluation.” American Economic Review.

Axelrod, R.
1984 The Evolution of Cooperation. New York: Basic Books.

Axelrod, R.
1986 “An evolutionary approach to norms.” American Political Science Review, 80: 1095-1111.

Eaton, D. H.
2002 “Valuing information: Evidence from guitar auctions on eBay.” Working paper.

Friedman, E., and P. Resnick
2001 “The social cost of cheap pseudonyms.” Journal of Economic and Management Strategies, 10: 173-199.

Greif, A.
1989 “Reputation and coalitions in Medieval trade: Evidence on the Maghribi traders.” The Journal of Economic History, 49: 857-882.

Greif, A.
1993 “Contract enforceability and economic institutions in early trade: The Maghribi traders coalition.” American Economic Review, 83: 525-548.

Houser, D., and J. Wooders
2000 “Reputation in auctions: Theory, and evidence from eBay.” Working paper, Department of Economics, University of Arizona.

Kahneman, D., and A. Tversky
1979 “Prospect theory: An analysis of decisions under risk.” Econometrica, 47: 263-291.

Kollock, P.
1994 “The emergence of exchange structures: An experimental study of uncertainty, commitment, and trust.” American Journal of Sociology, 100: 313-345.

Kollock, P.
1999 “The production of trust in online markets.” In E. J. Lawler, M. Macy, S. Thyne, and H. A. Walker, Advances in Group Processes, Vol. 16: 99-123. Greenwich, CT: JAI Press.

Lucking-Reiley, D., D. Bryan, N. Prasad, and D. Reeves
2000 “Pennies from eBay: The determinants of price in online auctions.” Working draft.

Olson, M.
1965 The Logic of Collective Action. Cambridge, MA: Harvard University Press.

Resnick, P., and R. Zeckhauser
2001 “Trust among strangers in Internet transactions: Empirical analysis of eBay reputation system.” Unpublished manuscript.

Resnick, P., R. Zeckhauser, E. Friedman, and K. Kuwabara
2000 “Reputation systems: Facilitating trust in internet interactions.” Communications of ACM, 43: 45-48.

Standifird, S. S.
2001 “Reputation and e-commerce: eBay auctions and the asymmetrical impact of positive and negative ratings.” Journal of Management, 27: 279-295.

Yamagishi, T.
1986 “The provision of a sanctioning system as a public good.” Journal of Personality and Social Psychology, 51: 110-116.

Yamagishi, T., and N.Takahashi
1994 “Evolution of norms without metanorms.” In U. Schulz, W. Albers, and U. Mueller (eds.), Social Dilemmas and Cooperation: 311-326. Berlin: Springer-Verlag.

Yamagishi, T., and M. Yamagishi
1994 “Trust and commitment in the United States and Japan.” Motivation and Emotion, 18: 129-166.

Table 1 Average index of honest evaluation and the correlation between the index of honest evaluation and total profit, buy profit, sell profit, and the level of dishonesty

Expt	Experiment and Condition	Index of honest evaluation	Correlation with			
			Total profit[a]	Buy Profit[b]	Sell Profit[c]	Dis-honesty
1	45 min. Control	.39****	.20	-.38	.39	-.24
1	20 min. Control	.34***	-.41	-.40	-.25	-.59*
1	45 min. Identity	.68****	-.18	-.16	-.14	.19
1	20 min. Identity	.63****	-.69*	-.45	-.75	-.47
1	45 min. Reputation	.67****	.01	-.56*	.26	.27
1	20 min. Reputation	.73****	-.12	-.37	.07	-.20
2	Variable identity	.57****	-.31+	-.28	-.15	-.42*
3	Negative Reputation	.44****	-.14	-.24	-.08	-.36+
3	Positive Reputation	.66****	.51	-.18	.63*	.20
Combined (excluding control and identity conditions in Expt. 1)		.56****	-.05	-.28**	.10	-.31**

+ p <.10; * p < .05; ** p < .01; *** p < .001; **** p < .0001

[a] Total profit includes profit from producing and selling and profit from buying.

[b] Buy profit is 1.5 times the true value of the purchased commodity minus the price of the commodity.

[c] Sell profit is profit is the price minus the true quality level of the commodity.

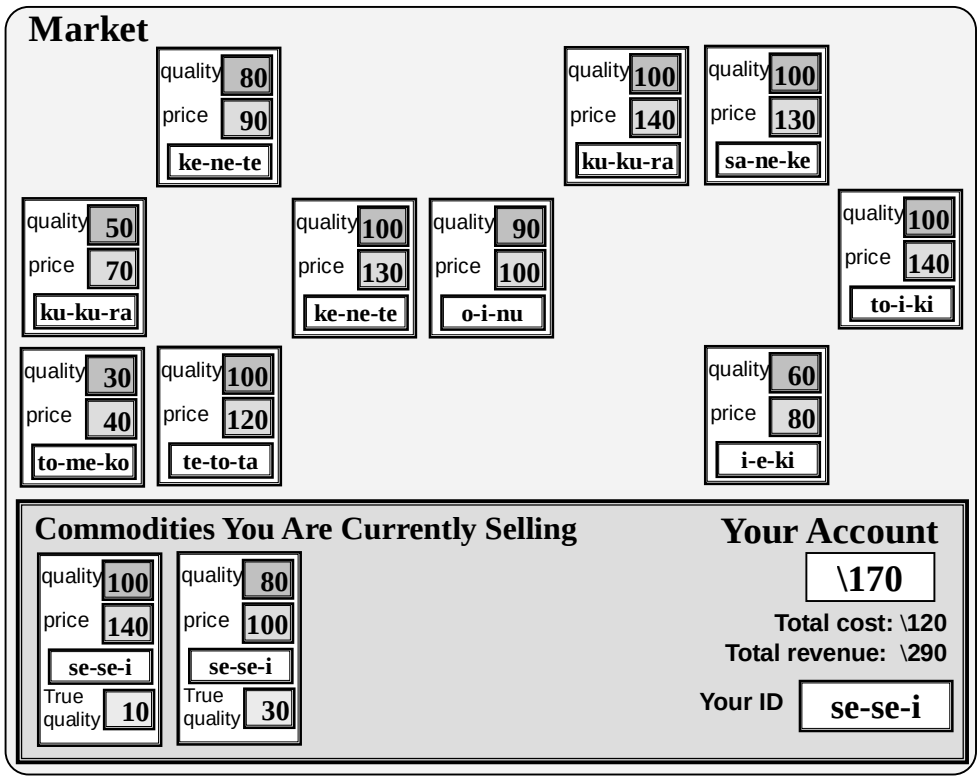

Figure 1 An image of the computer monitor of the identity condition of the first experiment. All characters are translated from Japanese.

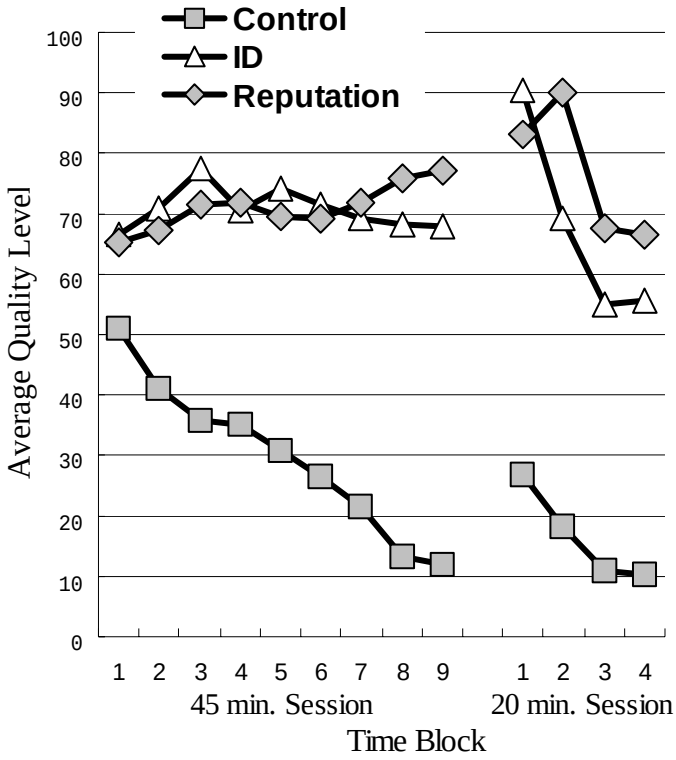

Figure 2 Average quality level of the produced commodities in Experiment 1.

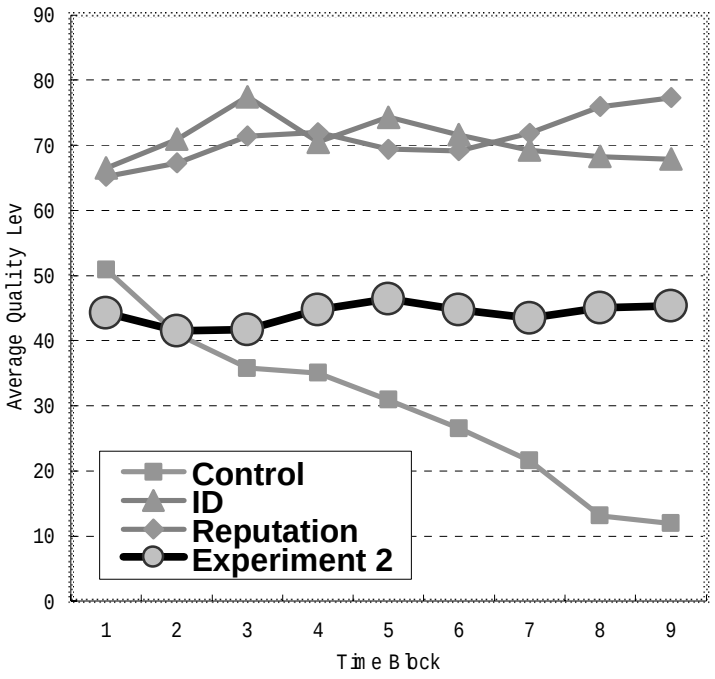

Figure 3 Average quality level in Experiment 2, compared to the levels in the three conditions in Experiment 1

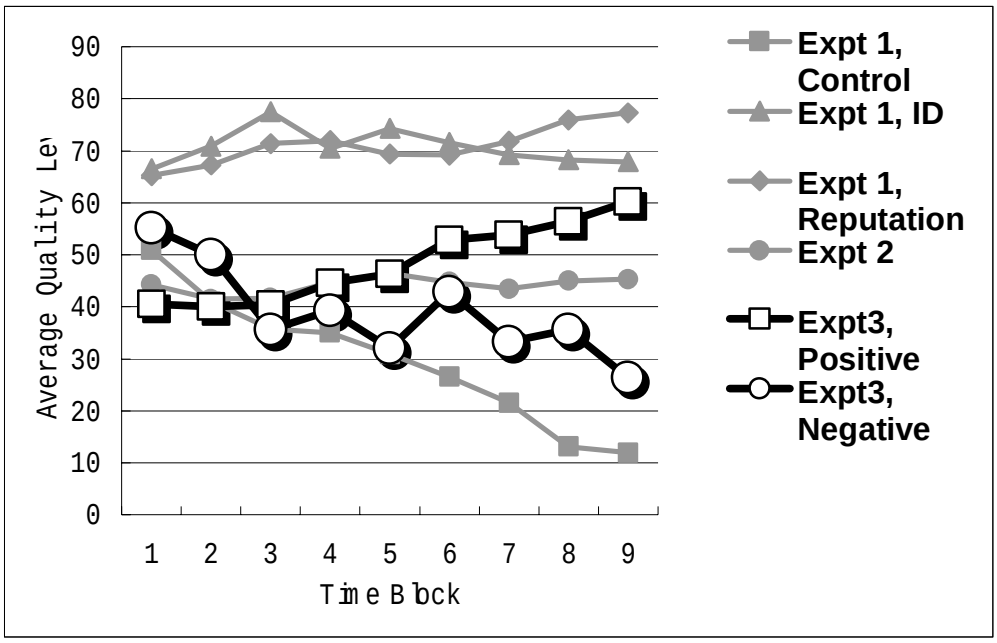

Figure 4 Average quality level in the positive reputation condition and the negative reputation condition in Experiment 2, compared to the levels in Experiments 1 and 2.