Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System

Paul Resnick, presnick@umich.edu*
Richard Zeckhauser, richard_zeckhauser@harvard.edu

Abstract

Reputations that are transmitted from person to person can deter moral hazard and discourage entry by bad types in markets where players repeat transactions but rarely with the same player. On the Internet, information about past transactions may be both limited and potentially unreliable, but it can be distributed far more systematically than the informal gossip among friends that characterizes conventional marketplaces.

One of the earliest and best known Internet reputation systems is run by eBay, which gathers comments from buyers and sellers about each other after each transaction. Examination of a large data set from 1999 reveals several interesting features of this system, which facilitates many millions of sales each month. First, despite incentives to free ride, feedback was provided more than half the time. Second, well beyond reasonable expectation, it was almost always positive. Third, reputation profiles were predictive of future performance. However, the net feedback scores that eBay displays encourages Pollyanna assessments of reputations, and is far from the best predictor available. Fourth, although sellers with better reputations were more likely to sell their items, they enjoyed no boost in price, at least for the two sets of items that we examined. Fifth, there was a high correlation between buyer and seller feedback, suggesting that the players reciprocate and retaliate.

1. Introduction

The Internet is about scale. By virtually any metric, orders of magnitude more people are connected to each other, and communicate cheaply with each other, than at any time in history. However, many of these communications are among strangers, people who do not know each other before they receive a communication, learn little about each other from the communication, and do not encounter each other again. Where such encounters involve merely e-mail or chat room messages, such exchanges are not surprising. Risks are small so not much trust is required.

What is surprising is the vast shuttling of both new and second hand goods among distant strangers on the Internet, through such mechanisms as eBay and the Yahoo auction site. Buyers, who must pay before inspecting or receiving their items, must put considerable

* We gratefully acknowledge financial support from the National Science Foundation under grant number IIS-9977999. We also thank eBay for providing data, and especially Patrick Firouzian and his team for doing the data extraction in the midst of many other responsibilities. Mihir Mahajan, Ko Kuwabara, and Kate Lockwood provided valuable research assistance. The participants in the NBER workshops on empirical studies of ecommerce provided useful comments at several stages in the development of the paper.

dollars at risk. This paper seeks to explain why buyers trust unknown sellers in this vast electronic garage sale. For data, we shall be drawing on all the transactions on the eBay auction site from February through June 1999. The scale of these operations is impressive. eBay's site now boasts an average of more than 5 million active auction listings and though there has been growth over time our complete dataset consists of millions of items.[1]

The basic assertion of this paper is that trust has emerged due to the feedback or reputation system employed by eBay and other auction sites. The primary task of this paper is to examine how and why that system works.

Trust in traditional transactions among strangers

This new phenomenon of trust among strangers in Internet auction exchanges relies on a reputation system that is fundamentally different from those that human societies have evolved over thousands of years to create trust, particularly trust in economic transactions. Our focus is on transactions at the retail level, namely when the scale and dollar volume of transactions is limited.

How is trust traditionally created when goods are exchanged? We identify eight factors; readers would add more. (1) Most retail transactions are conducted locally, which gives individuals the opportunity to inspect them, as say with fruit in a rural market. If quality is discernible, no trust is needed. (2) Retail operations tend to be large relative to their local market, be they vegetable sellers or the local department store. Buyers have frequent interaction with the same seller, and learn whom they can trust. (3) Even when one's personal interactions are limited, given that a retailer's sales are concentrated in a locale makes it easy to develop reputations so customers learn about retailers from their peers. (4) Retailer reputations are borrowed from other contexts. For example, retailers are likely to be pillars of the church and community, and would be highly reluctant to sacrifice the status that comes from such reputations.[2] (5) Reputations are built over many years; witness the reputations of Sotheby's and Christies, the leading auction houses, which are hundreds of years old. (6) Reputations are borrowed from others. Thus celebrities will attest to the quality of products. (7) New goods benefit from established brand names, and policing of quality by those who own them. The product, not the retailer, wins the reputation. (8) Significant expenditures – e.g., building a fancy store on Manhattan's Fifth Avenue[3] -- indicates that one will be reliable, lest this expenditure be wasted, a form of signaling.

Internet auctions have none of these mechanisms available. Sellers are not met, and little or nothing is known about their characteristics, or even their location beyond its city.

---

[1] We are not permitted to reveal the exact number of transactions per day in our data set.

[2] In recent years, a literature on social capital has documented the many positive effects of civic activities and informal social ties, including their effects on trust building. Robert Putnam's book Bowling Alone reviews this literature and also documents declining social capital in the United States from the 1960s to the mid-nineties (Putnam 2000).

3 Recently, H&M, a Swedish retailer of high quality but little known in the United States, had round the block lines at the opening of such a store.

Customers rarely repeat, and they do not run into each other. Putting items on the Web is a cheap activity. Some goods that are traded are not brand name, and when they are there is a risk of being counterfeit. Measured in relation to the age of significant retail operations, all of the sellers are new. No one attests about the sellers. Firms like eBay do not stand behind their auctioneers. Yet millions of transactions have taken place.

What has substituted for the traditional mechanisms that establish trust? The argument we develop below is that the Internet substitutes a much better distribution of what information there is for the much more limited, but more reliable information of traditional retail markets. At least as judged by sales volume, the system appears to be working.

Though the system appears to be working, none of its participants know exactly how it is working or what its properties might be. For example, we found that just over half of buyers provided feedback. Presumably, these buyers comprise an unrepresentative sample. If they are merely individuals who find it cheap to do so, the bias might not be severe. However, it may be that dissatisfied customers are substantially less likely to give feedback. If so, since the overwhelming majority of feedback is positive, the most important information is being lost. Similarly, there is no known correlation of feedback with the price of the transaction. Conceivably sellers are honest with small transactions, but deceive (cash in their reputations) with large ones.

The task of this paper is to determine, as best as is possible with data provided by eBay, how the system is working. A reputation system must meet three challenges (Resnick, Zeckhauser et al. 2000). It must: (1) provide information that allows buyers to distinguish between trustworthy and non-trustworthy sellers (2) encourage sellers to be trustworthy, and (2) discourage participation from those who aren't. In the terminology of asymmetric information, the second and third criteria are that a reputation system must deter moral hazard and adverse selection on the part of sellers (Milgrom and Roberts 1992, chapters 5 and 6).

The eBay reputation system is applied to buyers as well. However, buyers' reputations matter substantially less, since sellers can hold goods until they are paid. The greatest risk is that they will not get paid, in which case they can turn to the second high bidder. Moreover, even if sellers wished to rely on buyers' reputations it would do little good, since it is not possible to exclude buyers with bad reputations from one's auction.[4]

It is worth noting at the outset that the system need not be theoretically sound in order to work. It may only be necessary that both buyers and sellers believe that the system or some part of the system works. There is little published literature on the effective workings of reputation systems on the Internet, so it seems extremely unlikely that many participants are aware of frequency of feedback, disproportions in feedback among those having positive and negative experiences, etc. What matters therefore, is not how the system works, but how its participants believe it works, or even whether they believe it works even if they have no concern about why. To invoke an analogy drawn from

[4] In fact buyers' reputations are slightly better than sellers' reputations overall, as discussed in section 6.

grander considerations, the behavior of man in a world without a God might be fully moral and God fearing if its denizens believed there was a God who would judge them and possibly punish them in the hereafter.

Let us illustrate how a system might deter moral hazard and adverse selection, even if it did not allow buyers to actually distinguish trustworthy from untrustworthy sellers. Suppose that negative feedback is rarely given even when buyers are dissatisfied, suggesting that the system would not work if this fact were widely known. Say that new sellers have to pay an entry or initiation fee in terms of reduced prices and reduced frequency of sale/item listed.[5] (eBay charges a listing fee of between $.25 and $2, so that a lower probability of sale costs sellers a little money as well as time and effort.) If unreliable sellers know that they will have to pay their dues at the outset, and if they believe that the feedback system is likely to give them poor ratings, they will be deterred from participating. They will not make the investment of entering in the first place.

Internet-Based Reputation Systems

Why might we expect an Internet-based reputation system to work? Why might it fail? The big advantage of the Internet is that the out-of-pocket costs of providing and distributing evaluations are respectively minimal and zero.
Consider the contrast with the village retail store, which may have thousands of different customers in a month. Most Internet sellers will have far fewer actual buyers than this, and they will be much more dispersed. Both factors would suggest that reputations, both good and bad, would be harder to establish.

Cheap collection and distribution, however, can accomplish a lot. If one has a bad experience in the village, one may tell one's friends, at a cost of a few minutes for each telling. A bad experience with an Internet seller can be recorded in less than a minute, and spread to millions of potential customers. Each of those potential customers is much less likely to ever encounter the blameworthy seller, but they will have the relevant information on a seller when they need it. In short, if information were reliably provided, Internet auctions would provide much more information about sellers than normal retail operations.

One big question about Internet auctions is whether feedback information will be provided as a function of a variety of seller characteristics, such as the seller's current repuation, experience with the sale, and whether the seller first provides feedback on the buyer. Information about the behavior of others is a public good, and whether the information is bad or good, there is little incentive to provide it. Even keystrokes are costly. In fact, more than half of buyers do provide feedback. Assuming feedback is provided, will it be without bias? The disincentive to provide negative information may be far stronger, with the potential for lawsuits, and for retaliatory negative feedback.[6]

---

[5] Section 7 examines empirically whether sellers do pay dues in this way.
[6] A concern about retaliation for negative feedback is prominent in eBay's discussion forums and eBay has publicly acknowledged the concern, although their only remedy thus far is to exhort users to provide honest feedback despite the risk of retaliation. For example, founder Pierre Omidyar posted a public letter to that effect on June 9, 1998.

Once again, the reality may not matter. Even if there has never been a lawsuit, participants may feel the threat.[7] Beyond these concerns, most people do not like to provide negative feedback. And there may be a bias to the positive side, particularly if the seller goes first, if the feedback process is basically an exchange of courtesies: "You behaved well. You also behaved well."

In the section that follows we discuss the frequency with which feedback is provided as a function of a variety of seller characteristics. But why should anyone consider providing feedback? Why shouldn't they merely free ride? We suspect that many people do it as part of some quasi-civic duty. It is an encouraged activity, and does not cost much. Others do it as a courtesy. They have had a successful transaction and want to say thanks. Some expect reciprocity. Indeed, numerous sellers communicate with buyers that they always provide feedback for a successful transaction, and they hope the buyer will do so as well. These reasons do not apply, or apply much less forcefully, when experience has been bad. Bad evaluations, theory would suggest, are much less likely to be given, unless revenge is a strong motivating force.[8]

Customer-scored reputation systems to date rely overwhelmingly on voluntarily provided information. This creates strong incentives to free ride, and quite possibly to Pollyanna (disproportionately positive) feedback. An alternative framework would pay individuals for providing evaluations, and would reward them if their assessments correlated with future experience. Such markets for evaluations would rely on micropayments from potential buyers to past buyers. In auction markets, current potential buyers could check seller reputations, but only at a cost (Avery, Resnick et al. 1999).

2. The eBay Feedback System

The bulk of this paper is devoted to the assembly and simple analysis of factual information. Given the lack of information that even researchers have about these systems, it seemed that sophisticated game-theoretic analyses of feedback systems, however fascinating, simply could not capture reality. Few if any players could be fully aware of the game that they are playing. Indeed, it is likely that different participants view the feedback system quite differently. The principal goal of our analysis is to discover the empirical properties of these systems that allow them to work so well, not necessarily in terms of providing accurate feedback, but in encouraging participants to buy and sell such a large volume of sometimes nonstandardized items from strangers.

[7] Note the willingness of credit card companies to cover losses when one's card is stolen or used inappropriately. Presumably the expected cost is low. But assuring customers on this score would be far less effective than merely providing coverage and charging a bit more for the card. The advantage is not due to customer risk aversion, but rather that credit card companies could never assure customers that the level of risk is really low.

[8] Personal experience suggests that the threat of providing negative feedback may be employed when seeking to rectify a transaction (a counterfeit watch), but such threats do not always work, and the person issuing the threat (RJZ) does not always follow through.

Before turning to our actual data, it is worth looking at how the eBay reputation system actually works. Before participating in any auctions, as either buyer or seller, people (henceforth called users, as in users of the auction service) register, providing information to eBay, including name and contact information. The only information that eBay verifies, at least for buyers, is that the email address is valid.[9] Not all of the information provided to eBay is made visible to other users. In particular, as part of the registration process, the user chooses an online pseudonym, or ID. This, rather than the full name provided to eBay, is shown to other users when buying and selling. Some users choose to use their email address as their pseudonym but many others choose other short monikers. Other users can request the email address associated with any pseudonym, but must reveal their own email address in order to do so. EBay does not reveal the real names and physical addresses provided during registration to other users. Since there are many ways to sign up for free email accounts at services like HotMail and Yahoo, this system means that anyone who wants to remain anonymous has the option to do so.

Buyers and sellers can leave comments about each other after transactions, but are not required to do so. Each comment consists of one line of text, plus a numeric rating of +1 (positive), 0 (neutral), or -1 (negative). Initially, any user could leave feedback about any other user. Beginning in February, 1999, all negative feedback had to be tied to a particular transaction. Beginning in February, 2000, all feedback had to be tied to a transaction: i.e., only the seller and winning bidder can leave feedback about each other.

Users have the option of whether to make their feedback visible to other users or not, but this decision must be made about the entire feedback profile, not on a case by case basis. By default, the feedback profile is publicly visible, and we have never encountered a user who had chosen to hide their feedback profile.

When a buyer searches for items, she sees a summary listing of items, with item titles and current price, but not the sellers' pseudonyms or any indication of their prior feedback (Figure 1). After clicking through to the item's full description page, the seller's ID and a summary feedback score are displayed (Figure 2). The feedback summary is computed by taking the number of unique (distinct) users who left positive feedback, and subtracting the number of unique users who left negative feedback.[10] If the buyer chooses to click on the summary feedback score, she sees details of the seller's feedback profile (Figure 3). This includes a breakdown of positives, neutrals, and negatives, and also breakdowns for the most recent week, month, and six-month period. The buyer can then scroll down to view the text of the actual comments, most recent comments first (Figure 4). It is possible to scroll through any or all of the comments, but eBay does not provide an easy way to search for the negative and neutral comments. After placing a bid on an item, the seller can view the list of bidders and click through to see their feedback

[9] As part of the registration process, eBay sends an email to the user, which authenticates that the email contact information is a working a email account that the person registering has access to. EBay does not require buyers to provide a credit card number. Beginning October 22, 1999, eBay required sellers to provide a credit card number or go through an alternative identity verification process.

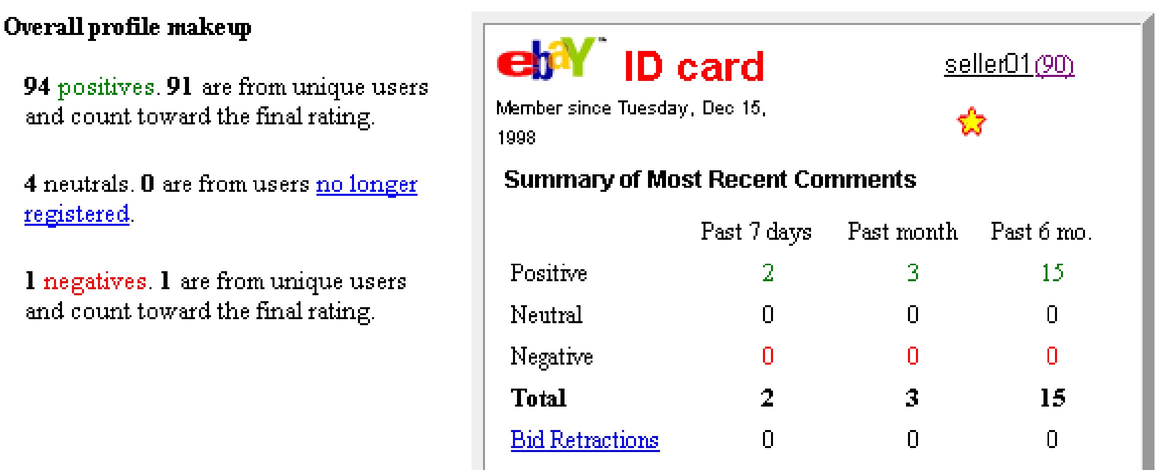
[10] This summary statistic is highly inappropriate in an information theory sense, since negative scores are so rare. A more informative single number would be positives - c*negatives, where c would be a number far greater than 1.

profiles. Sellers are permitted to cancel particular bids, but they do not have a way to proactively prohibit future bids from the same bidder, or all bids from bidders having bad feedback profiles. Since many bids arrive in the last few minutes of auctions (Roth and Ockenfels 2000), it is often impossible for a seller to avoid unwanted buyers.[11]

Item#	Current Items	Price	Bids	Ends PST
1202462934	BRILLIANT Cut Crystal Cruet by HAWKES #2	$65.00	-	Dec-22 19:10
1202254895	2 APRICOT CUT TO CLEAR WINE GOBLETS NR	$75.00	1	Dec-24 10:25
1203377353	2 Sandwich Plates Pinwheels	$9.99	-	Dec-25 22:54
1203323028	Cut Glass 2 Handled 9" Bowl - EXPANDING STAR	$15.00	1	Dec-27 20:36
1203341845	CUT GLASS - 1 1/2" SHALLOW BOWL - OLD	$99.95	-	Dec-27 21:18
1203505924	Cut Glass 11 1/2-in. High, Shouldered Vase	$150.00	-	Dec-28 18:39

Figure 1: the auction summaries listing

Figure 2: detail about one item, including feedback of seller and highest bidder. User names are disguised in this and other figures to protect the privacy of eBay users.



	Past 7 days	Past month	Past 6 mo.
Positive	2	3	15
Neutral	0	0	0
Negative	0	0	0
Total	2	3	15
Bid Retractions	0	0	0

Figure 3. Feedback profile of the seller.

[11] On the online forum that eBay runs, sellers complain about their inability to avoid problem bidders, especially those who do not pay after winning auctions. For example, see the discussion in Appendix B.

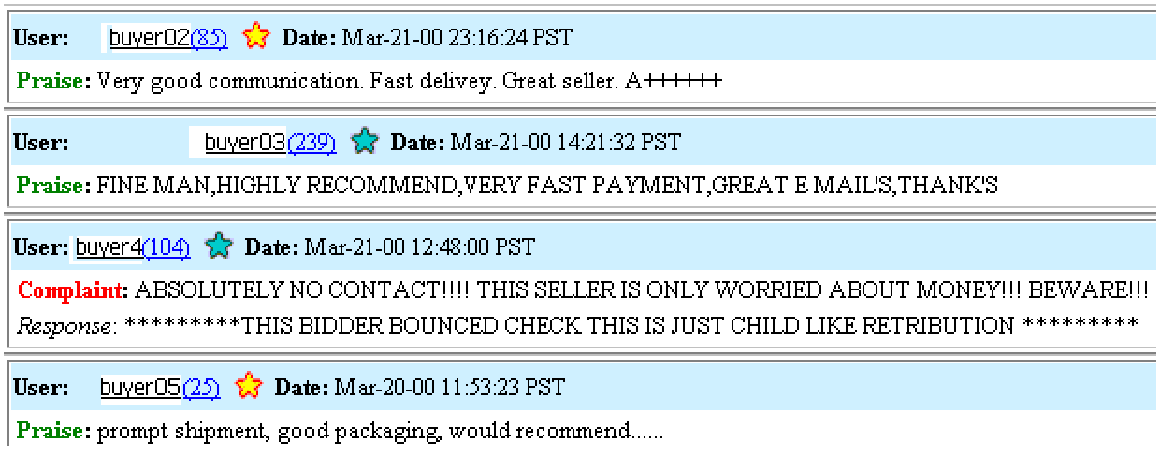


Figure 4: some of the individual comments about the seller. Note that the seller's response to the one negative feedback alleges that it was retaliatory in nature rather than a real indicator of the seller's performance.

A user can change the pseudonym presented to others at any time, but the previous feedback follows the user. It is possible, however, to get a new email address, and register completely anew, thus leaving behind one's previous feedback. Over time, eBay has taken efforts to make this increasingly difficult to return with a new identity (e.g., when registering with a hotmail.com email address, a user has to provide a credit card number and eBay may notice when the same number is used again). It is not, however, completely impossible to do so.

3. Data sets

Our primary data source consists of transactional data from eBay from February 1 - June 30, 1999, and all feedback data up to June 30, 1999. This data set contained more than 20 gigabytes, or the equivalent of more than 13 million pages of double-spaced text. In order to protect the privacy of users, the data set does not include email addresses, pseudonyms, or any other personally identifying information.[12] In order to protect eBay's commercial interests, we are not able to report data that could be used to infer revenue or profits, such as average number of items sold or average selling price. In order to make the data set manageable, we received only the item titles, and not the more extensive text descriptions of items that appeared on the item detail page (e.g., Figure 2). A few other data fields are also not available to us, in order to protect privacy interests of eBay users, or commercial interests of the company. For the most part, however, we have fairly complete data about items listed, bids placed, and feedback given and received.

Several smaller data extractions form the basis for analyses in this paper. The first we refer to as the listed items data set, LI. It is a sample of single items (eBay also permits Dutch auctions of multiple units of the same item) that were open for public bidding (eBay also permits private sales). 36,233 of these items attracted a bid at least as high as

[12] In our data set, users are identified with numeric identifiers, enabling tracking of user activity over time without revealing who the users are. eBay has lookup tables to match the numeric identifiers with the registration information for users, but this information was not provided to us.

both the starting bid and the reserve price and hence were officially sold (some buyers back out after winning auctions, which is the primary source of seller complaints against buyers). All the selected items were listed on February 20, 1999, but this was not a complete sample of items sold on that day. There were 13,695 distinct sellers who listed items for sale in the LI dataset (we refer to them as sellers even if the item did not sell), and 25,103 distinct buyers. This dataset was small enough to permit, for each transaction, computation of buyer and seller feedback profiles prior to the transaction and extraction of feedback about the transaction.

The second data extraction, which we refer to as the longitudinal listings data set, LL, contains all the items listed by a sample of 1000 sellers, during the full time period February 1 - June 30, 1999. The LL data set contains 168,680 item sales. It is too large to permit detailed analyses, but is useful for assessing the extent of repeated interactions between particular buyers and sellers.

A third data extraction focuses on negative feedback in particular, and we refer to it as NF. It consists of a random sample of 1580 negative feedbacks entered on May 1, 1999. For each recipient of this negative feedback, all transactions from February 1- June 30, 1999 were also extracted. This data set is considered in section 8.

Two additional small data extractions reflect transactions for particular items, Rio MP3 players and a particular type of Beanie Baby. These are considered further in section 7.

4. General characteristics

Before looking more specifically at the provision and effects of feedback, we first examine three characteristics of the eBay marketplace. First, are most transactions really between strangers, or do users develop ongoing trading relationships. Second, what is the distribution of prior feedback for buyers and sellers? Third, do most buyers also tend to sell and vice versa, or are the buying and selling roles distinct?

To assess the frequency of ongoing trading relationships, we examined the trading histories of a set of 1,000 sellers, in the LL data set. Overall, these sellers sold 168,680 items during the five-month period from February 1 to June 30, an average of 169 items each. There were 121,564 distinct buyers.

There were 138,458 distinct seller-buyer combinations. This indicates that 17.9% of all sales involved a buyer and seller who had done business with each other before. However, 89.0% of all seller-buyer pairs conducted just one transaction during the time period, and 98.9% conducted no more than four. In the vast majority of cases, multiple transactions between a seller and buyer occurred within a few days of each other (sellers often offer reduced shipping costs to buyers who buy several items that can be shipped together). In most cases, even these multiple transactions are best thought of as a single interaction. Thus, performance in the current transaction will rarely be directly remembered by future buyers who are, after all, different from the current buyer: an explicit reputation system is needed to make current performance visible to other buyers.

Sellers tend to be far more experienced, on average, than buyers. Figure 5 shows histograms of the overall feedback scores (unique positives minus unique negatives) of the two groups for the LI data extraction. The sellers are much more likely to have very high feedback scores. The sellers' median was 33 (the score of the median item was 82, reflecting that high feedback sellers listed more items in the data set). The buyers' median was 8. It's unlikely that many people accumulate scores of 55 (ln[score] of 4) or 148 (ln[score] of 5) selling items from their attics, so eBay must be attracting professional sellers who are deliberately acquiring items from other sources in order to sell them on eBay. On the other hand, about a quarter of items were listed by people with net scores of less than 20, and 18% were listed by sellers with scores of 10 or less. Thus, there is also a large contingent of amateur sellers. Among the buyers, the concentration of low feedback scores is much higher. Overall, then, eBay is not only a c2c (consumer to consumer) marketplace but also a b2c (business to consumer) marketplace.

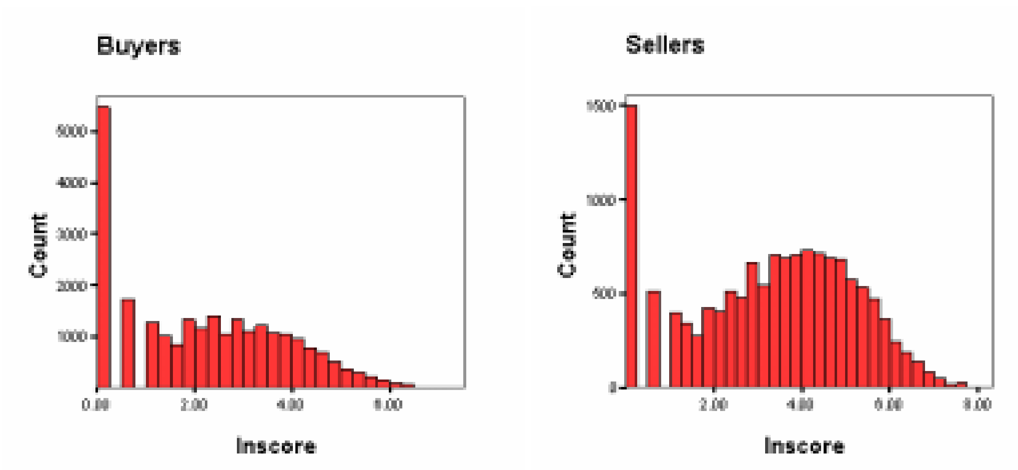


Figure 5. Histograms of feedback profiles for distinct buyers and sellers in the LI data set. Each bar shows the number of buyers and sellers, respectively, with ln (positives-negatives + 1) defining the range of feedback scores grouped in that bar. A ln[score] of 2 corresponds to a score of 7, 4 to 55, and 6 to 403. The sellers group includes all users who listed items, whether or not the items sold.

While some eBay users are traders, most are primarily either sellers or buyers. We traced the selling and buying of all users who listed or bought items in the LI extraction, for the entire period from February 1 through June 30, 1999. For the sellers in the LI extraction, the median number of items sold during the five month period was 76 but the median number of items bought was just 9. For the buyers in the LI extraction, the median number of items sold was 0 (three quarters sold fewer than 10 items) and the median number of items bought was 28.

5. Provision of feedback at eBay

The first and most basic questions about feedback concern the extent to which it is provided, and its content when provided. Clearly, there are some users with hundreds or even thousands of feedback points, so feedback is given at least some of the time. How often is feedback positive? A brief perusal of the ratings on Ebay or Yahoo suggests that negative feedback is rare, but hardly unknown.[13]

The LI data extraction of 36,233 sold items permits more systematic examination of these questions. During this time period, feedback was sometimes but not always explicitly marked as relating to a particular transaction. Using computerized matching, we were, however, able to find, for each transaction, the first feedback from the buyer about the seller that occurred in the six weeks following the transaction.[14] Table 1 summarizes the descriptive statistics for this data set. Buyers commented on sellers for 52.1% of the items, sellers on buyers 60.6% of the time. Thus sellers -- many of whom have quite extensive experience -- will have at least some expectation that there will be feedback on their transactions. Of feedback provided by buyers, 0.6% of comments were negative, 0.3% were neutral, and 99.1% were positive. Sellers were slightly less happy (the most common complaint by sellers is that the winning bidder simply doesn't follow through on the transaction). Taken at face value, the overwhelming majority of transactions lead to happy outcomes. A cursory examination of the follow-up comments after negative feedbacks suggests that even these transactions sometimes end with happy buyers.

	Buyer of Seller		Seller of Buyer	
	Frequency	Percent	Frequency	Percent
negative	111	0.3	353	1.0
neutral	62	0.2	60	0.2
positive	18,569	51.2	21,560	59.5
none	17,491	48.3	14,260	39.4
Total	36,233		36,233	

Table 1: The frequency of feedback after transactions in the LI data extraction.

Analysis of the text comments accompanying the 111 negative and 62 neutral comments revealed many reasons for dissatisfaction, as shown in Table 2. While both neutrals and negatives generally indicate dissatisfaction (except for 8 of the neutrals), they tended to used for different kinds of complaints. Surprisingly, items that did not match their description were somewhat more likely to receive neutral than negative feedback, perhaps reflecting that buyers may have thought discrepancies were honest mistakes on the part of sellers. Similarly, slow shipment was more likely to lead to a neutral than negative feedback. Not following through on a sale, or worse, not sending the item after receiving payment, tended to yield negative rather than neutral feedback.

[13] Informal surveys of economists, all of whom knew about eBay but only a small fraction had purchased from it, suggests that theory alone provides little indication to the answers to these questions. There was considerable variation in answer to the questions of: How frequently is feedback given, and how often is it negative? None of our respondents got close to the actual frequency of negative feedback.
[14] It is possible that such a feedback could be about a different transaction, if the buyer and seller engaged in multiple transactions during the time period. Thus, the data reported here probably slightly overestimate the percentage of transactions that get positive, neutral, and negative feedback.

	Neutral feedback	Negative feedback
Arrived in poor condition, item not as advertised, replica rather than original	24	17
Backed out of transaction (did not contact or respond to high bidder)	1	30
No item received after sending payment	2	38
Other communication problems	17	18
Slow shipping	13	6
Positive about transaction	8	0
All	62	111

Table 2: Buyers' most common complaints. Note that a single comment may be classified in more than one way, if it mentions more than one problem with the transaction.

6. Does Prior Feedback Predict Future Performance?

Before turning to the question of whether eBay users pay attention to feedback in deciding whether to bid for an item, and if so how much, we first consider the information content of a seller's repuation. Does a seller's feedback profile predict future performance? As suggested in the introduction, the feedback system may be useful even if it is not predictive of future performance, so long as people believe it is. There are also game theoretic models in which buyers should punish sellers for prior negative feedback (or punish newcomers for lack of prior positive feedback) even if prior feedback does not indicate future performance. Still, it is interesting to investigate whether prior feedback is predictive.

Table 2 indicated that neutral comments are typically used for slightly problematic transactions (delays, poor communication) while negative comments are used for very problematic transactions (never shipped, broken, counterfeit, etc.). Since both are rare, we have grouped them together as problematic transactions.

Negative feedback appears to be less frequently directed to experienced sellers than to those who are less experienced. Again considering the buyers and sellers involved in the transactions from the LI data extraction, Table 3 shows the percentage of negative feedbacks among users with varying amounts of positive feedback. Among sellers with fewer than 10 positive feedbacks, 2.83% of all feedbacks were neutral or negative. This figure steadily declines as sellers get more experienced, until the group with more than 1000 positives.[15] Removing one outlier – with 180 negatives, 125 neutrals and 3681

[15] There are many possible explanations for the experience/frequency of feedback relationship. For exampe, experience may make someone a better seller. Only sellers with stellar reputations will stay long-term on the system; with a bad record one comes back with a new identity. Buyers may be hesitant to give negative feedback to a long/strong record, perhaps doubting their own judgment.

Group	N (Sellers)	Percent neutral and negative (Sellers)	N (Buyers)	Percent neutral and negative (Buyers)
0-9 positive	4,018	2.83%	13,306	1.99%
10-49 positive	3,932	1.25%	7,366	1.09%
50-199 positive	3,728	0.95%	3,678	0.76%
200-999 positive	1,895	0.79%	738	0.60%
1000+	122	1.18%	15	0.92%
All	13,695	0.93%	25,103	0.83%

Table 3: The percentages of prior problematic feedback for buyers and sellers with different experience levels.

positives -- reduces the ratio of problematic transactions to 1.04%, still a lot higher than the 200-999 group and somewhat higher than the 50-199 group.[16]

The feedback profiles of buyers follow a similar pattern, although they have somewhat less problematic feedback overall than sellers. One possible explanation is that those who were buyers in the transactions in our sample tended to have more of their prior feedback from purchases than did our sellers. Since buyers offer a standard good (money) and offer it first, one might expect buyers to receive less negative feedback overall than sellers.

We next turn to the question of whether prior feedback could be used to help buyers avoid problematic transactions, again using the LI data extraction. The overall probability of a neutral or negative feedback was .46% (.89% of evaluations given). This may be an underestimate of the true percentage of problematic transactions, because problems may be resolved before users contribute feedback and even if the problem is not resolved, some buyers may not leave negative feedback. It could be a substantial underestimate if participants are hesitant to give negative feedback for fear of retaliation.

We model the feedback the buyer provides about the transaction as a function of the seller's previous feedback profile. It is not obvious exactly what features of the feedback profile should be most diagnostic of performance. The analysis above suggests that more experienced sellers are better, up to a point. We compute the log of the number of positive feedbacks, adding 1 to avoid the possibility of taking the log of 0.[17] To measure previous seller problems, we combined the neutral and negative feedbacks as an indicator of problematic transactions, added one and computed the log. A logistic regression the

[16] In future work, we hope to explore the decline in performance once an extremely large number of positives has been received. Initial hypotheses include that it becomes too expensive to leave the system once one has acquired so many evaluations, sellers are counting on many buyers not looking beneath the aggregate score, or that sellers become "worse" because they capitalize or ride on their established reputations.

[17] We also included the square of ln(pos) as a variable, to possibly account for the fact that performance declines with very high feedback, but its coefficient was not significant and was of the same sign as ln(pos), so we have omitted from the results reported here.

following coefficient estimates[18]: The coefficients are all highly significant.

$$\ln\left(\frac{\Pr(problematic)}{1-\Pr(problematic)}\right) = -3.9404 + .7712 * \ln(PROBLEMS+1) - .5137 * \ln(POS+1)$$

Thus, for a newcomer with no previous feedback, this estimates the probability of a problematic transaction at 1.91%. A veteran with 100 positive and no negative feedbacks would have an estimated probability of a problematic transaction of .18%. 100 positives and 3 negatives would yield an estimate of .53%. It is worth noting here that since negative feedback is rare, for experienced buyers the positive feedback score is almost the same as the net score of positives minus negatives. Thus, predicting performance based only the net score that eBay computes would treat the sellers with 100 positives and 0 or 3 negatives as almost the same, while this model suggests that the risk of problematic transactions is quite different for the two profiles.

The model accounts for only a small part of the variability in which transactions are problematic, but it does provide some value, as summarized in Figure 6 and Table 4.[19] For example, someone using this model who was willing to forego bidding on about half

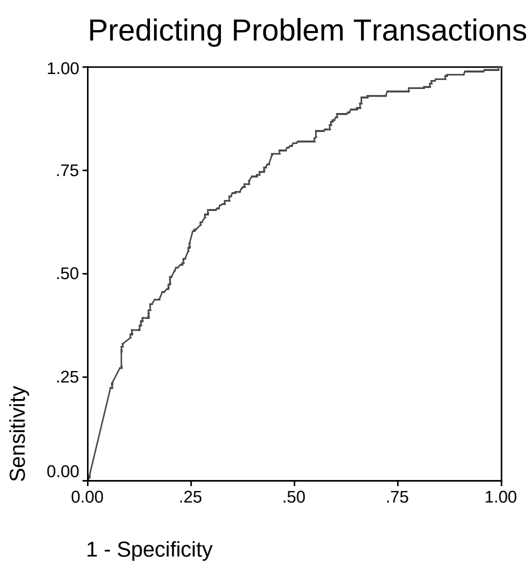


Figure 6: The tradeoff between sensitivity and specificity for predicting problematic

[18] We considered other models, including variants on the fraction of negative feedback, but the coefficient on this variable was never significant. We speculated that the value of the transaction may be correlated with the probability it will be problematic, since sellers can profit more from cheating on a high-priced item and the cost of a negative is independent of the price. But the effect of price was tiny and not statistically significant in models either with or without the prior feedback as covariates.

[19] More sophisticated models might do even better at predicting future performance. For example, Zaccaria has analyzed adaptive scoring functions that weight more recent activity more heavily (Zacharia, Moukas et al. 2000). Dellarocas has examined models that might discount the effects of a few ratings that were entered in a deliberate effort to manipulate reputations (Dellarocas 2000). The results from the simple model in this

of the items could avoid more than four fifths of transactions that were reported as problematic. This would cut the probability of a problematic transaction by almost two-thirds (from .48% to .18%).

1-specificity (% of unproblematic transactions rejected)	Sensitivity (% of problematic transactions rejected)	Cutoff predicted probability	% of accepted transactions that are problematic
75%	94.2%	.20%	.11%
50%	81.5%	.31%	.18%
25%	57.2%	.54%	.27%
10%	32.4%	1.09%	.36%
0%	0%	Accept all	.48%

Table 4: Some particular values for the tradeoff between reducing the danger of problematic transactions and the percentage of unproblematic transactions that would be rejected. The third column gives the cutoff: if the model predicts a probability less than this cutoff, then the transaction would be accepted. Since some of the accepted transactions have predicted probabilities far less that the cutoff, the figures in the right column tend to be somewhat lower.

7. Do Buyers Reward Better Reputations?

Independent of whether reputations are in fact diagnostic of future performance is the question of whether buyers reward sellers who have better reputations. They might choose to do so in a self interested fashion, because they believed reputations were diagnostic of future performance, i.e., they would pay more to get what was likely to be a better quality good. Or, perhaps seeking a warm glow, they could be selflessly policing the system, seeking to provide appropriate incentives to sellers.

The premium accruing to a good reputation should show up in some combination of higher probability that listed items will sell and higher selling prices. The seller may have some choice over how to allocate the premium between these two variables: for example, a seller may use higher minimum bids or reserve prices in order to extract higher prices, at the expense of a lower probability of sale, or vice versa. Thus, simply observing either prices or probability of sale could potentially present a misleading result. Our solution is to examine both: if they both are affected in the same direction we can infer an effect; if they are affective inversely, then we cannot conclude what the effect of reputations are.[20]

In computing the effects of reputations on buyer behavior, it is essential to control for the underlying value of the items being sold. For example, it may be that people with more negative feedback tend to sell more expensive items, which could lead to a spurious

paper merely demonstrate that there is enough information to do predictions and should motivate further development of prediction functions.

[20] Other researchers have dealt with this problem using more sophisticated statistics. For example, Bajari (Bajari and Hortascu 2000) develops structural models of buyers' bidding decisions and Lucking-Reiley et al (Lucking-Reiley, Bryan et al. 2000) compute maximum-likelihood censored normal regressions to account for the fact that bids that would have come in below the minimum bid will not be visible in the auction logs.

result if item type were not controlled for.[21] Two other studies (Bajari and Hortascu 2000; Lucking-Reiley, Bryan et al. 2000) have solved this problem by comparing item prices to book values. In both cases, they examined a set of sales of rare coins for which book values were available. A third study (Houser and Wooders 2000) considered a set of essentially identical items (computer chips). The results of these studies were slightly mixed, but most models reported showed a significant effect of negative feedback reducing sale price, and a trend toward a positive effect of positive feedback on sale price.

We have followed the second strategy of identifying comparable items that should sell for the same price if not for differences in sellers. This turns out to be somewhat tricky, since buyers may notice features of items (for example, by examining pictures) that we miss. Despite the huge dataset available to us, it was surprisingly difficult to identify collections of items that were clearly comparable based only on the item titles (the one-line descriptions shown in summaries like that of Figure 2). For this paper, we have identified two sets of items that seem to be pretty well described just by their item titles. One set consists of 456 listings of Rio MP3 digital audio players. The other consists of 180 listings of Britannia Beanie Babies.[22]

Tables 5 and 6 show the results of our regression analyses. While the effects on price were indeterminate, more positives and fewer negatives and neutrals did appear to affect the probability of sale, and in similar ways for the two groups. Doubling the number of negatives and neutrals seems to more than cancel the effect of doubling the number of positives, which suggests that sellers with a constant ratio of problematic feedback will be treated worse of over time. In future work, we hope to further mine this data set for insight into how buyers interpret and respond to seller feedback profiles. For example, it may be that the effects are different for items in different price ranges, for new or used goods, and for unique or standard goods.

	Price (if sold)	Ln[Pr(sale)/(1-Pr(sale))]
N	378	456
R^2	.004	
Ln (positives)	.544	.501**
Ln (problems)	1.23	-.771**
Constant	$141.93**	.962**

Table 5: regression analyses of effects of reputation on price and probability of sale of Rio MP3 players. The effects on price are indeterminate (neither coefficient approaches significance). The effects on probability of sale are both significant at .01 level.

The logistic regression model predicts that a seller with no prior feedback has a probability of sale of 72%, approximately the same as that for sellers with 2 positives and 1 problematic. 12 positives and no neutrals or negatives is approximately equivalent to 37

[21] Other researchers have in fact documented such spurious results (Lucking-Reiley, Bryan et al. 2000).
[22] Like coins and stamps, these “collectibles” are highly differentiated based on condition, manufacturing location, etc. We considered only those items that in their title indicated mint condition with mint tags, manufactured in Indonesia.

positives with 1 problematic, leading to a probability of sale around 90%. 70 positives and no problems brings the modeled probability of sale up to 96%.

	Price (if sold)	Ln[Pr(sale)/(1-Pr(sale))]
N	151	180
R^2	.015	
Ln (positives)	.176	.559**
Ln (problems)	6.169	-.813*
Constant	$122.45**	-.074**

Table 6: regression analyses of effects of reputation on price and probability of sale of Britannia Beanie Babies. The effects on price are again indeterminate (neither coefficient approaches significance). The effects on probability of sale are both significant (ln(positives) at the .01 level, ln(problems) at the .05 level).

The logistic regression model predicts that a seller with no prior feedback has a probability of sale of 48%. 12 positives and no neutrals or negatives is approximately equivalent to 36 positives with 1 negative, leading to a probability of sale around 80%. 78 positives and no negatives brings the probability of sale up to 91%.

8. Incentive Effects: Do Sellers Pay Attention to Their Feedback Profiles?

A necessary (though not sufficient) condition for feedback to have an incentive effect, encouraging better behavior from sellers, is for them to care about their feedback profiles. Anecdotally, we hear that sellers do indeed care, subtly or not so subtly encouraging buyers to provide positive feedback at the end of successful transactions, and going to great lengths to make buyers happy in order to avoid negative feedback. But perhaps these stories reflect only a small minority of sellers.

To assess this question more systematically, we examine two pieces of empirical evidence, the propensity of sellers to respond to negative feedback, and whether their behavior changes after receiving negative feedback. Starting in February 1999, eBay offered users the opportunity to respond to feedback. When feedback comments are displayed, the text of any response is displayed right below the text of the feedback. (The original issuer of feedback has the further opportunity to respond to the response, but the dialog ends there.) If sellers were concerned about how negative feedback would affect their reputations, we might expect them to provide explanations in response to negative feedback.

Since the opportunity to respond was introduced during the month of February, 1999, we first selected a sample of 1580 negative feedback events from May 1 to generate the NF data extraction, well after the feature was introduced, so that participants could have learned how to play. Of these, the recipient entered explanatory text in 457 cases, 29% of the time.

In future work, we plan to analyze whether sellers tend to alter their behavior in any detectable way after receiving a negative. For example, they might tend to buy a few items or sell lower-priced items in order to cause the negative feedback to no longer appear at the top of their feedback profile, or simply wait a week or a month, so that it no longer appears as recent feedback. We also plan to analyze whether negative feedback is correlated with decisions to stop participating altogether.

9. Equilibrium Considerations

Economists are used to studying interactive situations, setting up the game, and describing the equilibrium or equilibria that result. Strangers buying and selling on the Internet do not have the luxury of having a keen game theorist on hand. Many will not understand the game they are playing. Others would understand, but do not have the information to draw appropriate conclusions. No doubt, different people -- based in part on presuppositions and part on personal experience -- believe the games are quite different. Bearing these qualifications in mind, our empirical evidence tells a great deal about the nature of the game being played. We turn now to a discussion of considerations that likely inform the perceptions of the players and the performance of the feedback market.

High Courtesy Equilibrium

There are two surprising facts about the feedback game on eBay: (1) the high rate of providing evaluations, and (2) the extreme rarity of neutral or negative evaluations. The first suggests that free riding is overcome, the second that buyers are grading generously, or saying nothing after bad experiences.

We think of this as a High-Courtesy-Equilibrium. Manners frequently lead people to make small cost efforts, even when dealing with strangers that one will never again encounter, that promote general welfare and a sense of comity (Martin 1985). Such behaviors, say waving a car ahead of you at an intersection, can be self-reinforcing if people take satisfaction in doing the right thing, or experience discomfort from violating a possible norm. If buyers think it is the courteous or right thing to do to provide feedback, that is what they will do, if it is not too expensive.

The frequency of positive feedback can also be a result of High Courtesy. As children we have been told that: "If you can't say something nice about someone say nothing at all." If it appears that everyone, or near everyone, on eBay is behaving this way, then it is natural that we should as well. Etiquette is the art of getting people to coordinate on social conventions, and the convention of being positive in casual encounters is well known.

eBay itself has done what it could to create an environment where people will be strongly positive. For example, in computing an overall score, eBay merely subtracts the number of negatives from the number of positives, despite the former being much rarer and hence presumably more informative. eBay does not provide a facility for searching for a seller's negative feedback comments, leaving potential buyers with the task of scrolling through all the other feedback to find the rare but informative complaints. Finally, eBay encourages buyers to contact sellers to try to resolve problems and leave negative feedback only as a last resort.

Indeed, eBay may produce additional inducement of positive feedback through its policy of allowing sellers to provide feedback on buyers as well the reverse. If a seller provides positive feedback, it may create a reciprocal obligation to provide positive feedback in return.[23] There may also be a fear of retaliation for negative feedback.[24] Table 7a shows the distribution when both parties provided feedback. As before, we lump negative and neutral into the "problematic" category.

Buyer feedback about seller	**Seller feedback about buyer:** Problematic	Positive	None
Problematic	54	35	84
Positive	17	15,122	3430
None	342	6,403	10,746

Table 7a. Patterns of reciprocation.

There are two striking features about the table. First, there is a noticeable correlation in the propensity to provide feedback at all. Second, there is a very strong correlation of buyer and seller feedback when both provide feedback. The seller is positive 99.8% of the time when the buyer is positive but only 39.3% of the time when the buyer is neutral or negative. Similarly, the buyer is positive 99.7% of the time when the seller is, but only 23.9% of the time when the seller is neutral or negative. This overwhelming correlation is presumably due to two factors: (1) Some transactions just work out poorly -- say a shipment that gets delayed and elicits complaints, and both parties get upset. (2) Even when one side of the transaction has been without blame, there is retaliation.

Why do these features emerge? We might expect sellers to provide more feedback. First, since they engage in many more transactions, and they have more to do for each transaction, we would expect them to be more automated. Second, reputations count much more for them. If there is some degree of reciprocity, providing positive feedback early could be a good seller strategy (although retaining the threat of giving negative feedback would be a reason for sellers to delay giving feedback). Cutting in the opposite direction, buyers have much more to provide feedback about, promptness of shipment,

[23] Robert B. Cialdini explores the norm of reciprocation in his book Influence: Science and Practice (Cialdini 1993, pp. 19-49). At eBay, some sellers stimulate the reciprocity norm by e-mailing their buyers and assuring that they themselves always provide feedback. In other venues, shrewd marketers take advantage of reciprocation norms. For example, charities mail out address labels or holiday seals, hoping that respondents will find a need to reciprocate with a contribution.

[24] The fear of retaliation for negative feedback is mentioned frequently on eBay message boards. The situation is apparently serious enough that eBay founder Pierre Omidyar posted a message on June 9, 1998 exhorting users to give negative feedback when it was warranted, in spite of such fears.

adequacy of wrapping, and whether the product was as described, including sins of omission.

In fact, sellers do provide feedback more often, 60.6% as opposed to 51.7% for buyers. To address the other questions, we provide more disaggregated data, including information about whether the buyer or seller first provided feedback, in Table 7b.

	Seller feedback about buyer							
	Seller first (or only)			*Buyer first (or only)*				*No feedback*
Buyer feedback about seller	neg	neut	plus	neg	neut	plus	none	none
none	300	42	6,403					10,746
neg	20	1	3	18	0	11	58	
neut	2	4	5	5	4	16	26	
plus	0	6	5,091	8	3	10,031	3,430	

Table 7b: Timing of reciprocation.

We might expect that in most cases sellers would provide feedback first, since the buyer's responsibilities -- namely payment -- are completed first, and reputations are much more important to sellers, who therefore might provide early positive feedback to elicit reciprocity. Overall, buyers are the first provider of feedback 37.6% of the time, sellers are first 32.8% of the time, and neither provides feedback 29.7% of the time. Surprisingly, when both provide feedback, buyers actually go first about twice as often as sellers.

Can an early positive feedback by a seller strongly discourage a negative rating of the seller? When a satisfied seller provided feedback first, buyers gave them only 8 problematic (negative or neutral) feedbacks, a rate of 0.07%. By contrast, when the seller was satisfied but waited (either providing no feedback or positive feedback after the buyer evaluated) there were 111 problematic feedbacks, a rate of 0.81%, or 11 times as high. Of course, providing early feedback may simply be a covariate with some other cause of buyer satisfaction (e.g., good communication skill on the part of the seller). Still, the numbers suggest that early positive feedback may discourage negatives from buyers, even if they are dissatisfied.

Similarly, does the threat of retaliatory negative feedback from sellers discourage a negative rating of sellers? Sellers are far more likely to provide negative feedback after a buyer negative or neutral (19.4%) than they are overall (1.2%). Again, however, the evidence is not conclusive, since a botched transaction (e.g., lost in the mail) is likely to leave both parties unsatisfied. In such cases, reciprocal feedback may reflect misplaced blame by both buyer and seller, rather than retaliation.[25]

Skunk at the Garden Party, A Few Bad Apples

[25] One of our students, Ko Kuwabara, is conducting further theoretical and empirical analysis of the patterns of reciprocity and retaliation.

Assume that our High Courtesy equilibrium is in effect. Might not a few disreputable sellers seek to capitalize on the situation -- e.g., negative feedback unlikely -- and begin a process akin to Akerlof's lemons model (Akerlof 1970) that ruins the market for all? Suppose, for example that there were two types of people, goods and bad apples, with two types of sales, satisfactory and unsatisfactory. Suppose good apples produced an unsatisfactory sale 10% of the time due to chance variation. Bad apples are totally strategic. They behave like goods when it is desirable to do so. But they strategically exploit the system and reap profits from unsatisfactory sales, say from excessively generous descriptions of items.

If there is an infinite supply of bad apples, it would seem, and if the vast majority of buyers were willing to forego bad evaluations despite an unsatisfactory sale, then a bad apple could cruise along for a while before being detected. Consider a world filled with goods. With only 0.3% bad evaluations, this suggests that a negative evaluation conditional on a unsatisfactory sale is only about 3% as likely as a positive evaluation after a satisfactory sale (based on our assumption of 10% of sales being unsatisfactory and the observed levels of negative feedback at eBay). Thus, a bad apple could imitate goods half the time, presumably on cheaper items, and build up a reputation of 30 positives on average before receiving his first negative. Understanding this, more and more bad apples will come into the market, bad evaluations will increase in frequency, buyers will become less satisfied with the distribution, and presumably the whole process could spiral downward.

However, we do not observe such a downward spiral. We identify two phenomena that may help maintain the High Courtesy equilibrium while deterring bad apples: paying initiation dues and stoning bad behavior.[26]

Paying Initiation Dues

Many organizations have initiation dues to be paid before one can become a member, and this may well be the case on eBay, even though eBay does not charge an entry fee. For example, a new seller may find it difficult to make a sale, or may get a significantly lower price until his reputation is built. That is, buyers may consider the extent of a seller's record before deciding whether to do business with him, and what price to offer. If this is so, it may be worthwhile for goods to enter the market and pay the initiation dues. Their long-term returns from such dues may be greater than the amount the bad apples can get from exploiting the system for a short period of time. It may also be cheaper for them to engage in good behavior. An honest dealer will find it cheaper to sell lithographs, for example, than will a bad apple whose principal profits come from selling forgeries.

---

[26] Thomas Schelling addresses these possibilities in his Micromotives and Macrobehavior (Schelling 1978). If good sellers can still make a profit despite the suspicion raised by the presence of bads, and if bads hurt bads more than goods, there may be a stable equilibrium with goods and bads together in a stable mix. An increase in the number of bads might also make buyers attend more to details reputations, e.g., the nature of negative feedback, which would help goods.

Initiation dues in effect are a way of creating trust. Such dues are to sales on the Internet the equivalent of what a massive advertising campaign is for a new retail brand, or a fancy new store is to a retailer. They show the seller is willing to invest a lot of money, which suggests that he has a high quality product that he knows will sell well in the future.

In fact, Friedman and Resnick (Friedman and Resnick 2001) found that in environments where players can assume new identities and thus shed previous reputations, some form of initiation dues is inevitable if trust among veteran players is to be maintained. It has to be more valuable for veterans to maintain good behavior than to defect and then start afresh. The dues may be monetary. Or they may come in the form of worse treatment for newcomers, as we found with the lower probability of sale for newcomers.

Stoning Bad Behavior

Initiation dues, by themselves, may not be sufficient to deter bad apples, assuming that they had the ability to get away with their unsatisfactory behavior on a probabilistic basis. It is possible, however, that there is a contagion effect in reaping negative feedbacks. Once one has a black mark by one's name, others may be much more willing, indeed eager, to cast stones. Given the same level of service, they might become far more critical. Thus, the 3% number we mentioned above (P(neg|unsat)/P(pos|sat) might rise to 20%, going still higher after a second or third negative. If one seeks to exploit on a sustained basis, one will almost certainly get eliminated from the market.

If there is stoning, a bad apple might expect that he would have a very limited run of positive-profit exploitation after his introductory era of paying his dues and before getting stoned out of the market.[27]

Future empirical work will examine the frequency of negative feedback after a range of feedback profiles, and the likelihood of being driven quickly from the market. It will also examine differences in the contagion effects for different types of feedback. Some unsatisfactory experiences might be viewed as likely stemming from bad luck, say a delayed shipment. But others might indicate a difference in types, e.g., bad apple behavior of inflating the description of an item.

If the goal is to deter bad apples from playing in this market, it would be desirable if stoning were reserved for behaviors that are deliberate rather than merely careless, and where the perpetrator benefits significantly from the action that led to an unsatisfactory transaction.

[27] In situations where initiation does vary by player type (some types find it more difficult to carry out honest sales in order to establish good reputations) and players can trade names and hence reputations, stoning might be especially important. It could deter bad apples from acquiring and spending down good reputations: reputations would be spent down too quickly. Tadelis analyzes models of name trading, finding that reputations can still convey some information to buyers, but does not explicitly consider the effects of stoning (Tadelis 1999; Tadelis 2000).

10. Conclusion

This analysis describes the game between strangers on Internet auction sites. Significant trust is required to conduct transactions, yet the instruments that normally help sellers convince buyers they are reliable are for the most part not available in Internet auction transactions. Auction sites have developed ingenious feedback systems that enable sellers (and buyers) to build reputations from satisfied customers. It appears that the significant volume of such feedback -- the median seller on eBay has a net feedback score of 33 -- and the ability to spread it to all potential customers, makes up for the lack of traditional feedback mechanisms.[28]

The presumptive challenge to Internet-based feedback systems is to get buyers to provide feedback with reasonably high frequency, and to provide it honestly. Frequency is not a problem, presuming the feedback is unbiased. More than half of transactions receive feedback. However, the 0.3% negative feedback rate on transactions (.6% of those that provided feedback) and 0.3% neutral feedback numbers from eBay, our principal data source, are highly suspicious.

Yet the system appears to work. We consider two explanations: (1) The system may still work, even if it is unreliable or unsound, if its participants think it is working. Thus, if sellers believe poor behavior will elicit negative feedback, and that buyers depend strongly on reputations, then sellers will behave well and bad sellers will be deterred. It is the perception of how the system operates, not the facts, that matters. We suspect that few participants have conducted even cursory versions of the types of analyses conducted here. (2) Even though the system may not work well in the statistical tabulation sense, it may function successfully if it swiftly turns against undesirable sellers, a process we call stoning; and if it imposes costs for a seller to get established, what we label initiation dues.

We also suspect that norms drawn from elsewhere in society help enable Internet reputation systems to work. To illustrate, we mention merely the norm of reciprocity: The frequency of feedback from little-to-gain buyers on auction sites may derive from the positive but low value feedback to them by sellers. They reciprocate a low value favor with the only favor they have available. On the other hand, sellers gave feedback first only about half the time at eBay, despite the fact that they received the buyer's money before the buyer received the goods, suggesting that such seller-initiated reciprocity can not be the only reason that buyers provide feedback.

It is interesting to speculate about whether eBay would be better off with a system where mildly dissatisfied buyers recorded their dissatisfaction more than they do now. It would help buyers to differentiate among sellers, perhaps creating greater faith in the

[28]The typical customer of a traditional retail store, by contrast, has much more personal contact, and more reliable feedback from people close by. That feedback is hard to quantify, however, since much of the feedback provides summary impressions, and a single experience is likely to enter the assessments of many informants.

effectiveness of the feedback system. On the other hand, making dissatisfaction more visible might destroy people's overall faith in eBay as a generally safe marketplace.

Strangers encounter each other regularly on the hyperactive crossroads of the Internet. In some encounters -- e.g., discussing books in a chat room -- trust is natural, since neither party has much to gain by dissembling. With auctions of objects, however, sellers have strong incentives to exaggerate the quality of or misrepresent the authenticity of their items. Yet judging by the volume of transactions, sellers successfully build reputations of trust. This analysis provided some data on how seller reputations are created, and sketched some mechanisms that are turning such reputations into trust.

References

Akerlof, G. A. (1970). “The Market for "Lemons": Quality Uncertainty and the Market Mechanism.” *The Quarterly Journal of Economics* **84**(3): 488-500.

Avery, C., P. Resnick, et al. (1999). “The Market for Evaluations.” *American Economic Review* **89**(3): 564-584.

Bajari, P. and A. Hortascu (2000). Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights From eBay Auctions, Stanford. January 13 http://www.stanford.edu/~bajari/wp/auction/ebay.pdf.

Cialdini, R. B. (1993). Influence: Science and Practice. New York, Harper Collins.

Dellarocas, C. (2000). Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. ACM Conference on Electronic Commerce EC-00, Minneapolis, ACM. http://ccs.mit.edu/dell/ec00reputation.pdf

Friedman, E. and P. Resnick (2001). “The Social Cost of Cheap Pseudonyms.” *Journal of Economics and Management Strategy* **10**(1).

Houser, D. and J. Wooders (2000). Reputation in Auctions: Theory and Evidence from eBay. March .

Lucking-Reiley, D., D. Bryan, et al. (2000). Pennies from eBay: the Determinants of Price in Online Auctions, Vanderbilt**:** Working paper. http://www.vanderbilt.edu/econ/reiley/papers/PenniesFromEBay.pdf.

Martin, J. (1985). Common Courtesy: in which Miss Manners solves the problem that baffled Mr. Jefferson. New York, Atheneum.

Milgrom, P. and J. Roberts (1992). Economics, Organization, and Management. Englewood Cliffs, NJ, Prentice-Hall.

Putnam, R. (2000). Bowling Alone: The Crumbling and Revival of American Community, Simon & Schuster.

Resnick, P., R. Zeckhauser, et al. (2000). “Reputation Systems: Facilitating Trust in Internet Interactions.” *Communications of the ACM* **43**(12).

Roth, A. and A. Ockenfels (2000). Last Minute Bidding and the Rules for Ending Second-Price Auctions: Theory and Evidence from a Natural Experiment on the Internet., Harvard University**:** Working Paper. June 26 http://www.economics.harvard.edu/~aroth/papers/ebay16.pdf.

Schelling, T. C. (1978). Micromotives and Macrobehavior. New York, W W Norton & Company.

Tadelis, S. (1999). "What's in a Name? Reputation as a Tradeable Asset." *The American Economic Review* **89**(3): 548-563.

Tadelis, S. (2000). The Market for Reputations as an Incentive Mechanism, Stanford University**:** Working Paper. June 2000 http://www.stanford.edu/~stadelis/rep_incentives_1.pdf.

Zacharia, G., A. Moukas, et al. (2000). "Collaborative Reputation Mechanisms in Electronic Marketplaces." *Decision Support Systems* **29**(4): 371-388.