No. 3

UNIVERSITY OF COLOGNE

WORKING PAPER SERIES IN ECONOMICS

HOW EFFECTIVE ARE ELECTRONIC REPUTATION MECHANISMS? AN EXPERIMENTAL INVESTIGATION

GARY E. BOLTON, ELENA KATOK, AXEL OCKENFELS

Department of Economics
University of Cologne
Albertus-Magnus-Platz
D-50923 Köln
Germany

http://www.wiso.uni-koeln.de

How Effective are Electronic Reputation Mechanisms?

An Experimental Investigation

Gary E Bolton*

Smeal College of Business, Penn State University and Harvard Business School, USA

Elena Katok

Smeal College of Business, Penn State University and Harvard Business School, USA

Axel Ockenfels

Faculty of Economics and Social Sciences, University of Cologne, Germany

September 2003

Electronic reputation or "feedback" mechanisms aim to mitigate the moral hazard problems associated with exchange among strangers by providing the type of information available in more traditional close-knit groups, where members are frequently involved in one another's dealings. In this paper, we compare trading in a market with online feedback (as implemented by many Internet markets) to a market without feedback, as well as to a market in which the same people interact with one another repeatedly (partners market). We find that, while the feedback mechanism induces quite a substantial improvement in transaction efficiency, it also exhibits a kind of public goods problem in that, unlike in the partners market, the benefits of trust and trustworthy behavior go to the whole community and are not completely internalized. We discuss the implications of this perspective for improving feedback systems.

---

* Correspondence to Gary Bolton, 310K Business Administration Building, Smeal College of Business, Penn State University, State College, PA 16870; gbolton@psu.edu. Ockenfels gratefully acknowledges the support of the Deutsche Forschungsgemeinschaft through the Emmy Noether-program. Katok gratefully acknowledges the support of the National Science Foundation, award number SES-0214337.

1. Introduction: The problem of trust in online markets

Many online markets rely on electronic reputation or "feedback" systems to promote trust in transactions. The reliance on reputation *per se* does not distinguish online markets from traditional markets, legal enforcement being in all cases expensive. On balance, however, online markets have more problems with fraud. A recent report by the research group GartnerG2 (2002) concludes that "Internet transaction fraud is 12 times higher than in-store fraud." A U.S. Department of Justice (2002) survey also cites high levels of online fraud, pointing especially at frauds common on auction sites (many with online feedback systems) that "induce their victims to send money for the promised items, but then deliver nothing or only an item far less valuable than what was promised (e.g., counterfeit or altered goods)."[1]

The power of reputation to promote trust in business transactions is closely associated with networked communities, places where there is a good deal of interpersonal communication as well as exchange. Online and traditional markets are networked in different ways. Whether the networking pattern is critical to the amount of trust in the market is a matter of contention. The laboratory experiments we present in this paper investigate how differences in the *flow of information* – that is, the source of the information and how it disseminates in the market – might differentially influence trust. Our results imply that, even if information about reputation is shared and reliable, online feedback systems provide fewer incentives to trust or to be trustworthy than do traditional markets, where long-term relationships play a larger role. The results also suggest ways in which online incentives might be improved.

In traditional business communities, the patterns of information flow and contact that promote trust often interact in subtle ways. Vietnam's free market reform in the mid-1980's illustrates how effective informal reputational controls can be. At the time, there was little in the way of legal protection against exchange malfeasance, but markets nevertheless flourished. According to McMillan (2002), "People in the same line of business would meet each other

[1] Other evidence comes from the U.S. White Collar Crime Center and the Federal Bureau of Investigation (2001). The Center reports that 63 percent of the 49,711 formal complaints received in 2001 involved either non-delivery of merchandise, non-payment, or auctions, and that these numbers are rising rapidly.

every day in teahouses and bars …to discuss the reliability of particular customers. … About half of a sample of entrepreneurs said that they had had no prior connections with the businesses that were to become long-standing trading partners." Thus, in traditionally networked communities, interaction between members promotes trust in two ways. For one, the pattern of interaction promotes long-term relationships; a business partner whose trust has been rewarded is, all things equal, more likely to return to do future business. Second, information about individual reliability is transmitted by word-of-mouth to third parties, some of whom are prospective future trading partners.[2]

One of the great advantages of online markets is the opportunity to trade with a larger, fluctuating set of partners. This means lesser reliance on long-term relationships. In fact, in data collected from eBay over a five month period, Resnick and Zeckhauser (2002) found that 89 percent of all encounters were one-shot. To promote the exchange of information on the reliability of individual traders, many platforms, including Amazon, Cnet, eBay, Half, and Yahoo, have instituted online reputation mechanisms, known as "feedback" systems, to provide the kind of word-of-mouth available in traditionally networked markets. Online feedback systems arguably improve the flow of word-of-mouth reputational information in the sense that accessing information online does not require personal contacts in the trading community, and feedback information from large numbers of traders can easily be collected, processed and disseminated. Amazon's used books market platform for independent dealers (brick-and-mortar bookstores as well as private individuals) provides a simple but illustrative example of how these systems work. Sellers post the price and a description of the book's condition on Amazon's site. Buyers pay through Amazon, who takes a percentage, but sellers ship directly to buyers. The moral hazard problems inherent in the seller's side of the deal – stipulating the book's condition

[2] Use of reputation to enforce trustworthiness is perhaps as old as human social interaction. Alexander (1987) argues that reputation mechanisms are at the base of human moral systems. Greif (1989) investigates the reputation systems to facilitate trust among strangers, used by certain groups of Mediterranean traders during the Middle Ages. Milgrom et al. (1990) provide a historical as well as theoretical account of the role of trade fairs in disseminating reputation information during the Middle Ages. See Buchan et al. (2002) for a cultural comparison of trust and trustworthiness.

and the shipping – is addressed through the feedback system in which buyers are invited to post comments on the transaction that future buyers can view when deciding whether to make a purchase.[3] Recent field studies of online auction platforms find that feedback systems have at least some of the desired economic effect in the sense that reputable sellers are more likely to sell their items (Resnick and Zeckhauser, 2002), and can expect price premiums (e.g., Lucking-Reiley et al., 1999).[4]

One way of stating the difference between online and traditional reputation networks is to note that they emphasize different types of reciprocity. Traditional markets rely more on *direct* reciprocity: 'I trust you because you were trustworthy with me before.' Online markets rely more on *indirect* reciprocity: 'I trust you because you were trustworthy with others before.' In both cases, information about reputation enforces trust by inducing a reciprocal response; past trustworthiness is a prerequisite to future business. The two differ, however, in terms of the flow of information. In particular, in direct reciprocal dealings, traders make decisions based on reputational information culled from their own past transactions, and their present dealings produce information that they themselves will use in the future. But in indirect reciprocal dealings, reputational information is obtained from others, and others will use information from the present dealings in the future.[5]

Putting our investigation in these terms, we look at how well markets based on indirect reciprocity build and sustain trust in comparison to markets based on direct reciprocity. Game theoretic models imply that indirect reciprocity can be just as effective as direct reciprocity, so long as traders in indirect reciprocal relationships have access to sufficient information about

[3] Here is how eBay (2002) explains their feedback system: "A user's feedback is a key factor people use to determine whether or not they want to trade with that user. What feedback you give or receive is an important part of your trading reputation at ebay. […] If you're a buyer, checking a seller's Feedback Profile before you make a bid is one of the smartest and safest moves you can make. This Feedback Profile answers many questions about how a seller does business."

[4] Analogous results come from Ba and Pavlou (2002), Houser and Wooders (2001), Melnik and Alm (2002), and Ockenfels (2003); see Resnick et al. (2002) and Dellarocas (forthcoming) for recent surveys. Brynjolfsson and Smith (2000) compared pricing behavior at 41 Internet and conventional retail outlets. They identify internet sellers' trustworthiness as one important factor that affects market outcomes.

[5] See Brynjolfsson and Smith (2000), Dellarocas (2001), and Resnick and Zeckhauser (2002), for detailed comparisons of electronic and conventional dissemination of information about reputation.

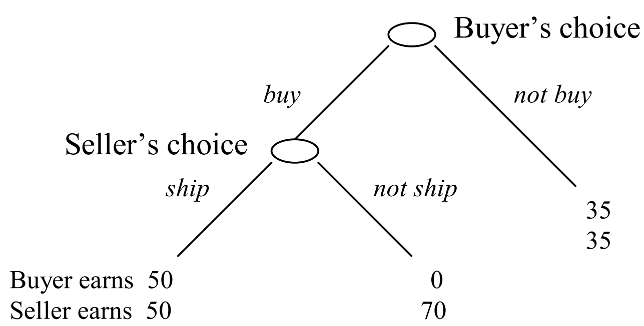
reputations; that is, game theoretic models imply what we will call the information hypothesis: It is the information *per se*, independent of its flow, that matters. Standing in counterpoint to the information hypothesis is an argument that says that the flow of information is, in fact, critical. Granovetter (1985), for example, in discussing market trust, argues that people put more stock in information acquired "from one's own dealings." This suggests that direct reciprocity is a more effective way of developing trust even when sufficient information for indirect reciprocity is present. But even if so, understanding *why* the information flows matter might help to improve online trust.

Section 2 develops the information hypothesis, and discusses the counter argument. Section 3 then details the design of the experiment. Section 4 lays out results. Section 5 discusses the implications for improving online markets.

2. Three markets and the information hypothesis

To put matters on a tangible footing, we describe the fundamental trust problem in the context of a simple market platform. Using this platform as a base, we characterize three specific markets that differ by the flow of information. We then describe the information hypothesis in the context of these markets.

Figure 1. The buyer-seller encounter



2.1 The basic trading platform for three markets

In each market, transactions take place over a series of rounds. (In the experiments, of course, there will be a finite number of rounds, but the number of rounds − finite or infinite − is not critical to our basic argument.) At the beginning of each round, a potential buyer is matched with a potential seller. The buyer then chooses whether to purchase an item at a fixed price. If the purchase order is sent, the seller decides whether to ship or simply keep the buyer's money. The moral hazard is that, on receiving the money from the buyer, the seller has no immediate pecuniary incentive to ship the item. Thus, a transaction that is in both parties interests may be impeded either because the seller proves untrustworthy, or because the buyer, anticipating this risk, chooses not to trust.

Figure 1 illustrates the exact moves in the buyer-seller encounter. Both the seller and the buyer are endowed with 35, which is the payoff when no trade takes place. The seller offers an item for sale at a price of 35 which has a value of 50 to the buyer. The seller's cost of providing the buyer with the item − costs associated with executing the trade, shipping, handling, as well as production costs − is 20.[6] So each successfully completed trade increases efficiency by creating a consumer surplus of 15 and a net profit of 15 for the seller. If the buyer chooses to *buy* the item, he sends his endowment of 35 to the seller, who then has to decide whether to ship the item. If the seller does not ship, he receives the price plus his endowment of 35 for a total of 70. If he ships, he receives the price minus the costs plus his endowment for a total of 50. If the buyer chooses not to buy the item, no trade occurs. In this sense, the buyers can choose with whom to trade and with whom not to trade.

The three markets are characterized as follows: In the *strangers market*, individual buyers and sellers meet no more than once and the buyer has no information about the seller's transaction history. Here the moral hazard has full force, since the actions of the seller are not conveyed to future perspective customers. In the *feedback market*, an online feedback system

[6] These are production costs where either the seller produces the item after he knows the demand, or the product is produced before the buyer's decision is known but costs are not sunk (e.g., when the item can be resold at a price equal to production costs).

tracks seller histories of shipping decisions and provides this information to prospective buyers. This affords the type of indirect reciprocity associated with online markets. In the *partners market*, the same buyer-seller pairs interact repeatedly, in every round. This affords the type of direct reciprocity associated with more traditional markets.

2.2. The information hypothesis

The critical question we investigate is whether the flow of information, in addition to the information content, affects trust and trustworthy behavior. If not, then feedback and partners markets should, in theory, produce just as much trustworthy exchange. If so, then partners markets may have an advantage.

A robust finding of game theoretic investigations of reputation building is that a market networked for indirect reciprocity can support just as much trustworthy exchange as a market networked for direct reciprocity. That is, standard game theory implies the *information hypothesis* (our null hypothesis), that information about reputation determines trusting and trustworthy behavior independent of the pattern of the flow of information.

To illustrate the reasoning behind the hypothesis, suppose that a buyer in a partners relationship suffers from short-term memory loss and always forgets the outcome of the last buying encounter. The incentive for the seller to be honest disappears with the information in the buyer's head. But now suppose that some feedback mechanism reminds the forgetful buyer of what the seller did last time. Now the incentive is restored – even if, be the buyer aware or not, the information is about a different seller than the one the forgetful buyer dealt with last. For the seller to have an incentive to be trustworthy, he need only expect that a future buyer will punish or reward his behavior; whether these punishments or rewards come from the same or from different buyers is irrelevant. The buyer, to induce this trustworthiness, need only be equipped with sufficient information about the sellers' histories; whether this information comes from one's own experience or from different sources is irrelevant. This is the basic message that derives from the game theoretic models: it is the information, not its source or its dissemination, that matters (ex., Kreps et al., 1982, Ellison, 1994).

That said, precisely how much exchange game theory predicts will occur in these markets depends on the specific modeling approach taken. In this regard, a critical issue is whether the model assumes that traders have complete information about one another's payoffs (different from information about reputation). The payoffs can include psychological as well as pecuniary awards. The simplest analysis assumes common knowledge of payoffs, and that the pecuniary rewards shown in Figure 1 are the sole rewards. Given this, when the market has a finite number of rounds there is a unique (Nash) equilibrium, found by backwards induction: In the final round of play, the seller's optimal action is *not ship*, and so the buyer's optimal action in the last round is to exercise the outside option. This is true regardless of whatever seller feedback the buyer has, or of whether buyer and seller are strangers or partners. And since there is no trading in the last round, there is no incentive to ship in the next to last round, and so the buyer's optimal action in the next to last round is again the outside option, and so on… back to the 1st round. The equilibrium implies no trade in any round of strangers, feedback or partners markets.

It turns out that this analysis is highly sensitive to the common knowledge assumption. Kreps et al. (1982) demonstrate, in the context of a finitely repeated prisoner's dilemma, that if each player assesses even a small positive probability that his partner receives some psychological reward beyond the pecuniary reward for cooperating, then there are (sequential) equilibria in which rational players cooperate, at least until the last few rounds of play. To see the intuition in the context of our markets, suppose buyers believe there is a small proportion of sellers who will ship out of, say, a sense of social responsibility. Buyers will buy only if the seller has shipped for orders received in the past (failure to ship reveals the seller as not honest). This gives all sellers a pecuniary incentive to ship since doing so is a prerequisite for future business. Of course, for this to work, seller past actions must be observable to the buyer – a condition that holds equally in the feedback and partners market, but not strangers. Thus, this model implies there should be the same number of transactions in feedback and partners markets, but few, if any, transactions in strangers.[7]

The analyses so far suppose that traders exit the market after a fixed and known number of rounds, as they do in our experiments. We note, however, that the infinite horizon case, where there is no maximal amount of time traders can be in the market, retains the critical implication of the information hypothesis; there is, from a game theoretic perspective, no reason to expect different exchange patterns in feedback and partners markets, and exchange will not take place in the strangers market (Ellison, 1994, Ockenfels, 2003, and the references cited therein). Thus, the information hypothesis, that equal information leads to an equal amount of trade, is robust to all of these models, even though the predicted amount of trade is model dependent.

Still, there are reasons to suspect that information flows matter. For one, in a partners market, the reputational information available about one's partner, and one's own history of transactions, precisely match. In a feedback market, however, the two are non-intersecting sets. The game theory models, and so the information hypothesis, imply that only the history of one's partner should matter to behavior, and so this difference in own and partner histories is superfluous to these models. But it seems plausible that one's history with one set of partners might influence trust and trustworthiness with other partners, which is in fact what our analysis will suggest. It will also suggest that there is a kind of information dilemma at work in the feedback markets, associated with the fact that, in these markets, reputational information is a public good benefiting all and not only those who produced the information.

[7] Kreps et al. (1982) prove an even stronger result: Even if *all* players' payoffs are correctly described as in our Figure 1, if this is not common knowledge, trust and trustworthiness may emerge in equilibrium. However, as with Kreps et al.'s model, the formal incomplete information model for our markets would be technically demanding, and we will not attempt to work such a model out here, but see Bolton and Ockenfels (2003). Game theoretic models in which people are assumed to care about social as well as pecuniary payoffs can explain data patterns across a wide variety of experimental games (ex., Fehr and Schmidt, 1999, and Bolton and Ockenfels, 2000; the latter make explicit connection to the Kreps et al. model, Section VI.D). Güth and Ockenfels (forthcoming) review the theory literature on the evolution of preferences in trust games not unlike the one we study. They show that depending on the institutional environment trustworthy behavior may survive evolutionary competition. Both of these lines of research suggest that trustworthiness and trust can have non-strategic as well as strategic causes. Our experiments measure the extent to which trust and trustworthiness is intrinsically or strategically motivated.

In addition, Granovetter (1985) puts forward a number of reasons why one might put special weight on one's own dealings rather than on reputation information provided by others, having to do with social ties or the costs of gathering or quality of information. While these differences are not relevant to our laboratory markets, they may well be relevant to real world online markets − and they tend to reinforce what we will find (as we explain in section 5).

3. Experimental design

Our experiment compares the strangers market to the feedback market as a method of measuring the improvement in trust and trustworthiness feedback brings about, keeping the trader matching scheme fixed. The comparison also permits us to gauge the extent to which information about reputation is used strategically. We then compare the performance of the feedback market with that of the partners market to judge the accuracy of the information hypothesis. A deeper analysis of this data permits us to investigate how the flow of information affects trust and trustworthiness.

Our experiment focuses on the role of differing patterns of information flow, and so our design carefully excludes other potentially confounding factors. For example, we control for the noise in feedback production, always truthfully provided in our experiment (see Ba, 2001, Dellarocas, 2001, and Ba et al., 2003, on the incentives and consequences of online identity and feedback manipulation). We control the distribution of individual valuations and knowledge of these valuations in all markets, and we focus on the effect of reputation on the probability of trade by keeping the price fixed across markets. Many of these factors, and their influence on trust, are interesting in their own right and deserving of broader investigation. Our experiment provides sufficient framework to address them in future investigations, in a way that might be difficult to do in the field.[8]

[8] See Roth (2002) and Ariely et al. (2002) on the relationship between experimental and field studies.

3.1 Implementing the markets

The payoffs in Figure 1 were used in all three markets, and (both buyer and seller) payoffs were known to all. Each payoff point in Figure 1 was worth $0.01 to the participant. The gains from trade are about 40% (a gain of 15 on the outside option of 35). This will prove to be sufficient to induce a high level of transaction in at least some of the markets (Bolton et al., 2002, show that the size of the gains from trade can influence trust in a reputation market).

The markets ran for 30 rounds. This length was made public knowledge to participants at the beginning of all three markets. Doing so allows us to observe endgame behavior as a measure of *strategic* reputation building since the strategic value of a good reputation diminishes in the last rounds. In addition, not stating the market end risks loss of experimental control.[9]

In the strangers market, buyers and sellers are randomly paired in each round under the commonly known restriction that nobody is matched with the same player in the same role more than once. Buyers and sellers are anonymous to one another, and no information about one another's history of action is made available. The random matching is public knowledge.

Matching is done the same way in the feedback market (we used the same pairing rotation used in the strangers market; in particular, no buyer is matched more than once with a specific seller), but here buyers are given the feedback on the seller's shipping decisions prior to choosing whether to buy. The feedback includes the total number of times the seller shipped in the past, as well as a round-by-round history of their shipping decisions. Figure A1 in Appendix A displays a typical computer screen. We find that the kind of tree representation we used leads to better understanding of the game. Also, the computer highlights the path of play on the tree as the game progresses, so subjects can know where they are in the game and how they got there.

In the partners market, matching is fixed; that is, the same two subjects are always paired together, and this is public knowledge. The same kind of feedback information available to

[9] Participants would still guess at the market end, making for unobservable "endgame" behavior, and since unobservable, potentially confounded with other factors. A design in which the market randomly stops or continues after each round introduces risk-over-stopping as a factor, again not directly observable. Note that, since participants in all markets were told of the 30 rounds, this cannot explain the differences we observe across markets.

feedback market buyers is available to partners market buyers. In this way, the information structure is parallel across the two markets, albeit the information is redundant in the sense that it is telling the buyers things the buyer has himself experienced.

3.2 Experimental protocol

The written instructions given to participants, reproduced in Appendix A, describe the protocol for the experiment in detail. Here is a synopsis: Each of the three markets consisted of three sessions. There were 16 subjects per session (48 per market) for a total of 144 participants in the experiment. All sessions were conducted in March and April of 2002 at the Laboratory for Economic Management and Auctions (LEMA) in the Smeal College of Business, Penn State University. Subjects were Penn State University students, mostly undergraduates, from various fields of study who volunteered through an on-line recruitment system. Cash was the only incentive to participate. Upon arrival at the laboratory participants were seated at the computers, separated by partitions. They were asked to read the instructions. To create public knowledge, the monitor read instructions to subjects out loud. Subjects then played several practice games in a sequence of roles that was chosen at random, with the computer as partner making its moves at random. To encourage subjects to explore the features of the game interface, practice game payoffs were displayed as the Marx brothers: Chico, Groucho, Harpo and Zeppo. Once familiar with the game interface, subjects played the 30 actual rounds. Upon completion of the session, each subject was privately paid his or her earnings in cash plus a $5 show-up fee.

4. Results

We first describe the basic treatment effects we observe. While access to reputation information induces a substantial improvement in transaction efficiency, partners markets perform significantly better than feedback markets. We then investigate the extent to which traders build and respond to reputation in a strategic manner. This establishes the foundation necessary to investigate why the feedback market's flow of information creates inferior incentives to trust and to be trustworthy.

4.1 Treatment effects

The major treatment effects have to do with trading patterns, to which there are three dimensions: *Efficiency* or the percentage of potential transaction completed (Figure 2), *trust* or the percentage of buy orders given (Figure 3), and *trustworthiness* or the percentage of shipped items conditioned on buy orders (Figure 4). In these figures, the treatment data has been aggregated across sessions.

Figure 2. Efficiency measured as how often the gain from trade is realized, by round

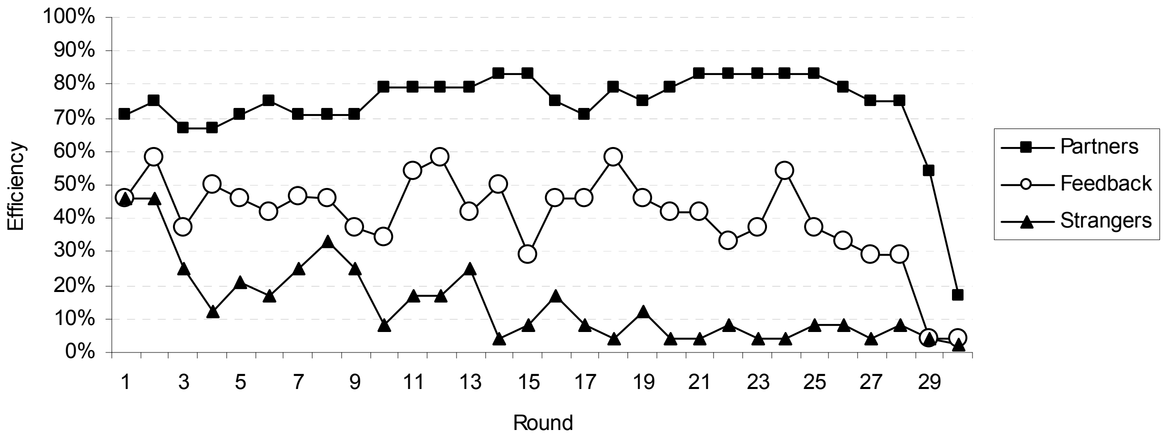


Figure 3. Trust measured as the percentage of buying per round

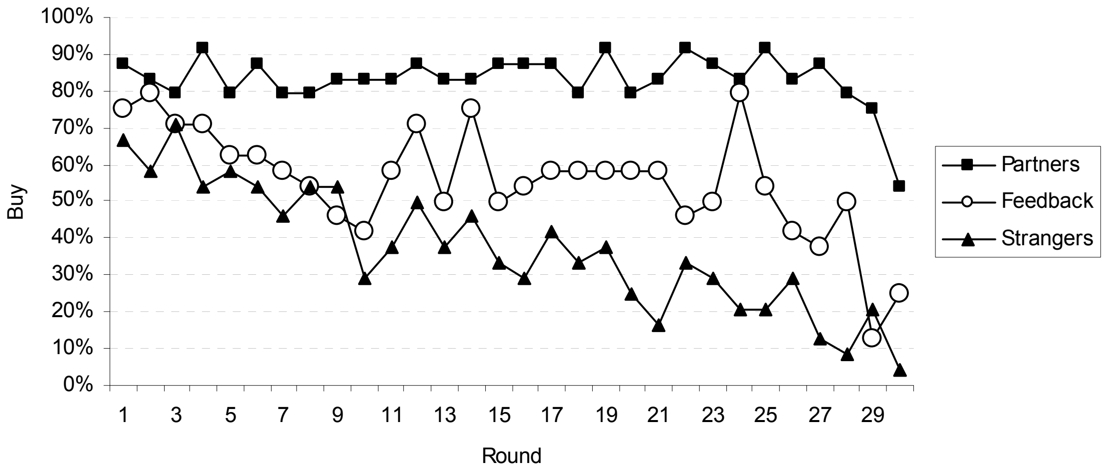
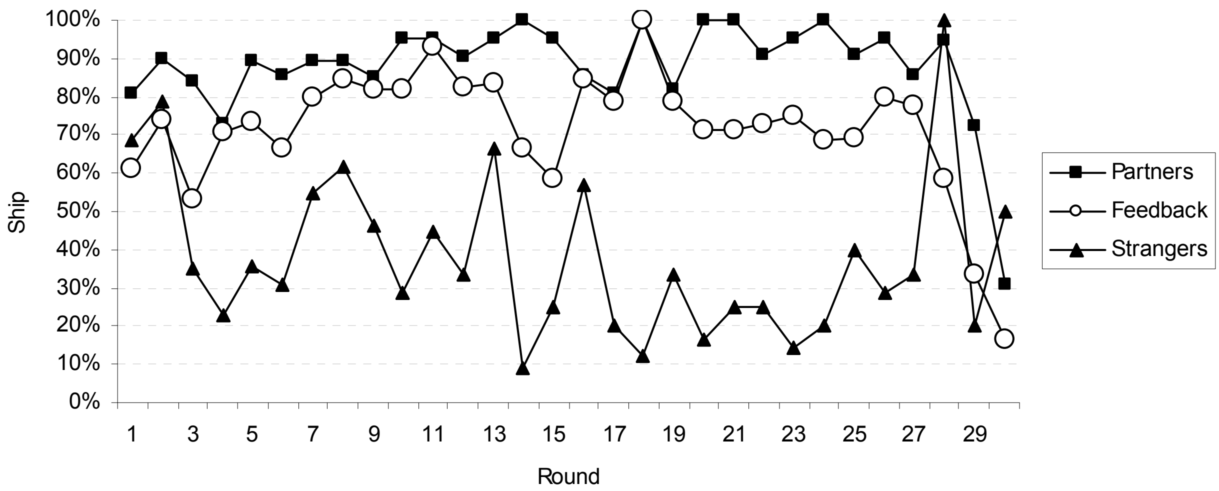


Figure 4. Trustworthiness measured as percentage of shipping per round



The same pattern is evident in all three figures: There is the least efficiency, trust and trustworthiness in the strangers market, more of all three in the feedback market, and more still in the partners market. For example, averaged over all rounds, feedback yields 2.8 times the efficiency of strangers, and partners yields 1.8 times the efficiency of feedback.

Comparing the treatments pair-wise, with sessions as the individual observations, support these conclusions (t-test p-value < 0.025 in all cases save for comparing average buying between strangers and feedback, p-value $= 0.08$, one-tailed tests, assuming equal variances). This rejects the information hypothesis.[10]

Note too from Figures 2-4, the marked differences across treatments in round-to-round trading dynamics. For the strangers market, there is a strong downward trend in efficiency, trust and trustworthiness. The trends for feedback and partners markets, however, appear to be quite stable, save for the last two rounds when trading collapses.

[10] A non-parametric rank test on session observations yields similar results, with $p = 0.05$ for all but the strangers-feedback comparison where $p = 0.10$ (one-tailed tests).

Table 1. Random effects probit models, buyers[a]

Maximum likelihood estimates [and two-sided p-values] for buyer behavior

Dependent variable = "1" for *buy*

Independent variable: Coefficient [p-value]	Model 1	Model 2	Model 3
CONSTANT	0.533	0.347	0.524
	[.0040]	[.0185]	[.0001]
FEEDBACK	-0.020	0.200	
= 1 if buyer is from feedback treatment, and 0 else.	[.9473]	[.4347]	
PARTNERS	0.963	1.48	0.852
= 1 if buyer is from partners treatment, and 0 else.	[.0001]	[.0000]	[.0011]
TOTALSHIPfeedback			0.0616
= number of seller ships prior to last order.			[.0014]
TOTALNOSHIPfeedback			-0.124
= number of seller no ships prior to last order.			[.0144]
SHIPLASTfeedback			0.212
= 1 if feedback seller shipped last order, and 0 else.			[.2111]
NSHIPLASTfeedback			-0.646
= 1 if feedback seller did not ship last order, and 0 else.			[.0005]
SHIPLASTpartners			1.330
= 1 if seller in partners shipped last order, and 0 else.			[.0000]
NSHIPLASTpartners			-.697
= 1 if seller in partners did not ship last order, and 0 else.			[.0100]
CBSH		0.045	-0.005[b]
= number of past times item was shipped to buyer.		[.0180]	[.8386]
CBNH		-0.412	-0.386[b]
= number of past times buyer bought but not shipped.		[.0000]	[.0000]
ROUNDstrangers	-0.062		
= round in strangers treatment, and 0 else.	[.0000]		
ROUNDfeedback	-0.019		
= round in feedback treatment, and 0 else.	[.0006]		
ROUNDpartners	0.006		
= round in partners treatment, and 0 else.	[.4806]		
LAST2ROUNDstrangers	-0.151	-0.390	-0.404
= 1 if round 29 or 30 in strangers treatment, and 0 else.	[.6414]	[.1649]	[.1671]
LAST2ROUNDfeedback	-0.903	-0.944	-0.974
= 1 if round 29 or 30 in feedback treatment, and 0 else.	[.0000]	[.0000]	[.0000]
LAST2ROUNDpartners	-1.15	-1.200	-1.322
= 1 if round 29 or 30 in partners treatment, and 0 else.	[.0000]	[.0000]	[.0000]
RHO	0.399	0.456	.444
(random effects)	[.0000]	[.0000]	[.0000]
Number of observations	2160	2160	2160
Log-likelihood	-1087.77	-1056.57	-988.67
χ^2 p-value	.0000	.0000	.0000

[a]Analogous estimates for fixed effects linear models are given in Table B1 of Appendix B.

[b]History for partners buyers does not include last transaction.

Model 1 of Table 1 refines our understanding of both the base treatment effects and trading dynamics. Model 1 estimates the probability a buyer will decide to buy using a random effect probit regression. The model includes three blocks of variables; one block for treatment effects (FEEDBACK, PARTNERS, with the default being strangers), another for ROUND effects, and a third for endgame effects (LAST2ROUND). Endgame effects include the last *two* rounds because, while player roles were switched randomly, they were told they would be a seller (buyer) for half the rounds, so that in round 29 a seller may be in his last round as a seller and thus has no strategic reason to be trustworthy. Random effects account for the effects due to the idiosyncrasies of the decision makers in the sample (Greene, 1992). The bracketed numbers in Table 1 are two-tailed p-values for the estimated coefficients. Analogous probit models for the seller behavior are omitted here but yield the same qualitative conclusions (see also Table 2 in section 4.2). Models 2 and 3 of Table 1 elaborate on Model 1 and are discussed in section 4.4.

The Model 1 FEEDBACK and PARTNERS coefficients permit a comparison of *initial* trust levels across treatments. The coefficient for FEEDBACK is small and insignificant, indicating that there is little initial difference in trust across strangers and feedback markets. The coefficient for PARTNERS, however, is positive and significant, indicating that partners market buyers are initially more trusting than their counterparts in strangers and feedback markets.

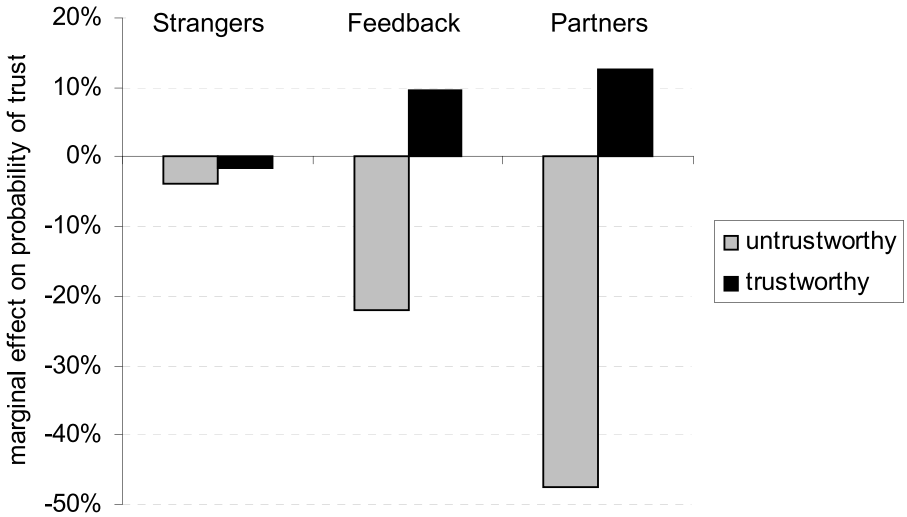
ROUNDstrangers captures the strong negative trend in trust across rounds of the strangers market. ROUNDfeedback is also significantly negative but substantially and significantly higher than ROUNDstrangers (two-tail p = 0.0106, Wald test). In contrast, the ROUNDpartners coefficient is slightly positive but not significant.

Hence, by Model 1, buyers in feedback markets are initially about as trusting as in strangers markets. Trust declines quickly, however, in strangers markets, and slowly in feedback markets. In contrast, trust in partners markets is stable at a level that is higher than in the other two markets. At the same time, there are large and significant end-game effects in both feedback and partners but not in strangers, indicated by the LAST2ROUND variables.

4.2 Comparing strangers and feedback: The strategic benefits of trust and trustworthiness

An online feedback mechanism can help promote trade *if* traders use reputation information strategically. Specifically, *if* buyers condition their behavior on the shipping history of their sellers, then this gives sellers a strategic incentive to avoid spoiling their reputation. Comparing buyer and seller behavior in the strangers market to that in the feedback market supports the view that buyer and seller behavior is largely, if imperfectly, strategic in nature.

Figure 5. Marginal trust conditioned on last feedback across treatments*



* The base rate (the zero line) is the average buy over all encounters for each treatment separately (37.08 percent in strangers, 55.56 percent in feedback and 83.22 percent in partners).

Figure 5 shows the marginal effects on the probability of trust in all markets depending on whether the seller shipped the *last* order ('trustworthy') or not ('untrustworthy'). In the strangers market, the marginal effect is close to zero since buyers do not see seller histories. (The effects are slightly negative since the bars do not include "newbies" − sellers who have not been trusted yet − and since buy rates in strangers decline rapidly.) Buyers in feedback markets, on the other hand, strongly condition their behavior on the seller's last feedback. In absolute terms, they trust with probability 33 percent if the seller did not ship the last order, and with about twice this probability, 65 percent, if he shipped the last order. (Formal statistical analysis for these effects is discussed in section 4.4.)

This conditional buying is sensible since the seller's history has predictive power for his future performance. Table 2 presents a random effect probit for the seller's decision. We see that shipping the last time for a feedback market seller who received a buy order is a significant predictor of whether the seller will do so this time. (The coefficient for LASTSHIPstrangers is significant as well but with a negative sign.)

Table 2. Random effects probit model, sellers

Maximum likelihood estimates [and two-sided *p*-values] for seller behavior

Dependent variable = "1" for *ship*

Independent variable	Coefficient [*p*-value]	Independent variable	Coefficient [*p*-value]
CONSTANT	0.190 [.2818]	SHIPLASTstrangers	-0.366 [.0713]
FEEDBACK	0.278 [.2666]	SHIPLASTfeedback	0.350 [.0168]
PARTNERS	0.557 [.0267]	SHIPLASTpartners	0.898 [.0000]
ROUNDstrangers	-0.037 [.0008]	LAST2ROUNDstrangers	-0.112 [.8711]
ROUNDfeedback	-0.005 [.6154]	LAST2ROUNDfeedback	-1.757 [.0066]
ROUNDpartners	-0.007 [.4589]	LAST2ROUNDpartners	-1.833 [.0000]
RHO	0.204 [.0000]	Number of observations	1267
		Log-likelihood	-561.63
		χ^2 *p*-value	.0000

[a]Variable interpretations are analogous to those given in Table 1.

Finally, the strong end-game effect exhibited by both buyers and sellers in feedback markets (Tables 1 and 2) also supports the view that it is the strategic incentive to ship created by the 'shadow of the future,' as opposed to, say, an intrinsic preference for being trustworthy, that largely drives the efficiency-enhancing effect of the feedback mechanism.

Of course, the fact that sellers in strangers ship some 36 percent of the time indicates that not all trustworthy behavior can be explained by strategic response to the pecuniary incentives. We might dismiss this as due to inexperience since most of the trading in strangers comes in the early rounds. However, even in the very final rounds of feedback and partners markets, some buyers still buy and some still ship. The buying might be attributed to naiveté, but the shipping is harder to explain this way since shippers know they are doing this for the last time. It would appear that some behavior reflects an intrinsic propensity for trusting or trustworthy behavior.

4.3 Comparing feedback and partners: The public benefits of trust and trustworthiness

Figures 2-4 show that partners markets are more successful in promoting exchange than feedback markets. Figure 5 additionally shows that buyers in the partners market respond more strongly to the sellers' histories than do buyers in feedback. Further, from Table 2, notice that a last decision to ship is more highly predictive of future shipping in partners than in feedback markets (two-tailed $p = 0.0121$, Wald test). This section investigates *why* reputation information has differential impact on trading patterns in feedback and partners markets.

The thrust of our arguments is that, unlike in a partners relationship, feedback and one's own past experience do not perfectly overlap in feedback markets. This creates effects of trusting and being trustworthy that are not internalized by the reputation mechanism in the feedback market as they are in the partners market. We first explain this phenomenon with some simple descriptive statistics to illustrate the points. We then discuss the analogous inferential evidence.

A trusting buyer in a feedback market generates valuable feedback information for *other* buyers who meet the same seller in the future. A trusting buyer in a partner relationship generates the same valuable feedback information − but entirely for himself. Thus, the informational benefits from trusting (as opposed to the pecuniary gains from the immediate trade) are internalized in transactions among partners but not in transactions among feedback market traders. Thus, all other things equal, the overall benefits from trusting are smaller in feedback. We call this the *informational dilemma*.[11]

[11] There is a related but distinct public goods problem of feedback provision observed by Resnick and Zeckhauser (2002), among others: Once the transaction is concluded, buyers have no incentives to provide others with feedback about their experience. (Resnick and Zeckhauser report that on eBay nevertheless feedback was provided in about half the time.) Since in our experiments feedback was produced automatically, this public good problem was not an issue (and in this sense our results overestimate the merits of a feedback mechanism that is based on *voluntary* feedback production).

In the data, the informational dilemma is particularly apparent if a buyer is matched with a newbie, a seller with no prior feedback history. The average trust in newbies ranges from 65 percent in strangers, to 77 percent in feedback, to 93 percent in partners. While the difference between the strangers market and the other markets can be explained by the fact that newbies in the strangers market have no strategic reason to be trustworthy, the difference between feedback and partners may be explained by the informational dilemma. In fact, buying from a newbie in the feedback market on average yielded a loss, namely 31 cents, less than the 35 cents from not buying. Buying from a seller who shipped the last order, on the other hand, yielded an average payoff of 40 cents, and buying from a seller who did not ship the last order yielded 17 cents.[12] (In contrast, trusting a newbie seller in the partners market had an expected value of 40 cents.) Thus, a buyer in the feedback market is, in fact, better off trusting somebody *only* if he or she has already been shown to be trustworthy.

Table 3. Buying probabilities

Marginal effects of the buyers' most recent history and the sellers' most recent decision on the buying probability (in percent)*

Market (average buy rate) / *Buyer's last experience*	*Seller's most recent decision* *Untrustworthy*	*Trustworthy*
Strangers (37 percent)		
Exploited	− 9.91	− 7.63
Rewarded	6.16	5.61
Feedback (56 percent)		
Exploited	− 27.35	2.20
Rewarded	− 20.01	11.34
Partners (83 percent)		
Exploited	− 47.62	n/a
Rewarded	n/a	13.42

* Zero level is normalized in each treatment to the respective average buy rate (as shown in parentheses), taken over all encounters.

In feedback markets, there are not only positive externalities from trust but also from trustworthiness. The critical evidence for this comes from examining buyers' reactions to their *own* history. Extending the scope of Figure 5, Table 3 provides some sense of the relative

[12] Four buyers in feedback never faced a newbie and so are not included in these statistics.

importance of own experience, defined either as ‘rewarded’ if the buyer’s *last* order was shipped, or ‘exploited’ if it was not shipped. Table 3 shows that, in all three markets, if a buyer was treated well in the past, he is more likely to trust in the future.

There is, however, on display in Table 3, a difference between feedback and partners markets: In partners markets, histories are perfectly aligned. That is, a buyer’s trust is rewarded if and only if his seller is trustworthy, so that the two history effects always cumulate. In feedback markets, on the other hand, the reputation effect is diluted by buyer’s experience based on his own history. A trustworthy seller is less trusted when the buyer’s own experience was bad and an untrustworthy seller is more trusted when the buyer’s own experience was good. Thus, there is a wedge driven between buyers’ and sellers’ histories in feedback markets that is (at least partly) responsible for why the marginal response to both the seller’s positive and negative feedback is on average weaker in feedback compared to partners (as shown in Figure 5).

Although the wedge does not *directly* explain why overall average trust is lower in feedback, it does explain why sellers under-invest in trustworthiness, which in turn should lead to less trust: A seller who ships in the feedback market benefits only through the improved own reputation, but *other* sellers will profit from the induced good history of his buyer. On the other hand, a seller who ships in partners contributes to his own good reputation and to a good history of *his own* future buyer. In other words, the trust-inducing effects of trustworthiness are fully internalized in partners but not in feedback. As a consequence, the overall private benefit from trustworthiness is weaker in feedback markets, which reduces the incentive to trade of both sellers and buyers.

4.4 A probit model of buyer behavior

Models 2 and 3 in Table 1 provide formal inferential support for the findings in the last section. They also provide some further insight into how buyers evaluate seller histories.

Model 2 replaces the round effect variables with variables reflecting the buyer’s history with shipping: the cumulative number of times he has bought and had the item shipped (CBSH) and the cumulative number of buys that were not shipped (CBNH). Comparing Model 2 to Model 1 permits a baseline check of the hypothesis that the experience effects we see in Model 1 are due to differences in experiences with shippers against the omnibus hypothesis that they are due to differing market structure. For example, there is less information for buyers to process in strangers, and this, instead of experience with shippers *per se*, might explain why the experience effect is greater in strangers than in the other treatments (see section 4.1). If differing market structure were in fact responsible for the dynamic across rounds, then we would expect both C··H coefficients in Model 2 to be about the same, and certainly Model 2 should fit no better than Model 1. In fact, the C··H coefficients differ sharply and the likelihood of Model 2 is higher than that of Model 1, indicating that buyer experience with shippers is a good explanation for the differing dynamics across markets.[13] What we see in Model 2 is that a single negative experience – making a purchase and having a seller fail to ship – erodes buyer trust in the market, while a single positive experience builds a smaller amount of trust.

The influence of buyer history persists in Model 3, where variables are added to reflect the information buyers have about sellers at the time they decide whether to purchase. For the feedback market, we add variables to reflect the cumulative shipping history of the seller (TOTALSHIPS and TOTALNOSHIPS) and the most recent shipping history (SHIPLAST and NSHIPLAST). In the partners market, the seller's cumulative history is already reflected in the buyers C··H variables. So for partners, we need add only most recent history variables. The FEEDBACK variable is dropped because it is not significant in the other two models.[14]

All of the information coefficients in Model 3 have the expected signs, save that for CBSH, but this estimate is small and not significant. There are three main observations. First, and most important for the present exposition, the CBNH variable is significantly negative,

[13] Models 1 and 2 are not nested and so not conducive to a formal inference test of the buyer experience hypothesis. If we re-estimate Model 1 with three CNBHx variables, one for each treatment x, all three of the new variables are negative and significant (two-tail $p < 0.01$ in all cases). A full model, with history variables broken out by treatment, is given in Table B2 of Appendix B.

[14] Including it would make little difference to the estimates of Model 3 and the coefficient of FEEDBACK would still not be significant. There are, however, some indications that FEEDBACK is collinear with SHIPLASTfeedback and NSHIPLASTfeedback.

verifying that seller behavior has the discussed externality. Second, both feedback and partner buyers weigh recent observations more heavily than older ones (p = .0000 for both markets, Wald test). Third, in feedback markets, NSHIPLAST has a more reliable effect on buyer decision than SHIPLAST, whereas partner buyers weight recent positive and negative history about the same.

Model 3 also captures the information wedge effects: First, buyers with the same buying history are more likely to trust newbies (who have no history) in the partners market than in either the strangers or feedback markets, as indicated by the coefficient for PARTNERS. Second, a partners seller who has been trustworthy in the recent past is granted higher trust than a feedback seller (SHIPLASTpartners is greater than SHIPLASTfeedback, two-tail p = 0.0007, Wald test), thus partners sellers have the greater incentive to be trustworthy. Third (and related to the second effect), a decision *not to ship* has very similar immediate total negative effect on buying in both feedback and partners markets (compare NSHIPLASTfeedback and NSHIPLASTpartners), but that cost is borne by the seller only in the partners market.

4.5 Payoffs

One would correctly conclude from the analysis of the efficiency reached by the different markets in section 4.1 that average payoffs are smallest in strangers, larger in feedback, and larger still in partners. In fact, the strategic incentives created by feedback systems also translate into positive *correlations* between trust(worthiness) and payoffs. The Spearman rank correlation coefficients between subjects' total payoffs and the total number of ship and buy in the strangers markets is negative (−0.307, p = .034, for the frequency of buy and −0.038, p = .795, for ship), making all trade activities unprofitable. Sixteen out of 48 subjects receive total payoffs that are smaller than in a market in which nobody ever buys. A feedback mechanism, on the other hand, makes trust and trustworthiness lucrative behavior. The corresponding coefficients in the feedback market are significantly positive (.288, p = .047, for buy and .504, p = .000, for ship) and not a single subject made payoffs smaller than in a market with no trust. Finally, the pecuniary incentives to trade are strongest in partners markets (.706, p = .000, for buy and .728, p = .000 for ship). While two subjects in this market received less than they could expect when no exchange ever takes place (both subjects never shipped to their partners but vainly tried to buy from them), 34 buyers and 31 sellers did their part of the exchange 14 or 15 times, numbers that are reached only once in strangers and feedback together.

5. Conclusions

The findings from our experiment have a number of implications for the design of online feedback systems, as well as for further research into online reputation mechanisms and research into reputation building in general.

One of the key findings concerns how a buyer's own experience affects his trust in the entire market. Specifically, a buyer whose trust has been betrayed tends to have diminished trust in all sellers, whereas a buyer whose trust has been rewarded continues to trust the market at about the same level as he did before. In practice, online feedback is vulnerable to manipulation, and so less reliable than the feedbacks truthfully generated in our experiment.[15] The value of own experience in evaluating the trustworthiness of the market is therefore likely greater than our finding already suggests. An important implication for online market design is that the public value of feedback information might be increased by informing market participants about the shipping probability in the *entire* market, and not only about the trustworthiness of individual traders. Information indicating positive overall trust in the market might mitigate the negative effects of a trader's own bad experience (while, of course, a negative overall market assessment might serve to aggravate it).

We also found that buyers in online markets are *rightly* reluctant to bear the costs associated with verifying the trustworthiness of newbies. In practice, many online market

[15] eBay distinguishes 4 forms of fraudulent feedbacks: defensive and offensive shill feedback (using secondary eBay user IDs or other eBay members to artificially raise the level of your own feedback or to leave negative comments for another user), feedback extortion (demanding any action of a fellow user that he or she is not required to do, at the threat of leaving negative feedback), and feedback solicitation (offering to sell feedback, trade feedback undeservedly, or buy feedback); see http://pages.ebay.com/help/community/investigates.html.

participants can change their identity at no costs, making it impossible for buyers to distinguish a 'real' newbie, trading for the first time, from a 'deceptive' newbie, an experienced seller who got rid of his bad reputation. The probability that a newbie is not trustworthy is therefore likely to be higher on real life platforms than on our laboratory platform. The important implication is that market platforms should try to gain control over the agents' identities (see Friedman and Resnick, 2001, for more theoretical reasons along these lines and how control can be realized; see also Ba, 2001, and Ba et al., 2003).

Another design implication from our study comes from our finding that buyers put more weight on negative than positive feedback, and more on recent than old feedback. The emphasis on recent negative feedback has also been reported in field studies (Lucking-Reiley et al., 1999, and Resnick and Zeckhauser, 2002, among others). In an experimental investigation, Keser (2002) finds that providing traders with only the most recent feedback information has by itself a significant impact on trust and trustworthiness, though efficiency is, in an intermediate phase of interaction, higher when trading partners are informed of the entire distribution of each other's previous feedbacks.[16] Together, these findings suggest that relying solely on a cumulative measure of trustworthiness may not be appropriate because it hides information critical to the buyers' decision to trust. Another problem with relying solely on cumulative feedback measures comes from the fact that the seller's actions have diminishing impact in influencing the buyers' assessment of the trustworthiness. This typically leads to increasing incentives to exploit one's good reputation (see Holmstrom, 1999, for a model along these lines in a different context).

We also found that our market participants exhibit a strong, systematic response to the strategic incentives created by a feedback mechanism. In other words, most of the trust and trustworthy behavior we observed appears to flow from the incentives in the environment as opposed to some intrinsic drive to be trusting or trustworthy. Nevertheless, we did identify some behavior – particularly sellers who shipped in the final rounds of play – that is not easily

[16] Keser's findings are, however, not straightforwardly comparable to our results because her experimental design differs in a number of design choices, such as the game itself (e.g., efficiency gains occurred after the buyer's move and were substantially higher), the matching and role assignment scheme (e.g., trading partners were matched more than once), feedback provision (voluntarily and endogenously given by buyers), etc.

explained by the incentives. One of the important tasks for further research is a more detailed characterization of the nature and extent of this intrinsic incentive. All present online feedback mechanisms rely on intrinsic motivation to some degree. For example, the mechanisms typically provide participants with no incentive to take the time to report their experiences and to report honestly. Since this sort of information is the heart of the system, it is important to know the extent to which intrinsic motivation is sufficient to provide it (Miller et al., 2003, recently proposed a payment-based system for getting buyers to rate transactions with sellers honestly and frequently).

Our data rejects the information hypothesis and is thus inconsistent with standard economic theories of reputation building. At the same time, our findings point to the reasons for the failure and thus suggest ways to capture the impact of information flows in new, extended models. To be more precise, our finding that a seller's reputation profile has predictive value for his future behavior, causing the information dilemma, is inconsistent with the sequential equilibrium model outlined in section 2.2. In standard sequential equilibrium approaches to our game, reputation information, though critical for the emergence of trust, is *not* valuable in deciding whether to trust or not: In the early phase of the finitely repeated game all sellers always cooperate regardless of their preference-type (trustworthy or strategic) so that in equilibrium early play cannot reveal valuable information. In all other periods, strategic sellers either play a mixed strategy such that buyers are indifferent between trusting and not trusting, or, once their reputation profile proves them as strategic players, never ship anymore. Thus, in this second phase, a buyer cannot make more than his outside option regardless of his seller's reputation profile. However, the dynamics of reputation effects are known to be sensitive to the set of preferences that exists (see, e.g., Diamond, 1989). It is conceivable that in a model in which some sellers are committed to behave trustworthy in all encounters, but some are committed to behave always untrustworthy (with the rest behaving strategically), a seller's reputation might have predictive power, and so be more consistent with the data we have here.[17]

[17] Avery et al. (1999) developed a related model of evaluations for goods with fixed but initially unknown quality. The more feedback is available, the less the uncertainty about the product's true quality. It may be unprofitable to

In addition, our observation that buyers condition their behavior on their own experience is not in line with sequential equilibrium approaches if one assumes that the seller-type distribution is commonly known. However, an extended model that allows buyers to update their beliefs about the distribution as they gain experience might capture that buyers are more willing to trust after a positive experience, and less willing to trust if trust was exploited, and thus would be sensitive to the trader matching scheme.[18]

We conclude by elaborating on a comment we made in the introduction, that goes to the importance of reputation building in general: Granovetter (1985) argues that people have greater confidence in information when it comes from "a trusted informant" that has dealt with the individual in question than when the same information comes from a stranger or an institution, and that they have even greater confidence in information gained from first hand dealings. Specifically, Granovetter argues that first hand information (a) is usually cheaper to gather, (b) is often more detailed, (c) offers the promise of future business that provides a great motivation to be trustworthy, and (d) often become overlaid with social content. These factors are likely important in practice, but they do not explain our data. The essence of our findings is that buyer trust provides information to the market about individual sellers' trustworthiness. Seller trustworthiness enhances individual buyers' confidence in the marketplace. Even if reputations are shared and reliable, these externalities will not be internalized in transactions among strangers, so that both trust and trustworthiness will be underprovided.

---

engage in behavior that would reduce uncertainty, and therefore reputable sellers may never be discovered. Dellarocas (forthcoming) surveys theories of reputation and relates them to online feedback mechanism.

[18] There are also non-strategic modeling approaches that are in line with the role of the buyers' histories. First, adaptation theories predict that people tend to choose strategies that performed well in the past (such as Roth and Erev's, 1995, reinforcement learning theory or Selten's, 1988, learning direction theory). Hence, if trust was rewarded it has a higher probability of being chosen again. Second, the own experience effect may reflect a non-strategic (backward-looking) reciprocal motive. Market participants may not be willing to cooperate in a market in which they were exploited. This argument appears to have more bite in the partners market where *not buy* can be straightforwardly interpreted as a reciprocal punishment for unfair behavior against oneself. The fact that the own experience effect also occurs in the reputation market suggests, however, that this unfairness aversion is also relevant among strangers (similar observations have been made by Blount, 1995, and Bolton et al., 2002, among others).

References

Alexander, R.D. (1987). *The Biology of Moral Systems*. New York: Aldine De Gruyter.

Ariely, D., A. Ockenfels, and A.E. Roth (2002). "An Experimental Analysis of Ending Rules in Internet Auctions." Working Paper, Harvard Business School.

Avery, C., P. Resnick, and R. Zeckhauser (1999). "The Market for Evaluations." *American Economic Review*, 89, 564-584.

Ba, S. (2001) "Establishing Online Trust through a Community Responsibility System." *Decision Support Systems*, 31(3), 323-336.

Ba, S., and P.A. Pavlou (2002). "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior." *MIS Quarterly*, 26, 243-268.

Ba, S., A.B. Whinston, and H. Zhang (2003). "Building Trust in Online Auction Markets through an Economic Incentive Mechanism." *Decision Support Systems*, 35(3), 273-286.

Blount, S. (1995). "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences." *Organizational Behavior and Human Decision Processes*, 63, 131-144.

Bolton, G.E., A.E. Katok, and A. Ockenfels (2002). "Cooperation among Strangers with Limited Information about Reputation." Working Paper, Smeal College of Business Administration, Penn State University.

Bolton, G.E, and A. Ockenfels (2000). "ERC: A Theory of Equity, Reciprocity and Competition." *American Economic Review*, 90, 166-193.

Bolton, G.E, and A. Ockenfels (2003). "Strategic Reputation Building as a Function of Information and Matching in Sequential Equilibria." Mimeo.

Brynjolfsson, E., and M.D. Smith (2000). "Frictionless Commerce? A Comparison of Internet and Conventional Retailers." *Management Science*, 46, 563-585.

Buchan, N., R. Croson, and R. Dawes (2002). "Swift Neighbors and Persistent Strangers: A Cross-Cultural Investigation of Trust and Reciprocity in Social Exchange." *American Journal of Sociology*, 108, 168-207.

Dellarocas, C. (2001). "The Design of Reliable Trust Management Systems for Online Trading Communities." Working paper, MIT Sloan School of Management.

Dellarocas, C. (forthcoming). "The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms." *Management Science*.

Diamond, D. (1989). "Reputation Acquisition in Debt Markets." *Journal of Political Economy*, 97, 828-862.

eBay (2002). http://pages.ebay.com/services/forum/feedback.html.

Ellison, G. (1994). “Cooperation in the Prisoner’s Dilemma with Anonymous Random Matching.” *Review of Economic Studies*, 61, 567-588.

Fehr, E., and K. Schmidt (1999). “A Theory of Fairness, Competition, and Cooperation.” *Quarterly Journal of Economics*, 114, 817-868.

Friedman, E., and P. Resnick (2001). "The Social Cost of Cheap Pseudonyms." *Journal of Economics and Management Strategy*, 10(2), 173-199.

GarnerG2 (2002). “Online Transaction Fraud and Prevention Get More Sophisticated.” http://www.gartnerg2.com/rpt/rpt-0102-0013.asp.

Granovetter, M. (1985). "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology*, 91, p. 481-510.

Greene, W.H. (1993). *Econometric Analysis*, 2nd edition, MacMillan: New York.

Greif, A. (1989). “Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders.” *Journal of Economic History*, 49, 857-882.

Güth, W. and A. Ockenfels (forthcoming). “The Coevolution of Trust and Institutions in Anonymous and Non-anonymous Communities.” In: M.J. Holler, H. Kliemt, D. Schmidtchen and M. Streit (eds.), *Jahrbuch für Neue Politische Ökonomie*, 20, Tübingen: Mohr Siebeck.

Holmstrom, B. (1999). “Managerial Incentive Problems: A Dynamic Perspective.” *Review of Economic Studies*, 66, 169-182.

Houser, D. and J. Wooders (2001). “Reputation in Auctions: Theory and Evidence from eBay.” Working Paper, University of Arizona.

Keser, C. (2002). “Trust and Reputation Building in E-commerce.” Working paper, IBM, New York.

Kreps, D.M., P. Milgrom, J. Roberts, R. Wilson (1982). “Rational Cooperation in the Finitely Repeated Prisoners' Dilemma.” *Journal of Economic Theory*, *27*, 245-252.

Lucking-Reiley, D., D. Bryan, N. Prasad, and D. Reeves (1999). “Pennies from eBay: the Determinants of Price in Online Auctions.” Working paper, Vanderbilt University.

McMillan, J. (2002). *Reinventing the Bazaar: A Natural History of Markets*, New York: Norton.

Miller, N., P. Resnick, and R. Zeckhauser (2003). „Eliciting Honest Feedback in Electronic Markets.” Working paper, Kennedy School of Government, Harvard University.

Melnik, M.I., and J. Alm (2002). “Does a Seller’s E-commerce Reputation Matter?” *Journal of Industrial Economics*, 50(3), 337-49.

Milgrom, P., D. North, and B. Weingast (1990). “The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs.” *Economics and Politics*, 2, 1-23.

Ockenfels, A. (2003), “Reputationsmechanismen auf Internet-Marktplattformen: Theorie und Empirie.“ *Zeitschrift für Betriebswirtschaft*, 73(3), 295-315.

Resnick, P. and R. Zeckhauser (2002). “Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay’s Reputation System.” In: M.R. Baye (ed.), *The Economics of the Internet and E-Commerce*, Volume 11 of Advances in Applied Microeconomics, Amsterdam, Elsevier Science.

Resnick, P., R. Zeckhauser, J. Swanson, and K. Lockwood (2002). “The Value of Reputation on Ebay: A Controlled Experiment.” Working Paper, University of Michigan.

Roth, A. E. (2002). “The Economist as Engineer: Game Theory, Experimental Economics and Computation as Tools of Design Economics.” Fisher Schultz lecture, *Econometrica*, 70, 1341-1378.

Roth, A.E., and I. Erev (1995). “Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term.” *Games and Economic Behavior*, 8, 164-212.

Selten, R. (1998). “Features of Experimentally Observed Bounded Rationality.” Presidential Address, *European Economic Review*, 42, 413-36.

U.S. Department of Justice (2002). “Internet Fraud.” http://www.internetfraud.usdoj.gov/#What Are the Major Types of Internet.

U.S. White Collar Crime Center and the Federal Bureau of Investigation (2001). “IFCC 2001 Fraud Report.” http://www1.ifccfbi.gov/strategy/statistics.asp.

Appendix A. Subject instructions and buyer screen

BELOW ARE THE WRITTEN INSTRUCTIONS THAT WERE GIVEN TO SUBJECTS IN THE FEEDBACK TREATMENT. INSTRUCTIONS FOR THE OTHER TREATMENTS WERE PARALLEL, THE ONLY DIFFERENCES BEING THE DESCRIPTION OF THE FEEDBACK SYSTEM (REMOVED FOR STRANGERS), OR THE DESCRIPTION OF PARTNER ROTATION (PARTNERS). The partners sessions were held on 3/27/2002 1 PM, 3/27/2002 2 PM and 3/39/2002 10 AM; the feedback sessions were held on 3/15/2002 1 PM, 3/15/2002 2 PM, and 4/08/2002 4 PM; the strangers sessions were held on 3/21/2002 9 AM, 3/21/2002 10 AM, and 3/29/2002 11 AM.

General. The purpose of this session is to study how people make decisions. If at any time you have questions, raise your hand and a monitor will happily assist you. From now until the end of the session, unauthorized communication of any nature with other participants is prohibited.

During the session you will play a game that gives you an opportunity to earn cash. At the end of the session, you will be paid your earnings plus a $5 show-up fee. Decisions and payments are confidential: No one will be told your actions or the amount of money you make.

[The figure that appeared here is the same as Figure 1 in the text of the paper.]

Description of the game. You and the other participants in the room (but not the monitors) are the players in the game. The game proceeds in a series of rounds. Each round, each player is randomly matched with another player to trade a (fictional) commodity. First, one of the players, the "Buyer," chooses to either *buy* or *not buy*. If the Buyer chooses *not buy*, then the game ends and both players receive $.35. If the Buyer chooses to *buy*, then the game continues and the other player, the "Seller," makes a decision to *ship* or *not ship*. *Ship* pays each player $0.50 while *not ship* pays the Buyer nothing and the Seller $0.70.

The game will last for 30 rounds. You will be a Buyer for half of the rounds, and a Seller for the other half. When you switch between roles is a matter of random chance, so you may be in one role for more than one round in a row before switching to the other role, and the pattern of switching may be different for you than for other players in the game.

Seller's feedback history. For each game played, the computer will record whether the Seller chose *ship* or *not ship* (if the Seller did not get to move, the computer records nothing). This feedback will then be made available to all future Buyers that are matched with this Seller. The feedback will include a summary of the number of times the Seller shipped in the past, as well as a round-by-round history of their shipping decisions, beginning with the most recent decision. Buyers will see this feedback history prior to making their buy decision.

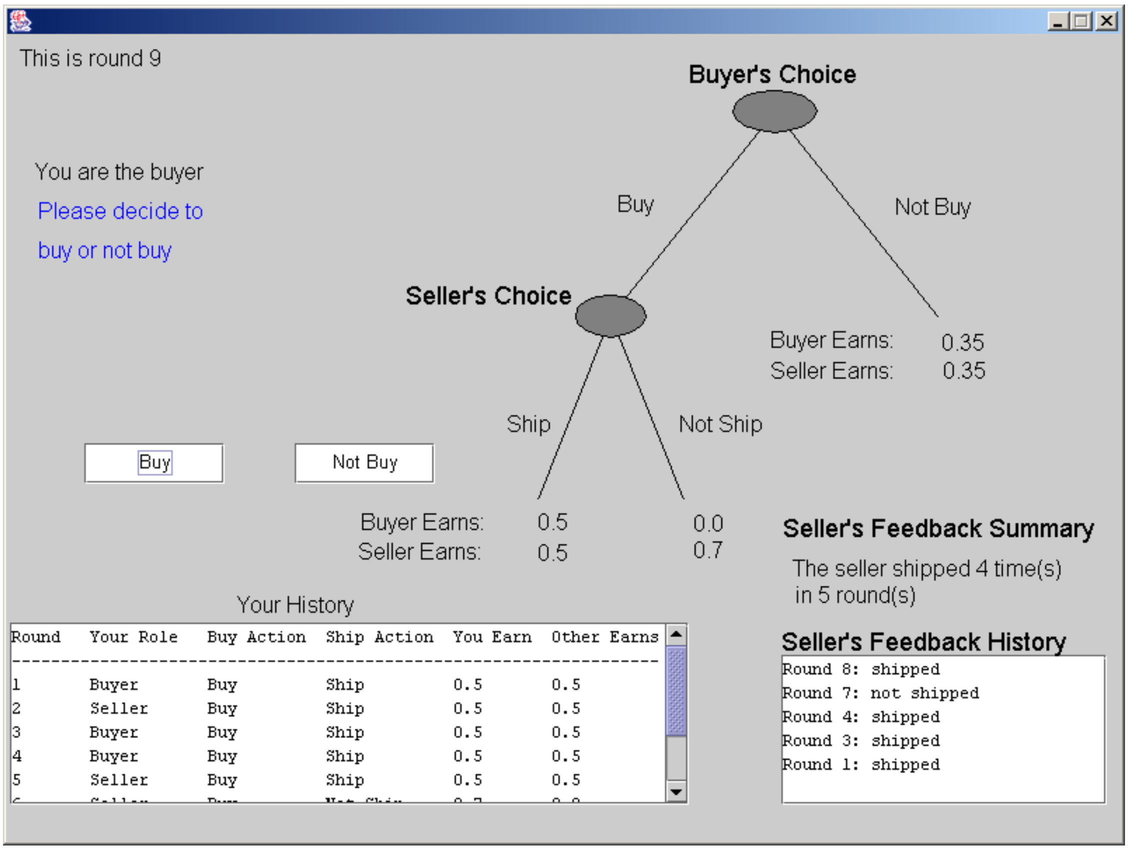
Pairings. All partner pairings are anonymous: Your identity will not be revealed to the person you are playing with either before, during or after the game. You will never be matched with the same player in the same role more than once.

Money earnings. You will be paid your earnings from all of the rounds of the game (plus a $5 show-up fee) in cash.

Practice games. When the monitor gives the OK, play some practice games. Your partner for the practice games will be the computer. It has been programmed to choose its moves at random. The practice games will allow you to experience the game from both the Buyer and Seller's perspective. Practice until you feel comfortable with the game and its rules.

Consent Forms. If you wish to participate in this study, please read and sign the accompanying consent form. The consent form explains your rights as a subject as well as the rules of confidentiality that will be adhered to regarding your participation.

Figure A1: A typical buyer screen



Appendix B.

Table B1. Fixed effects linear models, buyers[a]

OLS estimates (and two-sided *p*-values) for buyer behavior

Dependent variable = "1" for *buy*

Independent variable	Model 1	Model 2	Model 3
CONSTANT	---	---	---
FEEDBACK = 1 if buyer is from feedback treatment, and 0 else.	---	---	---
PARTNERS = 1 if buyer is from partners treatment, and 0 else.	---	---	---
TOTALSHIPSfeedback = number of seller ships prior to last order.			0.021 (.0000)
TOTALNOSHIPfeedback = number of seller no ships prior to last order.			-0.035 (.0092)
SHIPLASTfeedback = 1 if feedback seller shipped last order, and 0 else.			0.009 (.8697)
NSHIPLASTfeedback = 1 if feedback seller did not ship last order, and 0 else.			-0.261 (.0000)
SHIPLASTpartners = 1 if seller in partners shipped last order, and 0 else.			0.167 (.0023)
NSHIPLASTpartners = 1 if seller in partners did not ship last order, and 0 else.			-0.224 (.0005)
CBSH = number of past times item was shipped to buyer.		0.000 (.9103)	-0.006[b] (.1428)
CBNH = number of past times buyer bought but not shipped.		-0.133 (.0000)	-0.129[b] (.0000)
ROUNDstrangers = round in strangers treatment, and 0 else.	-0.184 (.0000)		
ROUNDfeedback = round in feedback treatment, and 0 else.	-0.007 (.0003)		
ROUNDpartners = round in partners treatment, and 0 else.	0.001 (.6651)		
LAST2ROUNDstrangers = 1 if round 29 or 30 in strangers treatment, and 0 else.	-0.005 (.9434)	-0.069 (.2393)	-0.070 (.2227)
LAST2ROUNDfeedback = 1 if round 29 or 30 in feedback treatment, and 0 else.	-0.301 (.0000)	-0.263 (.0000)	-0.263 (.0000)
LAST2ROUNDpartners = 1 if round is 29 or 30 in partners treatment, and 0 else.	-0.213 (.0010)	-0.136 (.0225)	-0.111 (.0610)
Number of observations	2160	2160	2160
Adjusted R-squared	0.368	0.401	0.442
F-test *p*-value	.0000	.0000	.0000

[a]These are analogous estimates for Table 1 in the text.

[b] History for Partner's buyers does not include last transaction.

Table B2. Random effects probit models, buyers[a]

Maximum likelihood estimates and two-sided *p*-values for buyer behavior

Dependent variable = "1" for *buy*

Independent variable	Coefficient	*p*-value
CONSTANT	0.512	0.0008
PARTNERS	0.835	0.0016
SHIPTOTALfeedback	0.077	0.0009
NOSHIPTOTALfeedback	-0.089	0.1194
LASTSHIPfeedback	0.270	0.1520
LASTNOSHIPfeedback	-0.566	0.0073
LASTSHIPpartners	1.184	0.0007
LASTNOSHIPpartners	-0.776	0.0127
CBSHstrangers	0.003	0.9691
CBSHfeedback	-0.067	0.1171
CBSHpartners	0.055	0.2151
CBNHstrangers	-0.382	0.0000
CBNHfeedback	-0.334	0.0000
CBNHpartners	-0.474	0.0000
LAST2ROUNDSstrangers	-0.420	0.1896
LAST2ROUNDSfeedback	-0.926	0.0001
LAST2ROUNDSpartners	-1.649	0.0000
RHO	0.442	0.0000
Number of observations		2160
Log-likelihood		-985.78
F-test *p*-value		.0000