

An Experimental Study of Reputation with Heterogeneous Goods

Anya Savikhin*

*Becker Center on Chicago Price Theory, The University of Chicago Booth School of Business
5807 South Woodlawn Avenue, Chicago, IL 60637, U.S.A.

January 17, 2011

Abstract

We use experimental methods to study a market with heterogeneous goods under a reputation system. In the experiment, sellers choose to offer a high value or a low value good, buyers choose whether to purchase the good, and sellers choose whether to send the good to the buyer. Successful transactions in high value goods are more profitable than transactions in low value goods, but traders are unable to take advantage of this when there is no reputation system. The reputation system increases market efficiency, primarily by increasing transactions in the high value good. Allowing for heterogeneous goods reduces the information carried by the reputation system. To resolve this issue, we investigate a new reputation system, which displays reputation separately for each type of good. The results offer suggestive evidence that the additional information provided by this new system is utilized by buyers and sellers.

JEL Classifications: D02, D83

Keywords: reputation systems, experiments, markets

Email Address: savikhin@uchicago.edu

I am particularly grateful to Tim Cason for excellent guidance and support. I thank Jack Barron and Roman Sheremeta as well as seminar participants at Purdue University and participants at the June 2009 Economic Science Association meeting for helpful comments. I thank the International Foundation for Research in Experimental Economics and John Umbeck for generous funding of this project. Any remaining errors are mine.

1. Introduction

The growing popularity of e-commerce has increased the need for technologies that strengthen trust among strangers, reduce information asymmetries, and facilitate transactions when there is an opportunity to cheat (Quelch and Klein, 1996; Corbitt et al., 2002; Grabner-Krauter, 2003).¹ Reputation systems have emerged to fill this role, and have been implemented in several e-commerce sites such as eBay and Amazon Marketplace. The different design aspects of reputation systems, such as the amount and type of information to display, play an important role in determining the success of organizations that conduct business online (Josang et al., 2007). The design of these mechanisms has potentially important implications for a wide range of management activities, such as brand building, customer acquisition and retention and quality assurance (Dellarocas, 2003). Effective reputation systems that promote trust and encourage successful transactions will increase the prevalence of e-commerce activities (Ganesan, 1994; Jarvenpaa et al., 2000).

We contribute to the study of reputation systems through an experimental investigation of how markets function when heterogeneous goods are available to buyers. In practice, there is a substantial difference in value among goods in electronic marketplaces. The popular online electronic exchange site eBay, which holds 80% of the online auction market share, allows sellers to offer goods of widely varying values.² For example, with a search on eBay one can see an e-book for sale with a value of less than \$1.00, as well as a house for sale with a value above

¹ According to a press release by ComScore, a company that tracks the digital world, the popularity of e-commerce continues to increase rapidly. Black Friday (November 27) in 2009 saw \$595 million in online sales, an increase of 11% as compared to Black Friday 2008. During the 2009 holiday season, over \$10.57 billion dollars was spent online.

http://www.comscore.com/Press_Events/Press_Releases/2009/11/Black_Friday_Boasts_595_Million_in_U.S._Online_Holiday_Spending_Up_11_Percent_Versus_Year_Ago

² eBay reported approximately 559 million listings and gross merchandise transaction value of \$14.46 billion in the second quarter of 2007 alone. Figures are cited from eBay's 2007 second-quarter earnings release, available at http://investor.ebay.com/financial_releases.cfm.

\$300,000. Evidence from eBay suggests that when goods of different value are available, sellers engage in strategic reputation building, selling many low value goods to build reputation and then defaulting on high value goods (Bhattacharjee and Goel, 2005; Brown and Morgan, 2006). This necessitates the development of a new framework for understanding reputation systems with heterogeneous goods, which will guide organizational decisions about which reputation systems to employ.

Previous work has emphasized the necessity of trust and reputation in the context of e-commerce applications in business-to-business (Archer and Yuan, 2000; Ratnasingham and Kumar, 2000), business-to-consumer (Egger, 2000; Lee and Turban, 2001; Ranganathan and Ganapathy, 2002) and consumer-to-consumer interactions (Standiford, 2000; Zacharia and Maes, 2000; Pavlou and Gefen, 2004). This research is relevant for organizations that facilitate trade among anonymous agents in consumer-to-consumer e-commerce, such as eBay, Amazon Marketplace, and others. This research is also applicable to business-to-consumer e-commerce where users purchase products and rate quality and value, such as Amazon or Overstock.³

We conducted a laboratory experiment to study consumer-to-consumer transactions in a setting with heterogeneous goods. Laboratory experiments are a valuable tool for understanding different market mechanisms and provide more control than data available in the field. In our experiment, sellers choose to offer a high value or a low value good, buyers choose whether to purchase the good, sellers choose whether to send the good to the buyer, and in some treatments the reputation of the seller is automatically updated. We find that the market is less efficient when there is no reputation system because fewer high value goods are traded, but that efficiency is increased with a reputation system. Here efficiency is defined as level of surplus that the

³ See Josang et al., 2007 and Dellarocas, 2003 for an overview of the role that feedback plays on several such review sites, including eBay, eLance, Slashdot, Amazon, Overstock, BizRate, Epinions, and others.

participants of the market are able to extract relative to the total possible surplus. Successful transactions in high value goods are more profitable than transactions in low value goods, but traders are unable to take advantage of this when there is no reputation system.

The first reputation system we investigate is similar to that used in laboratory experiments with homogeneous goods. In a homogeneous good setting, this reputation system carries complete information about the seller's past behavior. However, with heterogeneous goods, the information carried by the reputation system is reduced. Specifically, buyers cannot tell from the reputation system whether the seller previously transacted in high or low value goods. Therefore, we investigate a second reputation system that displays reputation separately for each type of good. Previous experimental work has ignored this issue; in fact, this question can only be addressed in a setting such as ours. We do not find a significant effect on efficiency as compared to the first reputation system. However, our results offer suggestive evidence that the additional information provided by this new system is used advantageously by buyers and sellers.

An extensive literature in economics exists on the role of reputation in several contexts, such as the helping/trust game (Bolton and Ockenfels, 2000; Bolton et al., 2005; Engelmann and Fischbacher, 2009; Seinen and Schram, 2006; Ule et al., 2009), labor market (Healy, 2007; Holmstrom, 1981; Shapiro and Stiglitz, 1982), prisoner's dilemma (Kreps et al., 1982), and firm behaviors (Fudenberg and Tirole, 1985; Kreps and Wilson, 1982; Milgrom and Roberts, 1982). We use the buyer and seller context because it is most applicable to the e-commerce setting. In the management and information systems literature, reputation systems have been recognized as a way to increase trust (Pavlou and Ba, 2000; Snijders and Zijdeman, 2004), signal information (Melink and Alm, 2005), reduce information asymmetry (Resnick et al., 2006; Standifird, 2001;

Chua and Wareham, 2002), and act as a collaborative sanctioning system (Josang et al., 2007; Zacharia and Maes, 2000). Although some previous experiments have studied the role of seller reputations with heterogeneous goods (Cason and Gangadharan, 2002), all previous market experiments that have addressed reputation mechanisms explicitly employ a setting of homogeneous goods (Bolton et al., 2005; Dellarocas, 2004, 2005; Resnick and Zeckhauser, 2002; Resnick et al, 2000; Resnick et al., 2006). The general finding of related work is that the reputation system increases the number of successful transactions that take place, thereby unilaterally improving efficiency of the market. Many of these experiments analyze the value of various reputation mechanisms; however, this understanding may be limited in practical applicability because buyers and sellers may behave differently in the framework with heterogeneously valued goods.

Empirical evidence from eBay suggests that buyers and sellers behave differently when heterogeneous goods are available. In particular, sellers may engage in reputation building by selling many low value goods to increase reputation, and then selling high value goods and defaulting (Bhattacharjee and Goel, 2005; Brown and Morgan, 2006).⁴ eBay's recent changes to the cost structure made the practice of arranging for false transactions in order to gain reputation even cheaper and more feasible (Dini and Spagnolo, 2009). Anecdotal evidence suggests that some users of eBay know that this type of behavior is occurring. For example, a participant in a focus group commented, "For example I've seen people selling laptops with pretty high ratings and I go and read their feedback and the only thing they've sold is like clothes or very cheap goods, and they go for very high feedback" (Haywood, 2006).

⁴ Another way of building reputation on eBay is to transact many times as a buyer and then transact as a seller, because the reputation score is updated the same way for the buyer and the seller (Cabral and Hortacsu, 2004). However, this is no longer possible because eBay eliminated buyer reputations in 2008.

Diverse empirical findings have been reported for different types of goods, suggesting that different feedback systems must be designed to take such differences into account (Lee et al., 2000). Empirical work has found that reputation matters in eBay auctions for several different types of homogeneous goods, such as computer processors and guitars (Houser and Wooders, 2006; Eaton, 2005). Other work has found that reputation matters for heterogeneous goods, such as rare coins of varying qualities (Melink and Alm, 2005). Our contribution is an experiment that allows us to control many factors that cannot be controlled in empirical work, such as quality aspects of the good and communication between the buyer and seller that can occur online.⁵ The unique benefit of approaching this problem using an experimental methodology allows us to directly compare different reputation systems.

The reputation mechanisms used in most e-commerce sites do not account for the value of previous transactions, which are only available in the detailed reputation.⁶ For example, on eBay, buyers must click through several pages to read details about transactions to find their value and this is time consuming. It has been suggested that buyers may not look at the detailed reputation on eBay (Resnick et al., 2006).⁷ Empirical work has shown that buyers pay more attention to feedback when buying expensive products and less attention when buying cheaper products on eBay (Resnick et al., 2006). Therefore, it is important to investigate whether detailed feedback is a necessary aspect of an effective reputation system.

⁵ While empirical work can compare quality in “collector” auctions, such as for rare coins, even very experienced buyers may not always be fully aware of the quality of the good, or the quality may be perceived differently by different buyers. The experiment is the best way to abstract from differences in knowledge about quality for heterogeneous goods, by simply using goods that have a different experimental dollar value, commonly known to all.

⁶ Note that eBay calls the “reputation” the “feedback”

⁷ Accessing the detailed feedback requires the user to click through to a detailed feedback screen. The user must then spend time looking through pages of information with each individual transaction in order to get an idea of the value of goods transacted and the outcomes of these transactions.

2. Design and Predictions

2.1 *Experimental Environment*

The objective of this project is to investigate various reputation systems in a heterogeneous good setting. In our experiment, sellers are asked whether they would like to offer for sale a high value or a low value item.⁸ Each seller is only allowed to offer one item. Then buyers are shown a screen with each seller's offer and reputation score, and buyers enter one by one in random order to buy items.⁹ Once an item is purchased, the item disappears from the buying screen and no future buyer can purchase it. Buyers are given the choice to buy an available item from any human seller, or a choice to buy an item from the computer (also called the "outside option"), where the item offered by the computer is lower in value than either the high or low value item offered by any human seller.¹⁰ Buyers are only allowed to buy exactly one item. After all buyers make their decisions, sellers are given the choice of whether or not to send the good to the buyer (where not sending the good is referred to as "reneging" in subsequent data analysis). The computer always sends the item to the buyer. If no buyer bought his good, the seller skips this part and waits until the next period. The payoffs are chosen to assure that a sufficient volume of trade occurs in each type of good. The setup does not allow for price-setting for consistency with previous literature (Bolton et al., 2004). The decision tree is displayed in Figure 1.

[FIGURE 1 ABOUT HERE]

⁸ In the experiment, "purple" and "orange" are used instead of high or low value and "pink" is used instead of the item offered by the computer to avoid biasing subjects. We also exclude all context (no "example" items) to avoid introducing bias.

⁹ Note that the random order is determined by the computer program, and the buyers have no influence regarding the order in which they enter.

¹⁰ In the experiment, the items offered by the computer are displayed as being from 2 different computer sellers. Both computer sellers have unlimited quantities of this item, and both computer sellers always send the item if it is purchased.

Six sequences of nine periods each were conducted in every session. Using multiple sequences per session gives subjects time to learn and gain experience. The use of sequences was previously suggested in Camerer and Weigelt's (1988) work. The computer records a change in the seller's reputation, and automatically displays the seller's reputation scores on the buying screen in treatments with reputation. The reputation scores are cumulative for the duration of the sequence, but revert to the starting values at the beginning of each new sequence. Buyers do not have a reputation.

There are seven human sellers and four human buyers in each session. We chose this design for two reasons. First, we wanted to maximize the chance that there would be an excess supply of each type of item. In this way, buyers were not constrained to buying one item over another for reasons of lack of supply most of the time. Second, sellers essentially "drop out" of the market over time by reneging, as buyers no longer want to purchase items from these sellers.

The most efficient outcome (which maximizes the buyer and seller's joint payoff) occurs when the seller offers the high value item, the buyer buys the item, and the seller sends the item. For example, total surplus for four transactions of this type is $4 \times 110 = 440$. If instead, four transactions occur for the low value item, efficiency falls to $4 \times 95 = 380$, or $380/440 = 86\%$ of maximum possible efficiency.

2.2 Treatments

The treatments and sessions are summarized in Table 1. We consider three different treatments: *No Reputation*, *Simple Reputation* and *Separate Reputation*. The *No Reputation* treatment serves as a baseline. In the case when there is no reputation system, the buyer does not have information about any seller's past actions. To measure the benefit of reputation, this

treatment is compared with the *Simple Reputation* treatment, which uses a reputation system similar to one previously studied in the homogeneous good setting and generally used in e-commerce settings. In the *Simple Reputation* treatment, reputation is updated in the same way regardless of whether the transaction was for a high or low value item. Specifically, the buyer sees two numbers – the first number indicates how many times the seller has sent an item and the second number indicates the percentage of times the seller has sent an item (number of times seller has sent an item, divided by the number of times an item has been bought from the seller, whether the seller sent it or not).¹¹

In the *Separate Reputation* treatment, reputation is updated separately for the high value good and low value good transactions. The reputation mechanism carries information similar to the information each buyer would have in a partner matching setting. Buyers see separate numbers indicating outcomes of transactions in the high value market and the low market. Specifically, the buyer sees four numbers – the first two numbers indicate how many times and the percentage of times the seller has sent a high value item, and the second two numbers indicate how many times and the percentage of times the seller has sent a low value item. This treatment is compared to the *Simple Reputation* treatment to examine the effect of restoring information to a level comparable with information provided in experiments with homogeneous goods.

¹¹ Note that the electronic exchange site eBay has a similar reputation system, called the “feedback market”, where two reputation numbers are displayed based on buyers’ voluntary rating of transactions as “Positive,” “Negative” and “Neutral.” This is similar to eBay’s reputation system, which updates a number indicating the seller’s reputation and a number indicating the percentage of times the seller has received a positive feedback. On eBay, the reputation number can either increase by one point if the seller receives a positive feedback, or decrease by one point if the seller receives a negative feedback. Because we have a low volume of transactions, because the “percentage of items sent” carries similar information, and because we wanted the reputation to be as easy as possible to understand for subjects, we have simplified from eBay’s system. In our experiment, the number does not decrease when the seller receives a negative feedback; rather, because our number is called the “number of items sent”, it remains the same when the seller receives negative feedback. See Bajari and Hortacsu (2004) for a full description of the eBay marketplace and its reputation system.

[TABLE 1 ABOUT HERE]

2.3 Experimental Procedures

The experiment was conducted at the Vernon Smith Experimental Economics Laboratory. Volunteers were recruited from a pool of undergraduate students from Purdue University. A total of 132 subjects participated in 12 sessions, with 44 subjects participating in each treatment and 11 subjects participating in each session. Subjects were randomly divided into groups of sellers or buyers, and remained in that designation for the entire session. Each subject participated in only one session of this study, although some subjects had participated in other economics experiments that were unrelated to this research. Upon arrival, subjects drew a number out of a bingo cage that determined which computer station they would be sitting at. Computer stations were designated as buyer or seller stations prior to the experiment.

The computerized experimental sessions were run using z-Tree (Fischbacher, 2007). Each session proceeded in two parts. Subjects were given the instructions, available in the appendices, at the beginning of each part and the experimenter read the instructions aloud. In the first part, subjects' risk attitudes were elicited using a multiple price list of 15 simple lotteries; see Appendix III (Holt and Laury, 2002). Subjects were asked to state whether they preferred safe option A or risky option B.¹² The second part was the market experiment; see Appendix I. Subjects also completed a quiz to assure that all participants understood the rules of the market experiment; see Appendix II.

At the end of the experiment, three periods out of each sequence were randomly selected for payment using a bingo cage draw. Thus, subjects were paid for 18 total periods, with

¹² Option A yielded \$1 payoff with certainty, while option B yielded a payoff of either \$3 or \$0. The probability of receiving \$3 or \$0 varied across all 15 lotteries. The first lottery offered a 5% chance of winning \$3 and a 95% chance of winning \$0, while the last lottery offered a 70% chance of winning \$3 and a 30% chance of winning \$0. At the end of the experiment, 1 out of the 15 lottery decisions made by subjects was randomly selected for payment.

experimental dollars converted to U.S. dollars at the rate of 50 experimental dollars to 1 U.S. dollar. Subjects earned about \$25 on average for their participation in each session, which in addition to market earnings also included a \$5 show-up fee, \$1 per correct answer on a quiz about the rules of the experiment, and risk elicitation task earnings (\$1-\$3). The duration of each session was approximately 1.5 hours.

2.4 Hypothesis Development

2.4.1 Seller Types and Related Models of Behavior

The goal of this paper is to document the behavior that occurs and the effect on efficiency when trade in differently valued goods is possible. We do not aim to test alternative theoretical models; instead, we proceed by providing some basic predictions given the findings of previous work. The model typically used to motivate research in reputations is Kreps et al. (1982). It is well known that multiple equilibria exist in simpler settings than ours, including mixed strategy (Camerer and Weigelt, 1988) and the pure strategy “full reputation” equilibria (Healy, 2007). We predict that even though our setting is more complex, including endogenous matching of buyers and sellers and the choice between a high value, low value, or outside option item, we may observe many of the same outcomes.

Previous literature distinguishes types of sellers between those who prefer to send and those who prefer to renege, separate from reputational concerns (Camerer and Weigelt, 1988). Types are innate behavioral characteristics of each subject and can range from strictly egoistic to strong welfare preferences (Charness and Rabin, 2002; Fehr and Schmidt, 2001). We propose that some sellers have higher propensity to renege than others, and that the propensity to renege is dependent on the possible gain from renegeing. Necessitated by the additional complexity of

heterogeneously valued goods, we propose the existence of one more medium preference type. The range of preferences for subjects could be classified into Standard Preference (SP), Medium Preference (MP), and High Preference (HP) types. HP sellers prefer to send both the high and low value good, MP sellers prefer to send the low value good but renege on the high value good, and SP sellers prefer to renege on both the high and low value good.¹³

2.4.2 Behavior under “No Reputation”

When there is no reputation, each buyer does not know the history of actions of any seller. While only HP types send the good in the high value market, both HP and MP types send the good in the low value market. Therefore, the probability that the seller sends the good is greater in the low value market as compared to the high value market. The probability that the buyer buys the high value good is lower as than the probability that the buyer buys the low value good. Knowing this, it is optimal for sellers to offer low value goods as compared to high value goods. It is more optimal for buyers to buy the low value good than the high value good.

2.4.3 Behavior under Treatments with Reputation

A reputation mechanism acts as a signal about the type of seller one is dealing with (one who prefers to send, or one who prefers to renege). While sellers who prefer to send the item are able to signal their type via their reputation, sellers who prefer to renege may send anyway to avoid punishment from buyers. Buyers can take advantage of the reputation system, “punishing” sellers who do not maintain a high reputation by refusing to purchase items from these sellers.

A key characteristic of the framework with reputation is that sellers of the type who prefer not to send are aware that buyers may “punish,” and act as if they prefer to send during the

¹³ Intuitively, one can think of the HP types placing the greatest weight on their partner’s payoff, and the SP types placing the lowest (or no) weight on their partner’s payoff. MP types place some weight on their partner’s payoff such that not sending (renegeing on) the high value good is desirable since the own-payoff is very high for that action, but sending the low value good is desirable since the own-payoff is not high enough to justify the lowered payoffs for the partner.

sequence in order to build reputation and attract buyers in future periods. This idea of “false reputation building” has been studied extensively (e.g., Kreps et al., 1982; Camerer and Weigelt, 1988; Healy, 2007). According to the model of Camerer and Weigelt (1988), once a seller reneges once, he reveals himself as a type who prefers to renege and buyers do not buy from this seller in future periods. Because renegeing in either market results in no future purchases of the item, sellers’ renegeing in early periods is reduced for both goods under a reputation system.

If the buyer believes that the seller is altruistic, the buyer’s best response is to buy the good. Because buyers know that sellers are less likely to renege in both goods in earlier periods, and because a successful high value good transaction results in the greatest payoff, more transactions of the high value good should take place as compared to under no reputation. Under a reputation system, the optimal response for sellers who prefer not to send is to continue to send until the last period, and renege in the last period only (this is often called the “end-period” effect in related experiments with a finite number of decision rounds). While it is optimal to offer the high value good most periods, it is optimal to offer the low value good in the last period, since buyers know that HP and MP types will send the low value good, but only HP types will send the high value good. For the above reasons, the introduction of a reputation mechanism improves efficiency by increasing buyer and seller surplus. We posit the following four hypotheses. While Hypotheses 1 and 4 could be tested in a setup with homogeneous goods, Hypothesis 2 and 3 can only be tested using our heterogeneous good setup.

Hypothesis 1: With a reputation system, there will be a greater number of overall transactions relative to no reputation.

Hypothesis 2: With a reputation system, there will be a greater proportion of transactions in the high value good relative to no reputation.

Hypothesis 3: With a reputation system, there will be less renegeing relative to no reputation.

Hypothesis 4: With a reputation system, we will observe an “end period effect” with fewer high value offers, fewer purchases and greater renegeing for both goods as compared to earlier periods.

2.4.4 Heterogeneous Goods and Information

In a setting with heterogeneous goods, the simple reputation system does not carry full information about seller behavior. Under the simple reputation system, sellers signal their type with some uncertainty; that is, sellers with a good reputation score may either be MP or HP types. We introduce a new reputation system, which we suggest improves the value of the signal and reduces the potential for sellers to use successful low value transactions as a reputation-building tool for transactions in the high value market (as in Bhattacharjee and Goel, 2005; Brown and Morgan, 2006). Under the separate reputation system, a good reputation score in the high value market sends a clear signal to the buyer that the seller is an HP type, while a good reputation score in the low value market sends a signal to the buyer that the seller is at least an MP type. We propose that buyers will react to this signal by being more likely to buy a high value good from a seller with a good reputation in the high value market.

Hypothesis 5: Under the separate reputation system, buyers will be more likely to buy a high value good from a seller with a good reputation in the high value market.

In a setting with heterogeneous goods, the seller who renegees in the high value market does not exit the experiment entirely as in settings with one good; rather, he may continue by offering low value goods and attracting buyers in the low value market if buyers believe he may be an MP type. The separate reputation system assists buyers in identifying MP types after a seller renege. Under the simple reputation system, renegeing in either market identifies the seller as either an SP or MP type. Under the separate reputation system, renegeing in the high value

market identifies the seller as either an SP or MP type, and reneging in the low value market identifies the seller as an SP type.

Suppose that under the simple reputation system a seller reneges in either market. This identifies the seller as an SP or MP type and no buyer will buy a high value good from this seller in the future. Since there is some probability that the seller is an MP type, buyers may still buy the low value good from this seller. Reneging in either market has the same result for the seller.

Suppose that under the separate reputation system a seller reneges in the high value market. This identifies the seller as an SP or MP type. Because the buyer knows with certainty that the renege occurred in the high value market, there is a greater likelihood that the buyer will buy the low value good after a renege in the high value market in separate reputations versus in simple reputations. Suppose that under the separate reputation system an SP type seller reneges in the low value market. This identifies the seller as an SP type, and no buyer will buy from this seller in the future. Under the separate reputation system, reneging in the low value market results in no future trade for the seller in either market.

Hypothesis 6: Under the simple reputation system, buyers may continue to buy low value goods from sellers with an imperfect reputation score; under the separate reputation system, reneging in the low value market results in no future trade for the seller in either market.

3. Results

Sub-sections 3.1-3.2 provide an analysis and discussion of the results from the *No Reputation* and *Simple Reputation* treatment, with *Separate Reputation* results relegated to the footnotes. Section 3.3 explores seller offer and buyer purchase decisions in more detail for all

treatments. Section 3.4 provides a discussion of the suggestive differences between *Simple Reputation* and *Separate Reputation*.

3.1 Overview

The results address the important question of whether transactions in the high versus low value good are affected in the same way by the introduction of reputation systems. This question has not been addressed in previous work. We find that individuals respond to the reputation system differently for high as compared to low value goods. As predicted in Hypothesis 1, we find that there are a greater number of overall transactions in the treatments with reputation relative to the treatment without reputation. This is in line with related work. As predicted in Hypothesis 2, we find that there are a greater proportion of transactions in the high value good in the treatments with reputation relative to the treatment without reputation. Without a reputation system, most transactions occur in the low value market, while with a reputation system, most transactions occur in the high value market. This is due to both the increase in seller offers and buyer purchases of the high value good in treatments with a reputation system.

Figure 2 shows the number of high and low value goods offered by sellers in the *No Reputation* and *Simple Reputation* treatments. Under no reputation, the high value good is offered a total of 254 times (27% of all offers), while under *Simple Reputation*, the high value good is offered a total of 648 times (69% of all offers).¹⁴

[FIGURE 2 ABOUT HERE]

Figures 3 and 4 show the number of each good purchased and the proportion of goods purchased among those that were offered, respectively. A greater number of goods overall are purchased in the *Simple Reputation* treatment relative to the *No Reputation* treatment (Wilcoxon

¹⁴ Under *Separate reputation*, the high value good is offered 570 times (60% of all offers).

Mann-Whitney p -value < 0.01).¹⁵ In aggregate, a greater proportion of available high value goods are bought in the *Simple Reputation* treatment relative to the *No Reputation* treatment. The high value good is bought 428 times (49% of all chances to buy) in the *Simple* reputation treatment, while it is bought only 26 times (9% of all chances to buy) in the *No Reputation* treatment. However, a similar proportion of available low value goods are bought in both treatments. The low value good is bought 142 times (36% of all chances to buy) in the *Simple* reputation treatment, while it is bought 328 times (34% of all chances to buy) in the *No Reputation* treatment.¹⁶ Despite the similar proportions, due to the level of offers, fewer low value goods are purchased in the *Simple* reputation treatment relative to the *No Reputation* treatment (Wilcoxon Mann-Whitney p -value < 0.01).¹⁷

Result 1: *Buyers are buying more goods from human sellers overall in the Simple treatment as compared to the No Reputation treatment.*

Result 2: *Sellers are offering and buyers are buying more of the high value good in the Simple treatment as compared to the No Reputation treatment.*

[FIGURE 3 ABOUT HERE]

[FIGURE 4 ABOUT HERE]

Figure 5 displays the proportion of reneges of all purchased goods and Table 2 provides total numbers of reneges for all periods and for “early” periods (periods 1-7). As predicted in Hypothesis 3, there is significantly more renegeing in both goods in the *No Reputation* treatment (65 total reneges) than in the *Simple* treatment (26 total reneges) (Wilcoxon

¹⁵ We aggregated the number of times each buyer bought any good from a seller (not from computer seller) across sessions 2-6, and conducted a non-parametric Wilcoxon rank-sum test. . When comparing *Simple* to *No Reputation* or *Separate* to *No Reputation*, we find statistically significant differences (p -values = 0.00 and 0.01, respectively).

¹⁶ The high value good is bought 402 times (48%) and the low value good is bought 137 times (32%) in the *Separate* reputation treatment – these numbers are similar to the *Simple* reputation treatment.

¹⁷ We aggregated the number of times each buyer bought the low value good across sessions 2-6, and conducted a non-parametric Wilcoxon rank-sum test. . When comparing *Simple* to *No Reputation* or *Separate* to *No Reputation*, we find statistically significant differences (both p -values = 0.01).

Mann-Whitney rank sum p -value = 0.00).¹⁸ The frequency of renegeing of high value goods looks highly volatile in the *No Reputation* treatment on the graph because very few high value goods are bought in this treatment.

Result 3: *There is greater renegeing in No Reputation versus in Simple Reputation.*

[TABLE 2 ABOUT HERE]

[FIGURE 5 ABOUT HERE]

Figure 6 displays the efficiency of the *Simple* and *No Reputation* treatments.¹⁹ Efficiency is calculated as a proportion, taking the total experimental dollars earned in the market in any period and dividing this total by the maximum total surplus that would have been achieved if all sellers offered the high value item, all buyers bought a high value item, and all sellers sent their items. Efficiency is greater in the treatments with reputation as compared to the treatment without reputation (Wilcoxon rank-sum p -value 0.00 for both tests).²⁰

Result 4: *There is significantly greater efficiency in the treatments with reputation as compared to the treatment without reputation.*

[FIGURE 6 ABOUT HERE]

We also observe an “end-period effect” in seller offers, buyer purchases, and seller renegees in the *Simple* treatment. As evidenced by Figure 3, sellers substitute offers of high value goods for low value goods in later periods (while 73% of offers in Period 1 are for the high value

¹⁸ We aggregated the number of renegees for each seller across sessions 2-6 separately for the high and low value good, and a Wilcoxon Mann-Whitney rank-sum test finds statistically significant differences in high value good renegees between *Separate* and *No Reputation* and *Simple* and *No Reputation* (p -values 0.02 and 0.00, respectively) and statistically significant differences in low value good renegees (both p -values 0.01). Aggregating by session, we find similar results – there are statistically significant differences between *Separate* and *No Reputation* (p -values 0.00 and 0.05 for high and low, respectively) and between *Simple* and *No Reputation* (p -values 0.09 and 0.00, for high and low, respectively).

¹⁹ See Figure 4.7 for a comparison of the *Simple* and *Separate* efficiencies

²⁰ We aggregated efficiency across each session for sessions 2-6 and conducted a Wilcoxon rank-sum test, efficiency is statistically significantly greater in *Separate* and *Simple* as compared to *No Reputation* (both p -values 0.00).

good, this proportion drops to 58% in Period 9).²¹ In the *No Reputation* treatment, this effect is not as great (31% of offers in Period 1 are for the high value good, and 25% of offers in Period 9 are for the high value good). As displayed by Figures 3 and 4, this “end period” effect exists for buyers as well. In *Simple Reputation*, 87% of buyers select an item offered by a human seller in period 1, compared to 8% of buyers in period 9. In *No Reputation*, 53% of buyers select an item offered by a human seller in period 1, compared to 42% of buyers in period 9.²² Finally, there is an end-period effect of renegeing in the last period in the *Simple Reputation* treatment, which is not observed in the *No Reputation* treatment (Wilcoxon Mann-Whitney rank-sum p -value = 0.01).²³ While the “end-period effect” has been observed in related work, the shift from offers of high value goods to offers of low value goods in final periods could only be investigated in our heterogeneous good design. Also, as evidenced by Figure 6, efficiency declines over time in the *Simple* treatment but does not have as great of a decline in the *No Reputation* treatment.

Result 4: *There is an “end-period effect” of fewer high value offers, fewer overall purchases, and more reneges in the last period for the treatments with reputation.*

Result 4 suggests that “false reputation building” may be a good explanation for behavior. In treatments with reputation, types who prefer to renege continue to send until the last period. Knowing this, buyers are less likely to purchase items from human sellers in the last

²¹ In *Separate*, 75% of offers are high value in period 1, compared to 47% in period 9.

²² In *Separate*, 57% of buyers select an item offered by a human seller in period 1, compared to 17% in period 9.

²³ We aggregated renegeing across sessions 2-6 for each subject both for the first 7 periods and for the “end periods”, periods 8 and 9. Even considering the greater number of periods in the non-end-period group and the fewer goods bought in the end periods, there is statistically significantly more renegeing in *Separate* and *Simple* in periods 8 and 9 as compared to all of periods 1-7 combined, a Wilcoxon Mann-Whitney rank-sum test finds statistically significant differences (p -values 0.06 and 0.01 for the high value good and p -values of 0.12 and 0.04 for the low value good, for respective treatments). No such statistically significant differences exist for the *No Reputation* treatment (p -values 0.48 and 0.56, for high and low value goods, respectively). Similar results are found when aggregating for each sequence rather than for each subject.

period. In the treatment without reputation, subjects do not renege more in the last period than in any other period.²⁴

3.2 Detailed Analysis of Buyer and Seller Behavior

To investigate seller offer decisions further, we estimated a probit model (with subject random effects) for each treatment with seller offer decisions as the dependent variables, and the results are summarized in Table 3. We use a dummy variable that is equal to 1 if the seller's reputation (percentage of items sent) is 100%. As predicted in Hypothesis 4, for the *Simple* and *Separate* treatments, an end period effect is present and significant – causing the choice to offer a high value good to decrease in periods 8 and 9. Because we do not predict that an end-period effect exists under no reputation, we do not include dummies for periods 8 and 9 for those regressions.²⁵ Having a buyer in period $t-1$ increases the likelihood of offering a high value good in period t in the treatments with reputation, but actually has a negative effect on offering the high value good in the treatment without reputation. This suggests that sellers respond to their reputation score in each type of market differently. Also, under *Separate* reputation, having an “unblemished” reputation of 100% in the high value market results in the seller being more likely to offer the high value item, while having an “unblemished” reputation of 100% in the low

²⁴ Since reputation no longer plays a role in the last period of the treatments with reputation, buyers can no longer utilize reputation to “punish” sellers who do not send the item. In this case, sellers who preferred to send continue to send, while buyers who preferred not to send under no reputation system renege. Thus, behavior in the last period should mimic the behavior under no reputation, where every period is like the last period of reputation treatments. However, we find that in the *Simple* and *Separate* treatments, the end period effect results in lower efficiency in the last period even as compared to any period in the *No Reputation* treatment. This may suggest that the initial existence of the reputation system acts to “crowd out” inherent trust in sellers. We conjecture that this finding may be due to the change in perception that buyers have about sellers when a reputation system is present. For example, buyers in the treatment with a reputation system may believe that it is needed because all sellers defect, while buyers in the treatment without a reputation system may not hold this belief since the experiment instructions suggest that buyers can trade with sellers over the course of the experiment.

²⁵ Including period 8 and 9 dummies in specification 3 results in qualitatively similar estimates for the remaining coefficients, while the period 8 and 9 dummies are not statistically significant, which is in line with the prior prediction that there should not be an “end period effect.”

value market results in the seller being less likely to offer the high value item. More risk averse subjects are less likely to offer the high value item, but this is only statistically significant in specifications (2) and (3).²⁶

[TABLE 3 ABOUT HERE]

To investigate buyer purchase decisions, we estimated a probit model with random effects for each treatment with buyer's purchase choices as the dependent variables, and the results are summarized in Table 4. As predicted in Hypothesis 4, significant end-period effects are observed in the treatments with reputation: buyers are less likely to purchase the high value good and more likely to purchase the outside option in periods 8 and 9. This effect is consistent with the interpretation that buyers are aware of "false reputation building." Because we do not predict that an end-period effect exists under no reputation, we do not include dummies for periods 8 and 9 for those regressions.²⁷ Availability of the low value good decreases the probability of purchasing a high value good, and this is statistically significant for the treatments with reputation. Availability of a high value good decreases the likelihood of purchasing a low value good, and this is significant in the *Separate* treatment.

[TABLE 4 ABOUT HERE]

Having a seller send one's item in $t-1$ increases the likelihood of purchasing a high value good in t for treatments with reputation, and decreases the likelihood of purchasing the outside option in all treatments. For the treatment without reputation, buyers learn to decrease purchasing the high value good over time but are influenced to buy more of the low value good

²⁶ We elicited risk preferences prior to the experiment using the Holt and Laury (2002) multiple lottery selection task. In this task, subjects choose the "risky" lottery or the "safe" sure payment for each of 15 different choices, ranging from risky to less risky. We use the variable "# safe options" which is the number of times that subjects choose the "safe" option.

²⁷ In fact, another probit model was run with the inclusion of Periods 8 and 9, and it was found that these are not statistically significant.

when the seller sent them a good in $t-1$. If the seller has a reputation of 100%, the buyer is more likely to buy any item from the seller under the simple reputation system.

We find support for Hypothesis 5 – buyers do react differently to the different reputation scores under the separate reputation system. Under the separate reputation system, a 100% reputation in the high value good positively influences buying the high value good, while a 100% reputation in the low value good positively influences buying the low value good, suggesting that buyers do respond to the different reputation system by linking the seller's reputation to the specific task.²⁸ Risk attitudes, which were elicited in a pre-experiment task, do not have a strong influence on the buying decision.

Result 5: Under the separate reputation system, buyers are more likely to buy a high value good from a seller with a good reputation in the high value market.

3.3 Simple versus Separate Reputation

On aggregate, similar buyer behavior is observed in the *Separate* versus *Simple* treatments. No statistically significant differences in seller offer frequencies, buyer purchase decisions, or seller renege decisions are observed across the two treatments.²⁹ The proportion of renegeing was not statistically significantly different in either treatment.³⁰ Efficiency is not

²⁸ Interestingly, a 100% reputation in the low value good actually statistically significantly reduces buying of the high value good.

²⁹ We aggregated the number of times each seller offered the high value item across sessions 2-6, and conducted a non-parametric Wilcoxon rank-sum test. Comparing *Simple* to *Separate*, there is no statistically significant difference (p -value 0.34). In regards to buyer purchase decisions, comparing *Simple* to *Separate*, there is no statistically significant difference (p -value 0.70). No statistically significant differences in renege rates are found between *Simple* and *Separate* (p -values 0.39, high value; 0.79, low value). Similar results are found when aggregating for each sequence rather than for each subject.

³⁰ We aggregated the proportion of reneges in each treatment by sequence and found that there is not a statistically significant difference in renege proportions in the high value market (Wilcoxon rank-sum p -value 0.36) or in the low value market (p -value 0.99).

significantly different across the two treatments (p -value 0.90).³¹ Given that the simple reputation system was very effective at increasing successful transactions, and especially successful transactions in the high value market, it is possible that this is partially due to the lack of room for improvement (average efficiency across the 9 rounds was 42% under no reputation, and increased to 72% under simple reputation and 73% under separate reputation.)

Table 5 summarizes renegeing and subsequent ability to successfully transact for the treatments with reputation. Consistent with our predictions, attracting a future buyer following a renege is unlikely (overall, 82% of renegees resulted in no future buys). Buying in the low value market after a high value market renege is more likely than buying in the high value market after a renege. We also observe that attracting a future buyer following a renege is not very likely under either reputation system; however, following a renege in the low value market, several buyers continue to buy under the *Simple* reputation system but no buyers continue to buy under the *Separate* reputation system. This is consistent with our prediction that there is some uncertainty about seller types in *Simple*, but less uncertainty about seller types in *Separate*.

[TABLE 5 ABOUT HERE]

Figure 7 displays renegeing percentages under each reputation system. Recall that sellers renegeing on the low value good under the simple reputation system may be identified as an MP or SP type, while sellers renegeing on the low value good under the separate reputation system will definitely be identified as an SP type. As predicted, sellers renege less on the low value good under the separate reputation system as compared to the simple reputation system, although these differences are minor (0% versus approximately 2% in periods 1-7). Theory predicts that under the separate reputation system it is never optimal to renege on the low value good in the

³¹ A Wilcoxon rank-sum test using efficiency in each session finds that efficiency is not different between *Separate* and *Simple* (p -value 0.90).

penultimate period, and we find a 0% renege rate in period 8 in *Separate*. However, theory suggests that under the simple reputation system, the seller may renege on the low value good in period 8, believing that buyers may be uncertain about his type in period 9. Consistent with this prediction, we find a renege rate of 16% in period 8 in *Simple*. The above analysis provides suggestive evidence to support Hypothesis 6, but statistical analysis is not feasible due to the limited number of renegees in this study. Future work is needed to explore these issues further.

[FIGURE 7 ABOUT HERE]

4. Conclusion

We study the functioning of reputation systems in heterogeneous good markets with anonymous agents who have the opportunity to cheat. Our finding is that when there is no reputation system, subjects do not trade sufficient quantities of the high value good. We also find that the reputation system affects transactions in high and low value goods differently. The reputation system is valuable, because efficiency is increased when a reputation system is present. Efficiency is increased for three reasons. First, a greater number of high value goods as compared to low value goods are offered under a reputation system as compared to under no reputation system. Second, a greater proportion of the high value goods are bought relative to low value goods under a reputation system. Third, less renegeing occurs under a reputation system in both goods. Renegeing results in lower efficiency in the current period. Renegeing under a reputation system also leads to a lack of buyers in future periods for that seller, resulting in lower efficiency in subsequent periods.

These results suggest that reputation systems are especially important for organizations that derive profits from sales of high value or quality goods, as customers are more likely to buy high value goods as compared to low value goods when a reputation system is available. An

important concern for e-commerce activity is customer retention (Bhattacharjee, 2001; Srinivasan et al., 2002). Because customers may exit the market after encountering a seller who defaulted, reputation systems that facilitate successful transactions are important for customer retention, especially for high valued goods.

The different design aspects of reputation systems, such as the amount and type of information to display, play an important role in determining the success of organizations that conduct business in electronic exchanges (Dellarocas, 2001; Josang et al., 2007). We also introduced a new reputation system, which displays reputation separately for each type of good, thereby increasing information. We find that efficiency is unchanged as compared to the simple reputation system, although there is suggestive evidence that this reputation system is valuable. First, buyers do respond to the new reputation system, being more likely to purchase high value goods from sellers with 100% high value reputation, and being more likely to purchase low value goods from sellers with 100% low value reputation. Second, the separate reputation system improves the identification of seller types, and we find suggestive evidence that buyers use the system for precisely this reason.

An extension of this study is to set up new experiments in which renegeing is more likely in both markets in order to further investigate differences between the simple and separate reputation systems. In addition, buyers incur a small implicit cost (time and effort) when viewing and posting a feedback – future research can investigate the impact of this cost on voluntary feedback provision in a heterogeneous good setting.

References

- Archer, N., & Yuan, Y. (2000). Managing business-to-business relationships throughout the e-commerce procurement life cycle. *Internet Research: Electronic Networking Applications and Policy*, 10(5), 385-395.
- Bajari, P., & Hortacsu, A. (2004). Economic insights from internet auctions. *Journal of Economic Literature*, 457-486.
- Bhattacharjee, R. & Goel, A. (2005). Avoiding Ballot Stuffing in eBay-like Reputation Systems. *SIGCOMM '05 Workshops*.
- Bhattacharjee, A. (2001). An empirical analysis of the antecedents of electronic commerce service continuance. *Decision Support Systems*, 32(2), 201-214.
- Bolton, G, Katok, E. & Ockenfels, A. (2005). Cooperation among strangers with limited information about reputation. *Journal of Public Economics*, Vol 89, 1457-1468.
- Bolton, G. & Ockenfels, A. (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *The American Economic Review*, Vol. 90, No. 1, 166-193.
- Brown, J., & Morgan, J. (2006). Reputation in online markets: The market for trust. *California Management Review*, 49(1), 61-81.
- Cabral, L. M. B., & Hortacsu, A. (2004) The dynamics of seller reputation: Theory and evidence from ebay. NBER Working paper.
- Camerer, C. & Weigelt, K. (1988). Experimental Tests of a Sequential Equilibrium Reputation Model. *Econometrica*, Vol. 56, No. 1, 1-36.
- Cason, T. N., & Gangadharan, L. (2002). Environmental labeling and incomplete consumer information in laboratory markets. *Journal of Environmental Economics and Management*, 43(1), 113-134.
- Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3), 817-869.
- Chua and Wareham, (2002), "Self-regulation for online auctions: an analysis," *Twenty-third international conference on information systems*.
- Corbitt, B.J., Thanakankit, T., & Yi, H. (2003). Trust and e-commerce: a study of consumer perceptions. *Electronic Commerce Research and Applications*, 2(3), 203-215.
- Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, 49(10), 1407-1424.
- Dellarocas, C. (2006). Reputation Mechanisms. *Economics and Information Systems*, 629.
- Dellarocas, C. Fan, M. and Wood, C. (2004). Self-Interest, Reciprocity, and Participation in Online Reputation Systems. *Center for eBusiness at MIT, Paper 205*.
- Dini, F., & Spagnolo, G. (2009). Buying reputation on eBay: Do recent changes help? *Int. J. Electronic Business*, 7(6), 581.
- Eaton, D. H. (2005). Valuing information: Evidence from guitar auctions on eBay. *Journal of Applied Economics and Policy*, 24(1), 1-19.
- Egger, F.N. (2000). "Trust me, I'm an online vendor": Towards a model of trust for e-commerce system design. *Conference on Human Factors in Computing Systems*.
- Engelmann, D. & Fischbacher, U. (2009). Indirect Reciprocity and Strategic Reputation Building in an Experimental Helping Game. *Games and Economic Behavior*.
- Fehr, E., Schmidt, K. M., & Str, P. (2001). Theories of fairness and reciprocity-evidence and economic applications. *Econometric Society Monographs*.

- Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.
- Fudenberg, D. and Tirole, J., (1985). Predation Without Reputation, Working papers 377, *Massachusetts Institute of Technology (MIT), Department of Economics*.
- Ganesan, S. (1994). Determinants of long-term orientation in buyer-seller relationships. *The Journal of Marketing*, 58(2), 1-19.
- Gefen, D., & Traub, D.W. (2004). Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *Omega*, 32(6), 407-424.
- Gneezy, U. R. I., & Rustichini, A. (2000). A FINE IS A PRICE. *Journal of Legal Studies*, 29.
- Haywood, A. J. (2006). Online auctions: User experience insights from Ebay. Working Paper.
- Healy, P.J. Group Reputations, Stereotypes, and Cooperation in a Repeated Labor Market. *American Economic Review*, Vol. 97, Issue 5, 1751-1773.
- Holmstrom, B. (1981) Contractual Models of the Labor Market. *The American Economic Review*, Vol. 71, No. 2.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 1644-1655.
- Houser, D., & Wooders, J. (2006). Reputation in auctions: Theory, and evidence from eBay. *Journal of Economics & Management Strategy*, 15, 353-369.
- Jarvenpaa, S.L., Tractinsky, N., & Vitale, M. (2000). Consumer trust in an internet store. *Information technology and management*, 1(1), 45-71.
- Josang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2), 618-644.
- Josang, A. (2007). Trust and reputation systems. *Foundations of security analysis and design IV*, 209-245.
- Kreps, D. & Wilson, R. (1982). Reputation and Imperfect Information. *Journal of Economic Theory*, 27, 253-279.
- Kreps, D., Paul M., Roberts, J. & Wilson, R. (1982). Rational Cooperation in the Finitely Repeated Prisoners' Dilemma. *Journal of Economic Theory*, 27, 247-252.
- Lee, M.K. & Turban, E. (2001). A trust model for consumer internet shopping. *International Journal of Electronic Commerce*, 6(1), 75-91.
- Lee, Z., Im, I., & Lee, S. J. (2000). The effect of negative buyer feedback on prices in internet auction markets. *Proceedings of the Twenty First International Conference on Information Systems*.
- Melink, M. I., & Alm, J. (2005). Seller Reputation, Information Signals, and Prices for Heterogeneous Coins on eBay. *Southern Economic Journal*, 72(2), 305-328.
- Milgrom, P. and Roberts, J. (1982). Predation, Reputation, and Entry Deterrence. *Journal of Economic Theory*, 27, 280-312 (1982).
- Pavlou, P.A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Information Systems Research*, 15(1), 37-59.
- Quelch, J.A., & Klein, L.R. (1996). The internet and international marketing. *Sloan Management Review*, 37(3), 60.
- Ranganathan, C., & Ganapathy, S. (2002). Key dimensions of business-to-consumer web sites. *Information and Management*, 39(6), 457-465.
- Ratnasingham, P., & Kumar, K. (2000). Trading partner trust in electronic commerce participation. *Proceedings of the International Conference on Information Systems (ICIS)*.

- Resnick, P. & Zeckhauser, R. (2002). Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. *Advances in Applied Microeconomics*, Volume 11, pp. 127-157.
- Resnick, P., Zeckhauser, R., Swanson, J., & Lockwood, K. (2006). The value of reputation on ebay: A controlled experiment. *Experimental Economics*, 9(2), 79-101.
- Seinen, I. & Schram, A. (2006). Social Status and Group Norms: Indirect Reciprocity in a Helping Experiment. *European Economic Review*, Volume 50, Issue 3, pp. 581-602.
- Shapiro, C. and Stiglitz, J. (1984). Equilibrium Unemployment as a Worker Discipline Device. *The American Economic Review*, 74(3), pp. 433-444.
- Snijders and Zijdeman, (2004), "Reputation and Internet Auctions: eBay and Beyond", *Analysis & Kritik*, 26, 258-184.
- Srinivasan, S.S., Anderson, R., & Ponnavaolu, K. (2002). Customer loyalty in e-commerce: an exploration of its antecedents and consequences. *Journal of Retailing*, 78(1), 41-50.
- Standifird, S.S. (2001). Reputation and e-commerce: eBay auctions and the asymmetrical impact of positive and negative ratings. *Journal of Management*, 27(3), 279.
- Ule, A., Schram, A., Riedl, A. and Cason, T. (2009). Indirect Punishment and Generosity Toward Strangers. *Science*, 326, 1701.
- Zacharia, G., & Maes, P. (2000). Trust management through reputation mechanisms. *Applied Artificial Intelligence*, 14(1), 881-907.

Tables and Figures

Figure 1 – Decision Tree for Each Period

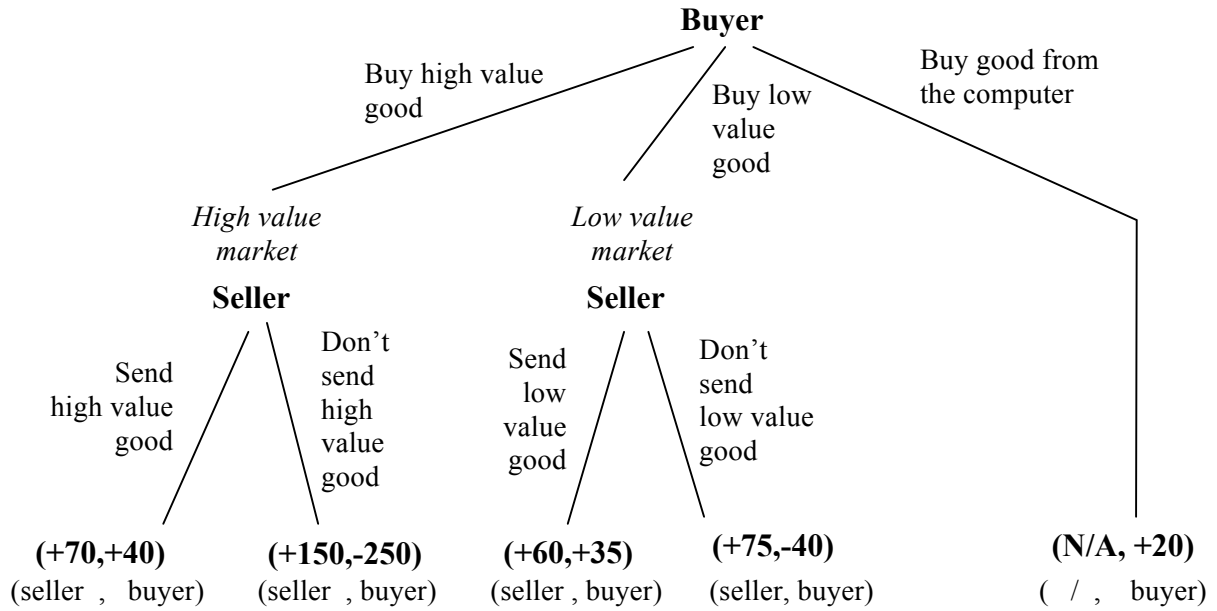


Figure 2: Seller's Choice of Good Offered over Periods 1-9, Sequences 2-6 Aggregated



Figure 3: Buyer's Choice of Good over Periods 1-9, Sequence 2-6 Aggregated

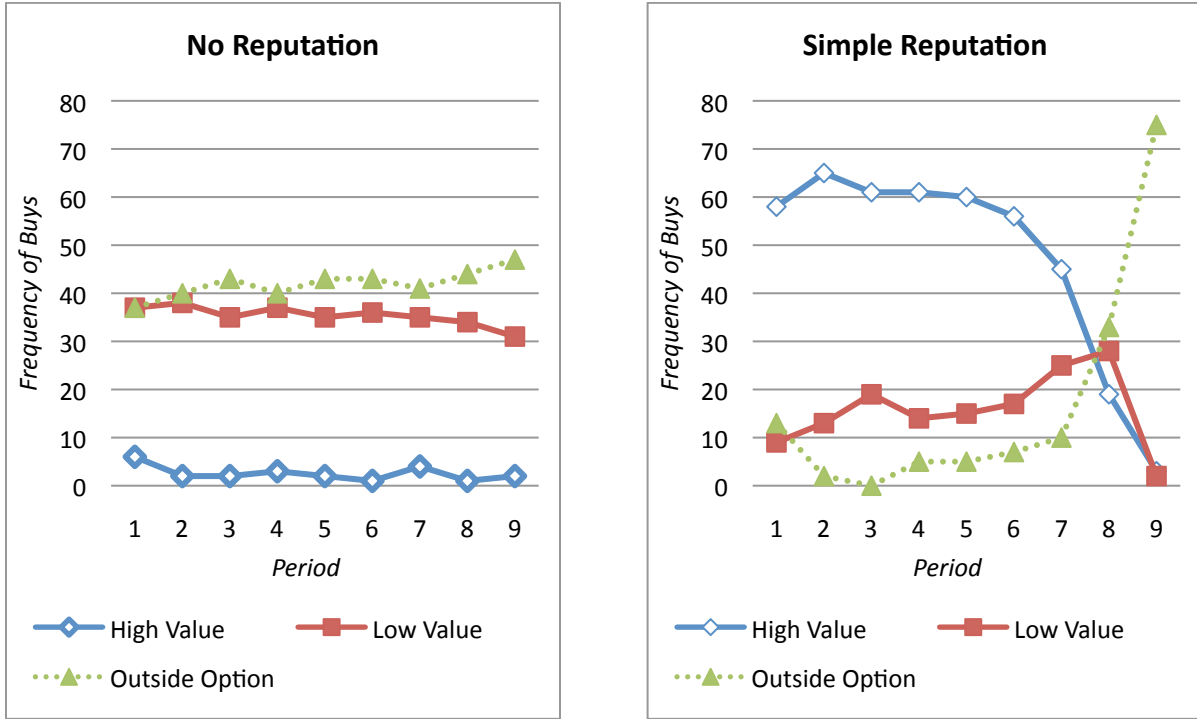


Figure 4: Percentage of Goods Bought of Offered, Periods 1-9, Sequences 2-6 Aggregated



Figure 5: Reneging Percentages

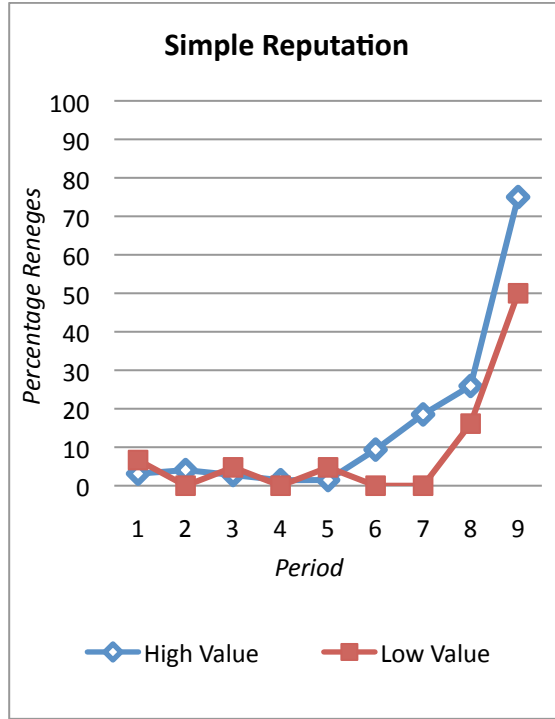
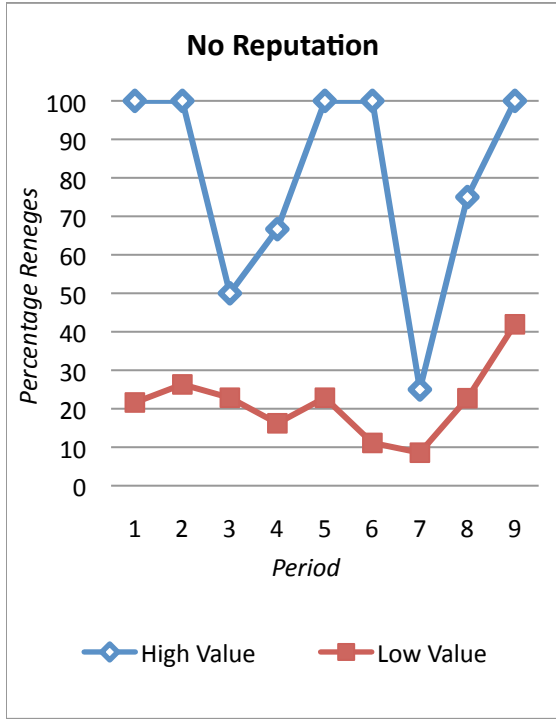


Figure 6: Efficiency, Periods 1-9, Sequences 2-6 Aggregated

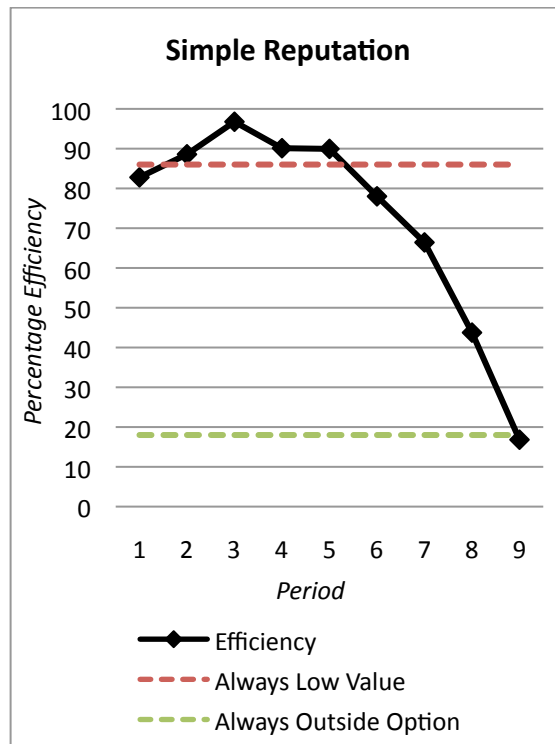
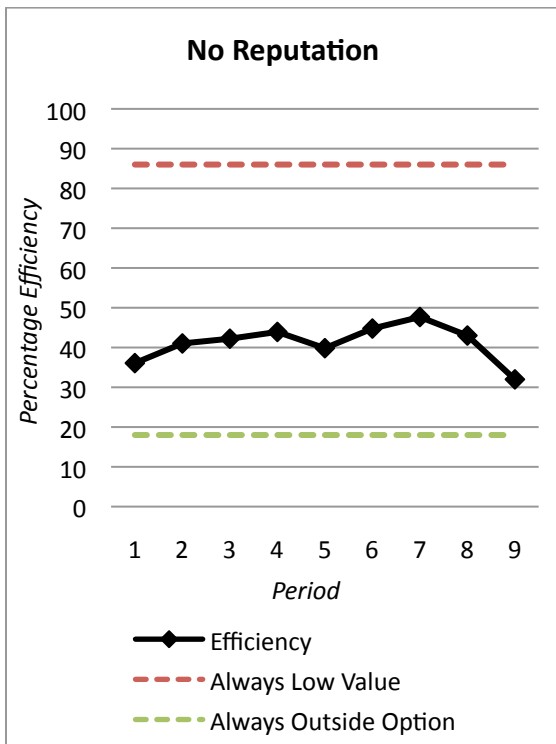


Figure 7: Reneging Percentages, Periods 1-9, Sequences 2-6 Aggregated

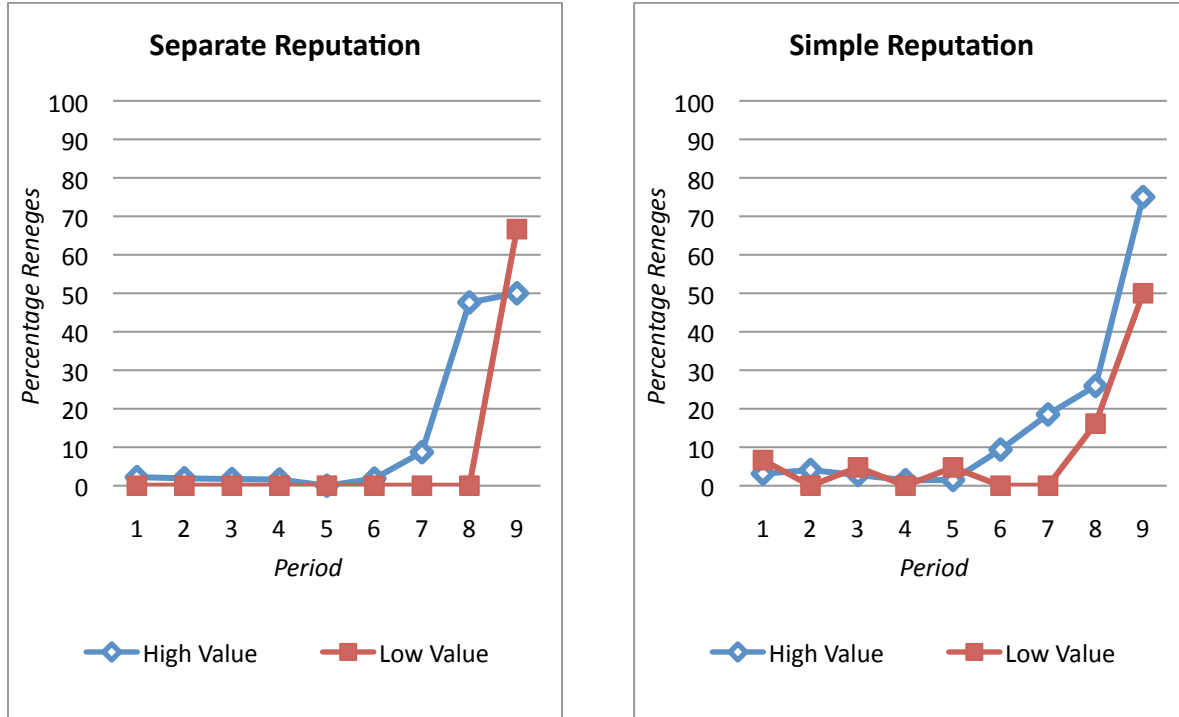


Table 1 – Summary of Treatments

	Type of Reputation	Number of Sessions	Number of Subjects
No Reputation Treatment	None	4	44
Simple Reputation Treatment	Simple	4	44
Separate Reputation Treatment	Separate	4	44
Total:		12 sessions	132 subjects

Table 2: Aggregate Number of Reneges

Treatment	Aggregate # High Value Reneges		Aggregate # Low Value Reneges	
	All periods (1-9)	Periods 1-7	All periods (1-9)	Periods 1-7
<i>No Reputation</i>	18	14	47	34
<i>Simple Reputation</i>	21	14	5	0
<i>Separate Reputation</i>	12	5	4	0

Table 3: Probit – Seller’s Offer Decision by Treatment

Dependent Variable: Seller’s Offer [1 if high value decision]	(1) Separate	(2) Simple	(3) No Reputation
Period 8	-0.31*	-0.60**	
[1 if $t=8$]	(0.14)	(0.13)	
Period 9	-0.47**	-0.81**	
[1 if $t=9$]	(0.13)	(0.13)	
1/Sequence	0.28	-0.39*	0.59**
[inverse of sequence order]	(0.16)	(0.15)	(0.15)
Has buyer lag	0.33**	0.31**	-0.31**
[1 if had buyer in $t-1$]	(0.10)	(0.09)	(0.10)
Reputation 100 Dummy - High	0.81**	0.18	
[1 if reputation is 100% in high (all) goods]	(0.11)	(0.12)	
Reputation 100 Dummy - Low	-1.09**		
[1 if reputation is 100% in low goods]	(0.12)		
# of safe options	0.01	-0.11*	-0.12**
[degree of risk aversion]	(0.02)	(0.06)	(0.04)
Observations	1260	1260	1260
# Subjects	28	28	28

Note: All results are from probit models with subject random effects. Standard errors in parentheses. Asterisks indicate ** $p < 0.01$, * $p < 0.05$

Table 4: Probit: Buyer Purchase Decisions by Treatment

Dependent Variable, Buyer's Purchase Decision [1 if (high, low) value]	(1) Sep. High Value	(2) Simple High Value	(3) No Rep High Value	(4) Sep. Low Value	(5) Simple Low Value	(6) No Rep. Low Value	(7) Sep. Outside Option	(8) Simple Outside Option	(9) No Rep Outside Option
Period 8	-1.29**	-1.56**		0.13	0.58**		1.69**	1.58**	
[1 if $t=8$]	(0.22)	(0.20)		(0.24)	(0.19)		(0.21)	(0.21)	
Period 9	-1.35**	-1.98**		-0.61*	-1.35**		2.55**	3.82**	
[1 if $t=9$]	(0.29)	(0.33)		(0.30)	(0.38)		(0.26)	(0.38)	
1/Sequence	0.48	-0.37	0.91*	0.03	0.58*	0.07	-0.43	0.02	-0.31
[inverse of seq. order]	(0.27)	(0.22)	(0.36)	(0.29)	(0.23)	(0.21)	(0.29)	(0.28)	(0.21)
Partner Coop. Lag	0.92**	0.55**	0.23	0.23	0.24	0.78**	-1.49**	-0.67**	-0.70**
[1 if received $t-1$]	(0.24)	(0.18)	(0.25)	(0.29)	(0.19)	(0.13)	(0.22)	(0.19)	(0.13)
Rep. 100 Dm. - High	1.85**	1.72**		-0.07	0.56**				
[1 if 100% high/all]	(0.18)	(0.16)		(0.21)	(0.18)				
Rep. 100 Dm. - Low	-0.86**			2.41**					
[1 if 100% low]	(0.23)			(0.27)					
Low Avail. Dm.	-1.17**	-0.90**	-0.14				0.15	-0.27	-1.01**
[1 if low good avail.]	(0.29)	(0.23)	(0.39)				(0.28)	(0.64)	(0.26)
High Avail. Dum.				-2.55**	-0.58	0.16	0.15	-1.23	-0.21
[1 if high good avail.]				(0.61)	(0.55)	(0.18)	(0.50)	(0.64)	(0.19)
# of safe options	-0.41	0.11	0.15	-0.06	-0.18	-0.13	0.11	0.02	-0.01
[risk aversion]	(0.04)	(0.08)	(0.09)	(0.10)	(0.11)	(0.13)	(0.07)	(0.09)	(0.16)
Observations	640	640	640	640	640	640	640	640	640
# Subjects	16	16	16	16	16	16	16	16	16

Note: All results are from probit models with subject random effects. Standard errors in parentheses. Asterisks indicate ** $p < 0.01$, * $p < 0.05$. Abbreviations can be translated with the following: "Coop" – Cooperative, "Rep." – Reputation, "Dm" – dummy, "Avail." – available."

Table 5: Reneging Behavior and Frequency of Attracting a Future Buyer³²

Find buyer for:	Separate Reputation			Simple Reputation		
	<i>high only</i>	<i>low only</i>	<i>neither</i>	<i>high only</i>	<i>low only</i>	<i>neither</i>
<i>Ren. High Only</i>	1	4	19	0	3	29
<i>Ren. Low Only</i>	0	0	0	1	2	3

³² These frequencies are obtained by counting the number of cases where a renege occurred in a period before period 9, and then looking to see if a future buy occurred. In all cases, the future buy occurred only one other time in the sequence.

Appendix I: Instructions for *Simple Reputation* treatment

INSTRUCTIONS

Welcome to the Experimental Economics Lab at Purdue University. Thank you for agreeing to participate in this experimental investigation on economic behavior. Please do not touch your computer until everyone is seated, and we have read the instructions together.

It is very important that you remain silent and do not look at other people's work. If you have any questions, please raise your hand and I will come to you. If you talk, laugh, or exclaim out loud, you will be asked to leave and will not be paid.

This is an experiment in the economics of decision-making. The instructions are simple and if you follow them carefully and make good decisions you will earn money that will be paid to you privately in cash at the end of the experimental session. Your earnings will be in experimental dollars and they will be converted into real dollars at the following rate: 50 Experimental Dollars = 1 Real Dollar. Notice that you earn more money by earning more experimental dollars. Everyone will also earn a \$5 show-up fee.

After we finish reading the instructions and before we start the experiment, we would like you to answer a set of questions relating to these instructions. You will be paid in cash (at the end of the experiment, in addition to your earnings from the actual experiment) at the rate of \$1.00 for each correct answer.

EXPERIMENT

The participants in today's experiment will be randomly split between two groups, designated as buyers or sellers. There will be 7 sellers and 4 buyers. Your buyer/seller designation remains unchanged throughout the experiment. The sellers and buyers will take turns making decisions. You will participate in this experiment for 6 sessions. There will be 9 periods in each session. After you have completed all sessions, 3 tokens will be randomly drawn out of a bingo cage containing tokens numbered from 1 to 9. The token numbers determines which period is going to be paid – 3 periods will be paid from each session. The tokens will be put back into the bingo cage, and this step will be repeated for sessions 2 through 6. Thus, in total within the 6 sessions, 18 periods will be paid. Since you don't know which period will be paid when you make your decisions, you should pay attention to the decision you make in every period of every session.

After this experiment is over, you will answer a quick questionnaire during which you will have the opportunity to earn more money.

Remember you have randomly picked a card when you arrived which told you what computer station to sit at. This card also told you whether you are a buyer or a seller for this experiment. As you continue to follow along with the directions, remember what your role will be for the remainder of the experiment.

YOU ARE A BUYER TODAY - Choose which item you will buy.		
<p style="color: purple; font-weight: bold; font-size: small;">SELLER - PURPLE ITEM</p> <p style="font-size: x-small;">Number of previous rounds seller sent the item is 2</p> <p style="font-size: x-small;">Percentage of previous rounds seller sent the item is 100</p>	<p style="color: orange; font-weight: bold; font-size: small;">SELLER - ORANGE ITEM</p> <p style="font-size: x-small;">Number of previous rounds seller sent the item is 1</p> <p style="font-size: x-small;">Percentage of previous rounds seller sent the item is 100</p>	<p style="color: purple; font-weight: bold; font-size: small;">SELLER - PURPLE ITEM</p> <p style="font-size: x-small;">Number of previous rounds seller sent the item is 1</p> <p style="font-size: x-small;">Percentage of previous rounds seller sent the item is 50</p>
<p style="color: orange; font-weight: bold; font-size: small;">SELLER - ORANGE ITEM</p> <p style="font-size: x-small;">Number of previous rounds seller sent the item is 1</p> <p style="font-size: x-small;">Percentage of previous rounds seller sent the item is 100</p>	<p style="color: orange; font-weight: bold; font-size: small;">SELLER - ORANGE ITEM</p> <p style="font-size: x-small;">Number of previous rounds seller sent the item is 1</p> <p style="font-size: x-small;">Percentage of previous rounds seller sent the item is 100</p>	<p style="color: pink; font-weight: bold; font-size: small;">COMPUTER SELLER 1 - PINK ITEM</p> <p style="font-size: x-small;">** Unlimited Items Available **</p> <p style="font-size: x-small;">Computer always sends the item.</p>
<p style="color: purple; font-weight: bold; font-size: small;">SELLER - PURPLE ITEM</p> <p style="font-size: x-small;">Number of previous rounds seller sent the item is 3</p> <p style="font-size: x-small;">Percentage of previous rounds seller sent the item is 100</p>	<p style="color: purple; font-weight: bold; font-size: small;">SELLER - PURPLE ITEM</p> <p style="font-size: x-small;">Number of previous rounds seller sent the item is 1</p> <p style="font-size: x-small;">Percentage of previous rounds seller sent the item is 50</p>	<p style="color: pink; font-weight: bold; font-size: small;">COMPUTER SELLER 2 - PINK ITEM</p> <p style="font-size: x-small;">** Unlimited Items Available **</p> <p style="font-size: x-small;">Computer always sends the item.</p>

SELLER

If you are a seller, you will move first. You will decide whether to offer a PURPLE item or an ORANGE item. Buyers in the room with you will decide whether they want to buy your PURPLE or ORANGE item.

BUYER

If you are a buyer, you will be able to buy a PURPLE item, an ORANGE item, or a PINK item. The PURPLE and ORANGE items will be sold by the sellers in the room with you. The PINK item will be sold by computer sellers. There will only be one item (either PURPLE or ORANGE) offered per seller, and once a buyer has bought an item, it will not be available to subsequent buyers. However, there are an unlimited number of PINK items offered by the computer sellers. Also, each buyer may only buy 1 item in each period.

At first, you will see a screen showing all the offers that are available.

Then, buyers will randomly enter the market one at a time in order to choose an item to buy. The random order of entry will be different in every period; that is, the order you enter in one period will not be related to the order you enter in any other period. Once a PURPLE or ORANGE item from a particular seller has been bought, it will no longer be displayed on the screen. However, PINK items will always be available to you.

Sellers in the room have numbers corresponding to the number of items sent and percentage of items sent during the current session. These numbers will be listed along with the seller's offer of an item. *Computer sellers do not have these numbers because they always send the item.*

Buying Screen

YOU ARE A BUYER TODAY - Choose which item you will buy.		
<p>SELLER - ORANGE ITEM</p> <p>Number of previous rounds seller sent the item is 1</p> <p>Percentage of previous rounds seller sent the item is 100</p> <p>Buy Item</p>	<p>SELLER - PURPLE ITEM</p> <p>Number of previous rounds seller sent the item is 1</p> <p>Percentage of previous rounds seller sent the item is 50</p> <p>Buy Item</p>	
<p>SELLER - ORANGE ITEM</p> <p>Number of previous rounds seller sent the item is 1</p> <p>Percentage of previous rounds seller sent the item is 100</p> <p>Buy Item</p>		<p>COMPUTER SELLER - PINK ITEM</p> <p>** Unlimited Items Available **</p> <p>Computer always sends the item.</p> <p>Buy Item</p>
<p>SELLER - PURPLE ITEM</p> <p>Number of previous rounds seller sent the item is 3</p> <p>Percentage of previous rounds seller sent the item is 100</p> <p>Buy Item</p>	<p>SELLER - PURPLE ITEM</p> <p>Number of previous rounds seller sent the item is 1</p> <p>Percentage of previous rounds seller sent the item is 50</p> <p>Buy Item</p>	<p>COMPUTER SELLER - PINK ITEM</p> <p>** Unlimited Items Available **</p> <p>Computer always sends the item.</p> <p>Buy Item</p>

Note that the location of each seller's offer will be random in every period. That is, a seller whose offer is listed in the top right corner of the screen in one period may be listed in any other or the same area in another period. Thus, as a buyer when making your buy decision you should pay attention to the seller's number of items sent and percentage of items sent and not the location of the offer.

SELLER

If you are a seller who offered a PURPLE or ORANGE item but no buyer bought this item, you will be shown this outcome and will wait to move on to the next period.

If you are a seller who offered a PURPLE or ORANGE item and this item was bought you will have one more choice to make.

The buyer sends the experimental dollars to buy the item but you will decide whether you will send the item to the buyer or not. **All future buyers will be able to see a number indicating how many times you have sent an item and a number indicating the percentage of times you have sent an item.** They will see these on the screen together with your sell offer when they make their decision regarding what to buy.

If you choose to send an item to the buyer, the number representing how many times you have sent an item will increase by 1. Also, the percentage of times you have sent an item will increase. If you choose not to send the item

to the buyer, the number representing how many times you have sent an item will remain the same. Also, the percentage of times you have sent an item will decrease.

These numbers are cumulative for the entire session of 9 periods. When the session terminates and another session starts, the number of times you have sent the item and the percentage of times you have sent the item revert back to the start.

The number of items sent will start at 0. The percentage of items sent will be blank until the seller has had a chance to choose to send or not send items.

To sum up, the numbers will be calculated as follows:

Number representing how many times you have sent an item:

- ◆ **Send an item in a given period:** +1
- ◆ **Do not send an item in a given period:** (Remains same)
- ◆ **Buyer does not buy the item:** (Remains same)

Percentage of times you have sent the item:

- ◆ **Send an item in a given period:** Percentage increases (unless it is already 100%)
- ◆ **Do not send an item in a given period:** Percentage decreases (unless it is already 0%)
- ◆ **Buyer does not buy the item:** (Remains same)

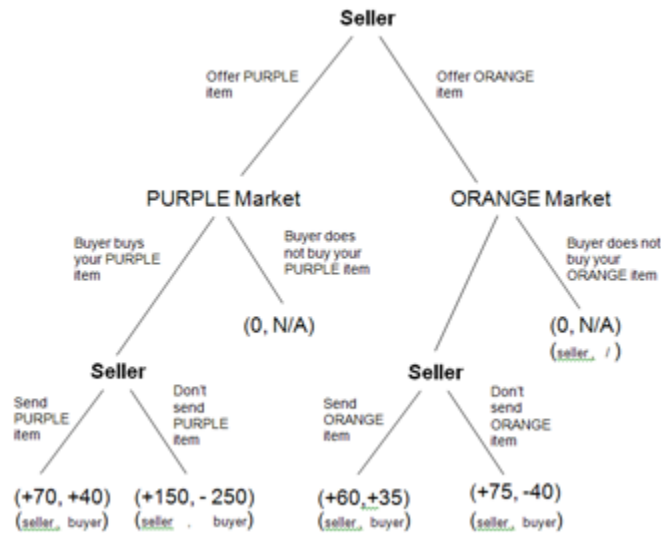
Examples:

- ◆ If the number of times you have sent the item is 0 in the first period, then you send an item, then in the next period the number is $0+1=1$. The percentage of times you have sent the item increases to $1/1=100\%$
- ◆ If the number of times you have sent the item is 1 at the start of the second period, and you choose not to send an item in the second period, then in the next period the number is still 1. Your percentage sent decreases to $(1/2)*100=50\%$.
- ◆ If the number of times you have sent the item is 1 in the current period, and the buyer chooses not to buy the item, then in the next period the number remains 1. Also, the percentage sent remains the same.

The experimental dollars earned in each period for both the buyer and seller will be determined by the actions that both parties take.

SELLER

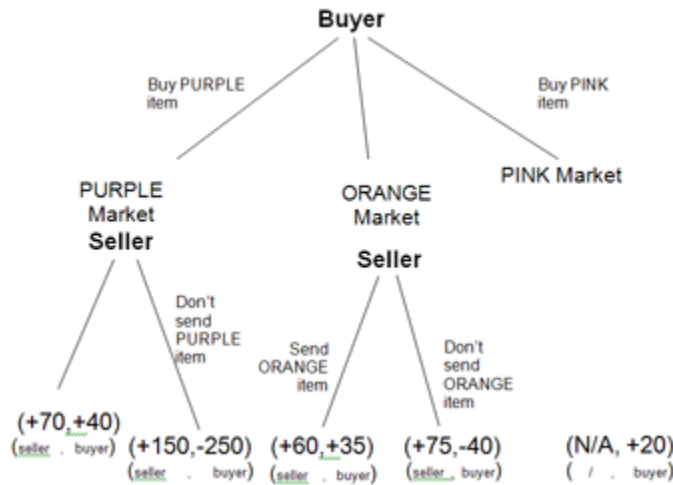
Here is a diagram describing the sequence of actions that the SELLER can take, and the corresponding earnings for the seller and the buyer:



The first number in parentheses is the earnings for that set of actions for the SELLER; the second number in parentheses is the earnings for that set of actions for the BUYER. N/A is also indicated when no buyer buys the item.

BUYER

Here is a diagram describing the sequence of actions that the BUYER can take, and the corresponding earnings for the seller and the buyer:



The first number in parentheses is the earnings for that set of actions for the SELLER; the second number in parentheses is the earnings for that set of actions for the BUYER. N/A is indicated for the computer sellers because they don't actually earn any money.

Examples for the SELLER

- ◆ The seller chooses to sell the PURPLE item. Start at the top of the tree and follow the line next to "Offer Purple item" down to the PURPLE market.
- ◆ Now the PURPLE item is awaiting a buyer. A buyer chooses to buy the item. Follow the line next to "buyer buys your PURPLE item." This is the left branch.
- ◆ Now it is the seller's turn again. The seller chooses to send the item. Follow the branch next to "send PURPLE item" under the word "Seller" to find the outcome. This is the left branch.
- ◆ **Earnings are: 70 for the seller, and 40 for the buyer**

Examples for the BUYER

- ◆ The buyer finds an offer for the PINK item from a computer seller and chooses to buy it.
- ◆ Now it is the computer seller's turn. The computer seller always sends the item.
- ◆ **Earnings are: 20 for the buyer**

Important: The earnings will given by this same tree for the entire experiment. That is, one set of actions in one period would result in the same earnings as the same set of actions in any other period.

OUTCOME of PERIOD

After a period is over, you will see the outcome. As a seller, you will see the actions that took place, your earnings in that period, and the earnings in that period of your buyer (if a buyer bought your item.) For your reference, you will also see the numbers representing how many times items have been sent and the percentage of times items have been sent for every other seller in the room. Note that the row corresponding to your number of items sent and percentage of items sent will be marked with a "You" in the first column of the table ("Subject ID"), see the screenshot below:

A buyer purchased your PURPLE item

You sent the buyer a PURPLE item

Your earnings in this period are: 70
The buyer's earnings in this period are: 40

Your updated numbers:
Number of previous rounds you sent the item is: 3
Percentage of previous rounds you sent the item is: 100%

Information about other sellers in the room

Subject ID	Number of previous rounds seller sent the item is	Percentage of previous rounds seller sent the item is
You	3	100
	1	100
	3	100
	1	100
	2	100
	1	50
	1	50

[Continue](#)

Note that the row that your information is listed on will change in every period, because the order is randomly assigned in every period. As a buyer, you will see the actions that took place, your earnings that period, and the earnings in that period of your seller. **Record these outcomes in your record sheet, before you press “continue.”** If you have any questions, please raise your hand and I will come to answer your question. Now we will take the quiz before we continue to the experiment.

Appendix II – Quizzes, buyer and seller

SELLER QUIZ

You are a seller. Please answer the following 4 questions. *Suppose the following happens:*

- ◆ The seller chooses to offer the PURPLE item.
 - ◆ A buyer chooses to buy the item.
 - ◆ The seller chooses not to send the item.
1. **What are the earnings for you, the SELLER?** _____
 2. **What are the earnings for the BUYER?** _____

Instead, Suppose the following happens:

- ◆ The seller chooses to offer the PURPLE item
 - ◆ No buyer buys this item
3. **What are the earnings for you, the SELLER?** _____

Instead, Suppose the following happens:

- ◆ The seller chooses to sell the ORANGE item.
 - ◆ Now it is the buyer's turn. The buyer chooses to buy the item.
 - ◆ Now it is the seller's turn again. The seller chooses not to send the item.
4. **What are the earnings for you, the SELLER?** _____

BUYER QUIZ

You are a buyer. Please answer the following 4 questions. *Suppose the following happens:*

- ◆ The buyer finds an available PURPLE item from a certain seller and chooses to buy it.
 - ◆ The seller chooses to send the item.
1. **What are the earnings for you, the BUYER?** _____
 2. **What are the earnings for the SELLER?** _____

Instead, Suppose the following happens:

- ◆ The buyer finds an available PURPLE item from a certain seller and chooses to buy it.
 - ◆ Now it is the seller's turn. The seller chooses not to send the item.
3. **What are the earnings for you, the BUYER?** _____

4. What are the earnings for the SELLER? _____

Appendix III – Risk Elicitation

For each line in the table in the next page, please state whether you prefer option A or option B.

Notice that there are a total of 15 lines in the table but just one line will be randomly selected for payment. You ignore which line will be paid when you make your choices. Hence you should pay attention to the choice you make in every line. After you have completed all your choices a token will be randomly drawn out of a bingo cage containing tokens numbered from 1 to 15. The token number determines which line is going to be paid.

Your earnings for the selected line depend on which option you chose:

If you chose option A in that line, you will receive 10 experimental francs.

If you chose option B in that line, you will receive either 30 francs or 0 francs. To determine your earnings in the case you chose option B there will be second random draw. A token will be randomly drawn out of the bingo cage now containing twenty tokens numbered from 1 to 20. The token number is then compared with the numbers in the line selected (see the table). If the token number shows up in the left column you earn 30 francs. If the token number shows up in the right column you earn 0 francs.

Now it is time for clarifications. Are there any questions?

Francs will be converted to dollars at the rate of 10 francs = \$1.

Decision no.	Option A	Option B	Please choose A or B
1	\$1	\$3 never	\$0 if 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15, 16,17,18,19,20
2	\$1	\$3 if 1 comes out of the bingo cage	\$0 if 2,3,4,5,6,7,8,9,10,11,12,13,14,15, 16,17,18,19,20
3	\$1	\$3 if 1 or 2	\$0 if 3,4,5,6,7,8,9,10,11,12,13,14,15, 16,17,18,19,20
4	\$1	\$3 if 1,2 or 3	\$0 if 4,5,6,7,8,9,10,11,12,13,14,15, 16,17,18,19,20
5	\$1	\$3 if 1,2,3,4	\$0 if 5,6,7,8,9,10,11,12,13,14,15, 16,17,18,19,20
6	\$1	\$3 if 1,2,3,4,5	\$0 if 6,7,8,9,10,11,12,13,14,15, 16,17,18,19,20
7	\$1	\$3 if 1,2,3,4,5,6	\$0 if 7,8,9,10,11,12,13,14,15,16,17,18,19,20
8	\$1	\$3 if 1,2,3,4,5,6,7	\$0 if 8,9,10,11,12,13,14,15,16,17,18,19,20
9	\$1	\$3 if 1,2,3,4,5,6,7,8	\$0 if 9,10,11,12,13,14,15,16,17,18,19,20
10	\$1	\$3 if 1,2,3,4,5,6,7,8,9	\$0 if 10,11,12,13,14,15,16,17,18,19,20
11	\$1	\$3 if 1,2, 3,4,5,6,7,8,9,10	\$0 if 11,12,13,14,15,16,17,18,19,20
12	\$1	\$3 if 1,2, 3,4,5,6,7,8,9,10,11	\$0 if 12,13,14,15,16,17,18,19,20
13	\$1	\$3 if 1,2, 3,4,5,6,7,8,9,10,11,12	\$0 if 13,14,15,16,17,18,19,20
14	\$1	\$3 if 1,2, 3,4,5,6,7,8,9,10,11,12,13	\$0 if 14,15,16,17,18,19,20
15	\$1	\$3 if 1,2, 3,4,5,6,7,8,9,10,11,12,13,14	\$0 if 15,16,17,18,19,20