Int. J. Electronic Business, Vol. 7, No. 3, 2009

The actual structure of eBay's feedback mechanism and early evidence on the effects of recent changes

Tobias J. Klein*

Department of Econometrics and OR, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands Fax: +31 13 466-3280 E-mail: T.J.Klein@uvt.nl *Corresponding author

Christian Lambertz

Department of Economics, University of Mannheim, 68131 Mannheim, Germany Fax: +49 621 181-1874 E-mail: lambertz@econ.uni-mannheim.de

Giancarlo Spagnolo

Faculty of Economics, Department of Economics and Institutions, University of Rome 'Tor Vergata', Via Columbia 2, 00133 Rome, Italy Fax: +39 06 2020500 E-mail: giancarlo.spagnolo@uniroma2.it

Konrad O. Stahl

Department of Economics, University of Mannheim, 68131 Mannheim, Germany Fax: +49 621 181-1874 E-mail: kos@econ.uni-mannheim.de

Abstract: eBay's feedback mechanism is considered crucial to establishing and maintaining trust on the world's largest trading platform. Yet, there is confusion among users about its exact institutional details, which changed substantially in May 2007. Most importantly, buyers now have the possibility to leave additional, anonymous ratings on sellers on four different criteria. We provide a thorough description of the institutional details of eBay's feedback mechanism, including those changes. Then, we provide first descriptive evidence on the impact of those changes on rating behaviour.

301

Keywords: eBay; reputation mechanism; strategic feedback behaviour; informational content; reciprocity; fear of retaliation; electronic business.

Reference to this paper should be made as follows: Klein, T.J., Lambertz, C., Spagnolo, G. and Stahl, K.O. (2009) 'The actual structure of eBay's feedback mechanism and early evidence on the effects of recent changes', *Int. J. Electronic Business*, Vol. 7, No. 3, pp.301–320.

Biographical notes: Tobias J. Klein is an Assistant Professor at Tilburg University and is also affiliated with CentER, IZA, Netspar, and TILEC. He obtained his PhD from the University of Mannheim. His research interests are industrial organisation, policy evaluation, the economics of aging, and various other topics in applied microeconomics.

Christian Lambertz is a PhD student at the University of Mannheim. He wrote his undergraduate thesis on eBay and also some of his dissertation will consist of research that is related to this paper.

Giancarlo Spagnolo is Professor at the University of Rome 'Tor Vergata'. He obtained his PhD from the Stockholm School of Economics. He is also affiliated with SITE – Stockholm School of Economics, C.E.P.R., and ENCORE. He has published, amongst others, in the *RAND Journal of Economics*, the *Journal of Economic Theory*, the *European Economic Review*, and the *Journal of Money*, *Credit, and Banking*.

Konrad O. Stahl is Professor at the University of Mannheim. He obtained his PhD from the University of California, Berkeley. He is also affiliated with C.E.P.R., CESifo, and ZEW. He has published numerous papers, amongst others in the *Bell Journal of Economics*, the *Journal of Industrial Economics*, the *RAND Journal of Economics*, and the *Journal of Public Economics*.

1 Introduction

EBay is the largest market ever to exist in terms of number of participants. It brings together about 83 million active users. In 2007, the number of listings exceeded 2.3 billion, and eBay's gross merchandise volume amounted to more than 59 billion US dollars.¹ This great success is often attributed to eBay's feedback mechanism, which is argued to foster trust in the platform (see, e.g., Resnick and Zeckhauser, 2002).

Trading on electronic platforms involves a particularly wide room for opportunistic behaviour on both sides of the market. Anonymity and distance allow sellers to cheat on the quality of the good. Likewise, buyers can be dishonest concerning their payment behaviour (Lin et al., 2007).²

The need to enforce agreements and foster trust for exchange amongst strangers is not limited to eBay. North (1990, 1991) argues that in general *institutions* play a major role in reducing uncertainty by establishing a stable structure to human interactions. Institutions consist of both informal constraints, such as traditions, and formal rules that provide the incentive structure in an economy. On eBay, an escrow service, the availability of buyer insurance, and the feedback mechanism mitigate the thread of opportunistic behaviour. The escrow service is relatively expensive and typically only used for purchases of \$500 and more.³ Buyer insurance is provided as part of the

Paypal payment service and is an effective way of buying insurance given that the seller offers this payment option and the buyer is willing to pay the fees which amount to about 3% of the transaction value.⁴ In this paper, we concentrate on the feedback mechanism and investigate its ability to discipline the transaction parties by providing the right incentives.

As long as feedback reflects the trading partners' experience in a transaction, accumulated feedback provides valuable information for potential trading partners. Thereby, the feedback mechanism potentially disciplines the agents so that the aforementioned forms of opportunism are attenuated by the threat that, if an agent misbehaves today, she will receive bad feedback and will therefore be avoided by other traders in the future. Public statements by eBay emphasise this (desired) effect of the feedback mechanism.⁵

On eBay, both the seller and the buyer of an object can rate each other after a transaction. It has been argued that this reciprocity might lead to *strategic* reciprocation and retaliation: Leaving a positive rating could be driven by expectations to receive a positive rating in return. This is consistent with the observation that usually a positive rating is accompanied with an unrealistically favourable text comment such as "Great transaction. A++++++ Seller". At the same time agents dissatisfied with their trading partners could refrain from leaving negative feedback at all, as they anticipate revenge. Both forms of behaviour would bias the reputation index towards more positive evaluations (Fehr and Schmidt, 1999; Resnick and Zeckhauser, 2002).⁶

The informational content of any user's feedback record depends on the incentives to report truthfully rather than strategically, that is, on the design of the feedback mechanism. For example, if feedback is reciprocal and traders are hesitant to report a bad transaction because they fear revenge from their trading partner, then a fixed length of the time period in which feedback can be left would allow for leaving a negative rating at the very last moment of this period. By 'sniping' negative feedback – i.e., leaving it so late that the other party cannot react any more – users would be safe from retaliation.

Retaliating against deserved negative feedback and reciprocating positive ratings may be used to build a reputation of being an imitator who always replies strategically to a positive feedback with a positive one, and to a negative feedback with a negative one. Such behaviour is in principle observable to other users on eBay as the feedback a user gets and the replies she leaves can be inferred from her feedback record. It may be valuable for a seller to be known as an imitator because it encourages her trading partners to give a positive feedback and discourages them from giving a negative one. EBay even sells a service to sellers allowing them to automatically reciprocate positive feedback.⁷ At the same time, if potential trading partners are fully aware of a seller being an imitator, some of them will probably abstain from participating in the auction, knowing that a negative first feedback would always be retaliated.

The fear of retaliatory negative feedback is regularly expressed in discussion boards and newsgroups on eBay, and users indeed seem to be hesitant to leave negative first ratings. Statements similar to the following from a buyer can be found in many forums:⁸

"In the past I've not left any neg[ative] feedback as I'm afraid of revenge feedback that'll paint me as a bad trading partner ... the dodgy seller ends up with getting away with it just to rip someone else off."

The idea to leave negative ratings so late that the other party cannot react is often put forward to solve the dilemma:⁹

"The secret ... is to wait until the 90 day feedback period is nearly up and then zap em w[ith a] negative feedback when they only have a few hours remaining to respond ... That way they can't retaliate ... This only wor[ks] if you are able to hold a grudge for 90 days ..."

Accordingly, McDonald and Slawson (2002) note that "some users attempt to avoid retaliatory negative feedback by 'sniping a negative'". Auctionhawk, a company specialised on offering services around eBay, even developed and advertised a service, for payment, to give feedback in the last minute.¹⁰ One important contribution of this paper is to show that 'feedback sniping' is in fact technically impossible on eBay, as will be explained below.

Reputation is of central importance in electronic markets and the functioning of a feedback mechanism depends on its design. In this paper, rather than focusing on the *effects* of reputation, for example on prices or the probability of selling, we provide detailed information on the precise *design* of the feedback mechanism on eBay, on which there appears to be substantial confusion, and on how it changed in May 2007. Important characteristics are the sequential nature of the feedback, the ending rule for the time feedback can be left, the possibility to mutually withdraw ratings, and the additional, anonymous seller ratings that can now be left by buyers.

Until the end of April 2007, a user's reputation on eBay consisted of all ratings received from his trading partners, buyers or sellers, on past transactions. In May 2007 the system changed. Since then, buyers have the possibility to leave additional ratings on sellers, one to five stars, reflecting their satisfaction with the accuracy of the auction listing, communication, shipping speed and shipping charges. These ratings can be left anonymously and are summarised in additional summary statistics. After describing the change, we provide descriptive evidence that is obtained from recently collected data. The results suggest that the changes are a useful means to increase the likelihood that dissatisfaction is actually communicated to the community.

The literature on strategic aspects of rating behaviour is sparse. Dellarocas (2006) and Mailath and Samuelson (2006) discuss reputation building mechanisms in general. Dellarocas et al. (2006) provide a practical guidance to design them. In the theoretical literature on reputation, it is typically assumed that either the reputation bearer's behaviour is publicly observed by potential future trading partners, or that privately observed behaviour is truthfully communicated from one market participant to another, e.g. through "word of mouth communication". An exception is Lippert and Spagnolo (2006) where incentives to pass on information within a network of relationships are analysed as well. However, there is no empirical research on this because word of mouth communication is rather hard to document.

Electronic markets, especially their publicly observable feedback records, offer economists the important chance to empirically analyse both the way reputation as a collective good is built and how it acts, albeit in a somewhat special (anonymous and public) environment. Resnick and Zeckhauser (2002) were among the first to provide a descriptive analysis of rating behaviour on eBay. They report that, in their data, sellers commented on buyers 60.6% of the time and buyers on sellers 52.1% of the time. Ratings left by sellers and buyers were positive in 99.1% and 98.1% of the cases, respectively. Furthermore, they find a high correlation between first and second

ratings. They interpret this as evidence for reciprocity and retaliation. Other papers acknowledge the possibility of strategic rating behaviour but focus on the effects of reputation on selling prices and the probability of selling a good. The effects of seller reputation on prices and the probability of selling the object are usually found to be negligible or positive. See, for example, Melnik and Alm (2002), Bajari and Hortaçsu (2003), Cabral and Hortaçsu (2006), Livingston and Evans (2004), Lucking-Reiley et al. (2007), and Houser and Wooders (2006). In contrast, Anderson et al. (2007) focus on typical listing strategies employed by occasional vs. large scale sellers and study the effects of these strategies on prices and probability of sale. See also Bajari and Hortaçsu (2004) as well as Resnick et al. (2004) for an overview. Reichling (2004) presents descriptive evidence on rating behaviour and argues that users have an incentive to leave untruthful ratings. Cabral and Hortaçsu (2006) provide a model of reputation building on eBay but maintain the assumption that rating behaviour is truthful. Dellarocas and Wood (2008) estimate a model of rating behaviour on eBay which is based on the assumption that rating behaviour, if users decide to rate, is truthful. This precludes the possibility that retaliative untruthful negative feedbacks are left. Using their model they infer that buyers are satisfied with a transaction in about 80% of the cases whereas sellers are satisfied in about 85% of the cases. These numbers compare to about 99% positive feedbacks that are left on eBay.

The remainder of the paper unfolds as follows. Section 2 gives a detailed analysis of the 'classic' part of the feedback mechanism. Section 3 describes the May 2007 changes to the system and the newly introduced reputation measures. In Section 4 we provide first empirical evidence on the effect of these changes. We conclude in Section 5, suggesting additional simple ways to discourage opportunistic feedback giving, and to improve on agents' incentives to truly express their satisfaction or dissatisfaction on eBay.

2 The 'classic' eBay feedback mechanism

When an auction ends or an item is sold, eBay sends a notification to the seller and the buyer who should then contact each other and arrange payment and shipping. As an intermediary, eBay assumes no responsibility in the process after the auction has ended, and only gets involved if a dispute arises.

As soon as the auction is over (or the item is sold, if not through an auction), both the seller and the buyer can leave feedback on their trading partner regarding this transaction. Neither party is required to leave feedback, but is actively encouraged by eBay to do so.

A feedback is a 'positive', 'neutral', or 'negative' rating accompanied by a textual comment. When a rating is left, it is immediately observable to the counterpart and to the community. The ratings a user receives are used to calculate his 'feedback score'. Generally, a positive rating increases the user's feedback score by one point, a negative rating decreases it by one point, and a neutral rating leaves it unchanged. A special rule applies if a user interacts repeatedly with the same trading partner and receives more than one rating from her: Then, the balance of these ratings is calculated. If the balance is positive, the user's score increases by just one point, if it is negative, the score decreases

by just one point. Thus, even if users rate each other repeatedly, a member cannot affect another member's feedback score by more than one point.

The 'feedback score' is the most prominent indicator of a member's reputation on eBay and shown in parentheses next to his user ID wherever it is displayed on eBay, no matter in what context. When the user acts as a seller – that is, on every item page –, in addition to his feedback score, the percentage of positive feedback amongst all positive and negative ratings she has received is reported. More detailed information on any user, seller or buyer, as well as several summary statistics are available in the 'feedback profile' that eBay provides for every member. This profile includes a record of all feedback that the user has ever received from or given to other members. Members can make their feedback profile 'private'. In that case, other users can see only summary statistics of the feedback record. However, while his feedback profile is private, a user is prohibited from selling on eBay. The ratings in a feedback profile can be sorted so that only ratings received as a seller or ratings received as a buyer are displayed.¹¹

2.1 Last minute feedback?

With the reciprocity of the system, a detail of crucial importance is the 'ending rule' of the period in which trading partners can post their feedback (hereafter 'feedback period'). When retaliation is possible it is important whether this period is of fixed or stochastic length.

We have argued that dissatisfied participants may be (rightly) concerned that if they post a justified negative feedback the trading partner could retaliate with a non-justified negative. This could only be avoided if it were possible to leave a rating in the 'last minute' of a deterministic feedback period. Then, the trading partner would have no time to retaliate.¹² In the opposite case, i.e., without a fixed ending time of the feedback period, the trading partner may retaliate so that many users will find it unattractive to leave a first negative feedback. Consequently, if truthful reporting is a welfare concern, the presence of a 'last minute' is desirable in the context of feedback.

While eBay only *guarantees* that feedback can be left within 90 days after an item is sold, it seems to be a widespread perception that feedback cannot be left after this 90-day period. Examining our data base and again the structure of the eBay platform we found, instead, that *feedback can in fact be left as long as the auction details are available*. Moreover, if a first feedback is left, the party that received the rating has the opportunity to reply for at least *another* 90 days. This is because whenever feedback is left for a particular transaction, the rating is recorded in the feedback profiles of the member that leaves and of the member that receives the rating. From there the rating is linked to the item details for another 90 days – not after the end of the auction, but starting with the time the first feedback has been left. The eBay system is built in a way so that the details will be available for at least another 90 days. As a result, the receiver of a first feedback will always have time to reply to this rating.

In light of the considerations about feedback retaliation and truthful reporting, it is surprising that the end of the feedback period is stochastic on eBay, and in fact automatically extended after the first rating. In fact, the current real ending rule of the feedback period on eBay is similar to the one of Amazon for auctions.¹³ With this structure, the eBay feedback mechanism in principle *discourages truthful negative*

ratings by the potential first feedback giver, by giving the opponent enough time to retaliate.

2.2 Mutual feedback withdrawal

Another important feature of the mechanism in this context is the possibility to mutually withdraw feedback. Once recorded, a rating cannot be unilaterally removed. However, what probably not all users have noticed, feedback *can* be withdrawn if both parties agree to it.¹⁴ Because of this option, the receiver of a first negative feedback has an incentive to strategically reply with a negative feedback in order to induce her trading partner to agree to a mutual withdrawal of ratings. She might consider doing this even if she is in fact happy with the behaviour of her trading partner in the underlying transaction. If both trading partners then agree on a feedback withdrawal, again, no negative feedback would be observed. If an eBay user is not fully aware of this fact, she might overestimate the informational content of feedback records (Jin and Kato, 2007; Resnick et al., 2004).¹⁵

To summarise, key elements of the classic feedback mechanism are the following:

- 1 feedback is immediately observable
- 2 there is always a time window allowing a trading partner to react to a rating as long as she has not entered one yet
- 3 feedback can be withdrawn upon the mutual consent of both trading partners.

Figure 1 contains a state chart representing these key elements. Each circle in the graph represents a state and each arrow a transition from a state into another state or into itself. Such a transition, and if only from a state into itself, happens at every instant in time, i.e., at every moment, we move along some arrow in the state chart. We can describe the feedback game using the state chart: We enter from the left. We are in the state in which nobody has left feedback so that it can still be left by both. In the next instance of time, either no one rates and the feedback period is not over, or only one of the two parties rates, or both rate each other simultaneously, or the feedback period is over. Depending on the actions of the players we transition into another state.

The last (grey shaded) state is always the payoff state. Payoffs are to be understood as expected payoffs from future transactions on eBay, conditional upon the feedback outcome. Depending on the history, either no feedback has been left, or one or two ratings have been left without being withdrawn – the usual case –, or ratings have been left and were withdrawn thereafter.¹⁶

The dashed part of the graph represents the misperception of the existence of a last minute of the feedback period that was discussed above. To be more specific, the misperception is that after 90 days there is a transition into a 'last minute' state in which the trading partner cannot react to a rating. Most importantly, the chart (without the dashed part) shows that once a first feedback is left, in fact, the trading partner *always* has the opportunity to react with a second feedback.





3 May 2007 changes

As already mentioned, the eBay feedback mechanism changed substantially in May 2007. Since then buyers can, in addition to the 'classic' feedback, leave detailed ratings on a seller regarding four different criteria: item as described, communication, shipping time, and shipping and handling charges. A buyer can award the seller one to five stars on each of these aspects. Detailed seller ratings are optional, and it is left to the buyer's discretion whether she wants to rate the buyer in any of these areas.¹⁷ If a buyer wants to leave detailed ratings, she must do so at the same time as she leaves the classic 'overall' feedback with textual comment. Classic feedback can be left without providing detailed ratings, but not the other way around. Concerning their substance, however, overall feedback and detailed ratings need not be connected in any way. For example, the buyer could leave a positive classic feedback, and at the same time give only one star on each of the four criteria. The two systems co-exist to allow for a smooth transition from the old to the new system. This is necessary because users have built up their feedback records over many years and classic feedbacks cannot be converted into detailed ones. However, as more and more detailed ratings will be collected classic feedbacks are likely to become less and less important.

There are two important ways in which the detailed ratings are different from the classic ratings. First, the *granularity* of the type of rating that can be given is different. In the classic system, ratings consist of a positive, neutral or negative overall assessment and every mark that is not positive is usually considered as a negative mark, see, e.g., Resnick and Zeckhauser (2002). The detailed ratings consist of 1-5 stars along four dimensions. This allows users to be more precise in their assessment of their trading partners' performance (Dellarocas, 2005). Moreover, eBay attaches meanings such as 'satisfactory' to the different star ratings, see Table 1 for details.

	1 star	2 stars	3 stars	4 stars	5 stars
How accurate was the item description?	Very inaccurate	Inaccurate	Neither inaccurate nor accurate	Accurate	Very accurate
How satisfied were you with the seller's communication?	Very unsatisfied	Unsatisfied	Neither unsatisfied nor satisfied	Satisfied	Very satisfied
How quickly did the seller ship the item?	Very slowly	Slowly	Neither slowly nor quickly	Quickly	Very quickly
How reasonable were the shipping and handling charges?	Very unreasonable	Unreasonable	Neither unreasonable nor reasonable	Reasonable	Very reasonable

Table 1Meaning of stars

In the empirical analysis we multiply the numbers of stars by 100.

Second, the new detailed seller rating is *anonymous*. For each criterion, only the average of all ratings is given, and a seller must receive at least ten ratings before the average is reported in his feedback profile. eBay stresses that "sellers will not be able to see the detailed seller ratings you've given them",¹⁸ and when actually leaving detailed ratings, buyers are notified again that "sellers will not see your individual ratings" and that

"only the average of all buyer ratings can be seen by the seller". In other words, buyers are assured that, with the new detailed seller ratings, they are safe from retaliation.

As with the classic feedback where a specific buyer can affect a seller's score by only one point, if a buyer leaves detailed ratings on the same seller for different transactions, only the average of her ratings enters the seller's statistic for that criterion. In contrast to the classic feedback score, averages for detailed ratings will be calculated over the preceding 12 months only.¹⁹

At the time eBay introduced detailed seller ratings, it also introduced minor changes to the other part of the feedback system, in particular to the way in which 'classic' ratings are displayed. Before, apart from the auction number, only the rating – 'positive', 'neutral', or 'negative' –, the textual comment, the feedback giver and the date and time when the rating was left were displayed. For at least 90 days after the end of the auction there was a link to the auction details and users had to follow that link to learn more about the transaction. Today, in addition to this information, the item title and the price at which it was sold are also reported in the feedback profile. Potential buyers can thus see at a glance from what kind of item a recent rating stems, and whether it was given for a high value or on a low value transaction.

To summarise, we predict that users are more inclined to report dissatisfaction in the new system. This is for two reasons. First, the change in granularity of the rating scale makes it less likely that the best rating is given, simply because the scale is finer. Second, the incentive to rate truthfully, if unsatisfied, is considerably stronger in the new system because ratings are anonymous.

4 Empirical results

The data we use were collected in several steps. First, in May 2007, we obtained the usernames of, respectively, the last 3,000 users that were selling in the five categories listed in Table 2.²⁰ This is a representative sample of those who sold at least once in those five categories. Then, on the 1st of June, July, August, and September, we collected their feedback overview pages, one per user. This gave us detailed information on user characteristics such as the overall percentage of positive marks received over their life as an eBay user. Most importantly it gave us information on ratings received in the five months prior to the change and in the first four months after the change.²¹ From this raw data we construct percentages for the classic ratings in order to compare them to the average number of stars obtained in the new part of the feedback mechanism.

Table 2 Categories

Home > All categories > Computers and Networking > Laptops, Notebooks	
Home > All categories > Consumer Electronics > Apple iPod, MP3 Players	
Home > All categories > Toys and Hobbies > Model RR, Trains	
Home > All categories > Collectibles > Trading Cards	
Home > All categories > Home and Garden > Food and Wine	

These categories are from the USA eBay system.

Table 3 contains summary statistics. They are reported for our core sample consisting of users for which detailed ratings are available.²² They show that there are both some big players with large feedback records but also smaller ones. The distribution of feedback scores is highly skewed with a mean of 1,998 and a median of 616, see Table 4. About 29% of these sellers are power sellers. Interestingly, in the first four months after the introduction of the new system a seller in our sample received 317 classic ratings and 163 detailed ratings on average. This reflects the fact that detailed ratings can only be received once users act as sellers.

Variable	Obs.	Mean	Std.	Min.	Max.
Duration membership in years	3,704	4.71	2.79	0.27	11.35
Member is power seller	3,704	0.29	0.45	0.00	1.00
Member has 'About me' page	3,704	0.21	0.40	0.00	1.00
Member has store	3,704	0.29	0.45	0.00	1.00
Received ratings 1/12/2006-30/4/2007	3,704	424.52	1,279.44	0.00	29,913.00
Positive	3,704	420.17	1,258.00	0.00	29,371.00
Neutral	3,704	2.46	15.58	0.00	565.00
Negative	3,704	1.89	14.91	0.00	599.00
Neutral or negative	3,704	4.35	29.56	0.00	1,164.00
Percentage ratings 1/12/2006-30/4/2007					
Positive	3,600	0.99	0.02	0.50	1.00
Neutral	3,600	0.00	0.01	0.00	0.33
Negative	3,600	0.00	0.01	0.00	0.50
Neutral or negative	3,600	0.01	0.02	0.00	0.50
Received ratings 1/5/2007-31/8/2007	3,704	317.02	805.75	9.00	16,376.00
Positive	3,704	312.87	790.78	6.00	16,099.00
Neutral	3,704	2.28	10.59	0.00	189.00
Negative	3,704	1.88	9.71	0.00	229.00
Neutral or negative	3,704	4.15	19.57	0.00	390.00
Percentage ratings 1/5/2007-31/8/2007					
Positive	3,704	0.99	0.03	0.43	1.00
Neutral	3,704	0.01	0.01	0.00	0.21
Negative	3,704	0.01	0.02	0.00	0.49
Neutral or negative	3,704	0.01	0.03	0.00	0.57
Detailed ratings 1/5/2007-31/8/2007 (from 100 to 500)					
Item description	3,697	478.57	17.22	240.00	500.00
Communication	3,686	472.92	21.30	260.00	500.00
Shipping time	3,691	467.24	27.07	280.00	500.00
Shipping charges	3,683	458.19	21.84	330.00	500.00

Table 3Summary statistics

 Table 3
 Summary statistics (continued)

Variable		Mean	Std.	Min.	Max.
Mean number of detailed ratings 1/5/2007–31/8/2007	3,704	162.63	492.94	10.00	13,955.50
Item description	3,697	163.60	495.46	10.00	14,016.00
Communication	3,686	163.00	492.53	10.00	13,875.00
Shipping time	3,691	163.26	494.00	10.00	13,990.00
Shipping charges	3,683	163.10	493.32	10.00	13,941.00
Member feedback score	3,704	1,998.01	5,647.45	1.00	148,980.00
Overall percentage positives	3,704	99.48	1.50	48.60	100.00

The observational unit is a user who has offered a good in one of the categories listed in Table 2 and for whom detailed ratings are available. Received ratings are classic ratings which are either positive, neutral or negative. These ratings are not anonymous and are directly observable to the trading partner. Detailed ratings are the ratings that can now be left additionally to the traditional ratings. Here, 1 to 5 stars can be left, respectively. We multiply the average number of stars received, per user, by 100. The number of ratings differs across detailed and traditional ratings because detailed rating are optional and can only be left for sellers. The member feedback score is the number of unique users who left a net positive rating minus the number of unique users who left a net negative rating. The overall percentage negative ratings is calculated from all ratings the user has received.

 Table 4
 Percentiles of the distribution of feedback scores by availability of detailed ratings

	Percentile									
Detailed rating	1	5	10	25	50	75	90	95	99	
No	0	1	3	15	60	189	446.5	705	1,650	
Yes	26	59	99	229	616	1,646	4,199	7,612	24,714	

Our data confirm the well known empirical finding that the overwhelming share of classic ratings is positive. Interestingly, this share is declining once we compare the five months prior to the introduction of the new rating possibilities to the first four months after the system has been put in place. This decline is small but significant with a *p*-value of 0. One interpretation of this is that the recent changes have motivated some users to expose themselves to the risk of retaliation. It is an open question whether this can be reconciled with rational behaviour.

Figure 2 contains box plots for the average number of stars received by the users in our sample. Overall, these ratings are very favourable. In part this could be driven by the fact that detailed ratings are only available for active players, namely those who have received at least ten detailed ratings in their role as a seller over the first four months the new system was in place. As we have started with 3,000 sellers in each of the five categories this is only the case for 3,700 out of 15,000 sellers. Still, it is interesting to see that the typical seller receives many ratings that are not of the best type, 5 stars. This is in contrast to the finding that almost all classic ratings are of the best type, positive, and is well consistent with the arguments that have been made in Sections 2 and 3.



Figure 2 Box plots for detailed ratings

The horizontal axis is the average number of stars for a given user multiplied by 100. The box is bounded on the left by the first quartile and on the right by the third quartile. The observations outside the two whiskers are considered to be outliers and are reported. The observation on the very left is the lowest observed value, respectively.

Table 5 contains, by category and percentage positive ratings received within the same period, sample means of the average number of stars received by the user. It illustrates the dependence between the two measures and shows that even for users who have received less than 95% positive feedbacks the detailed ratings are very good, e.g., on average between 'accurate' and 'very accurate' for the item description.

Table 5Detailed ratings

	Obs.	Item description	Communication	Shipping time	Shipping charges
	3,669	478.624	472.976	467.291	458.237
100% feedbacks positive	1,919	483.715	480.130	475.393	464.914
At most 95% feedbacks positive	185	446.703	426.595	416.054	425.676

This table shows, in the columns, respectively, the condition on the percentage positive feedbacks received, the number of observations and the average rating received across users. The average number of stars is multiplied by 100. For example, 483.715 is the average score for the item description among all users who received only positive ratings. Reported for users for which all four summary statistics are reported.

Finally, Table 6 contains results from a regression of the average star rating on a rich set of user characteristics. It illustrates that the number of stars depends positively on the number of detailed ratings received, the percentage positive feedbacks received, and the overall percentage positive feedbacks. It depends negatively on the feedback score and is smaller for power sellers.

Table 6Regression results

	Item description	Communication	Shipping time	Shipping charges				
Category notebooks	-5.803	-6.069	-2.855	-8.883				
	(7.75)**	(7.21)**	(2.37)*	(8.65)**				
Category mp3 players	-6.691	-5.708	-3.592	-10.939				
	(7.93)**	(6.01)**	(2.65)**	(9.44)**				
Category model rr trains	-3.310	-1.275	-2.140	-0.061				
	(4.92)**	(1.69)	(1.98)*	(0.07)				
Category trading cards	-0.414	-1.559	-4.749	-0.109				
	(0.68)	(2.28)*	(4.85)**	(0.13)				
Log of mean number of	1.287	1.454	2.567	0.071				
detailed ratings	(4.18)**	(4.20)**	(5.18)**	(0.17)				
	-391.187	-509.203	-697.092	-421.991				
Percentage neutral ratings	(18.71)**	(21.69)**	(20.64)**	(14.73)**				
Percentage negative	-40.905	-204.224	-192.535	-4.601				
ratings	(3.60)**	(16.00)**	(10.02)**	(0.30)				
Log feedback score	-0.710	-1.661	-1.970	-0.586				
	(2.44)*	(5.09)**	(4.22)**	(1.47)				
Overall percentage	4.816	4.825	4.738	4.651				
positive feedbacks	(410.77)**	(365.79)**	(251.89)**	(289.25)**				
Duration membership	0.669	0.811	1.292	0.852				
	(6.26)**	(6.75)**	(7.52)**	(5.81)**				
Member is power seller	-0.384	-0.985	-0.996	-2.261				
	(0.62)	(1.43)	(1.01)	(2.68)**				
Member has 'About me'	0.223	1.852	0.842	1.585				
page	(0.37)	(2.72)**	(0.87)	(1.91)				
Member has store	1.637	0.244	-0.883	0.390				
	(2.84)**	(0.38)	(0.95)	(0.49)				
Obs.	3,697	3,686	3,691	3,683				
R-squared	1.00	1.00	1.00	1.00				
Absolute value of t statistics in parentheses								

* Significant at 5%; ** significant at 1%

The columns contain coefficient estimates and corresponding t statistics that were obtained from a regression of the respective dependent variable on the covariates listed above. The dependent variable is the average number of stars received multiplied by 100.

We conclude from these first empirical results that the rating possibilities that were recently added to the feedback mechanism tend to add valuable information. In particular, about 99% of classic ratings are of the best type, positive, but this seems not to be the case for the detailed ratings.²³ In our previous discussion we have related this to the fact that here, the fear of a retaliatory feedback does not play a role as the type of rating cannot be inferred by the seller.

5 Conclusions

EBay's reputation mechanism is an important institution in an environment of impersonal exchange where information about the attributes of the goods traded and the performance of the agents involved is imperfect. Milgrom et al. point out that

"a good reputation can be an effective bond for honest behaviour in a community of traders if members of the community know how others have behaved in the past – even if any particular pair of traders meets only infrequently." (Milgrom et al., 1990)

Whether the system employed by eBay succeeds in this respect depends on its setup. In this paper we have investigated the institutional details of both the classic part of eBay's reputation mechanism and its newly added features. It is well known that the former allows for the immediate observation of feedbacks given by the trading partners. Therefore, the choice of the timing of feedbacks and their type may be guided by strategic considerations. Leaving a truthful negative feedback is a risky endeavour because it may be retaliated untruthfully, and purely for strategic reasons. We have argued that the existence of a deterministic last minute of the feedback period would allow for negative first feedback giving without retaliation. However, closer scrutiny of eBay's reputation mechanism reveals that a trading partner has always enough time to react to a first rating. This implies that a negative first feedback can never be given without the fear of retaliation. One aim of this paper was to emphasise this point.

A direct consequence of this institutional design is that a first feedback giver, if behaving opportunistically towards establishing a good feedback record, may not give a negative feedback even if it were justified. Instead, she might wait for her partner to give the first feedback. If both wait, no feedback is left at all (Dellarocas and Wood, 2008).²⁴ Conversely, leaving a positive feedback might be driven by expectations on *feedback reciprocity* which induces the trading partner to react with a positive mark. For this reason, many studies have argued that classic ratings do not directly reflect the true performance of the participating agents. Importantly, this does not apply to the recently added rating possibilities that allow buyers to unilaterally rate sellers, as they ensure anonymity.²⁵ The difference between the two systems has been at the centre of the empirical analysis that has been presented in Section 4 of this paper. One important finding is that whereas about 99% of the classic ratings are of the best type, positive, this is not the case in the new system.

In addition to the changes that were made in May 2007 there may be scope for further improvements of the classic part of eBay's feedback mechanism using easy-to-implement measures. In particular, our analysis suggests that to reduce concerns for retaliation and to foster expression of deserved dissatisfaction, the revelation of information should be delayed.²⁶ Most importantly, feedback should only be revealed to the trading partners and thus, the public, *if no more feedback can be left.*²⁷ This could alternatively be done after a fixed period, or after both trading partners have given their feedback. A direct consequence could be that less users might be willing to leave a first (positive) rating because they cannot hope any more to thereby induce their trading partner to leave a rating in return. However, this might be a desired effect as, on average, the rating would be closer to the true satisfaction of the user who left it. In addition, incentives to rate (e.g., reductions on future fees) could be provided or ratings could be made (quasi-) mandatory by not allowing users to bid in

another auction or add a new auction listing if they have not rated within a specified time. An important point is that this change alone would potentially make things even worse as long as the mechanism would still allow for the mutual withdrawal of feedback. We therefore advocate to also remove this option. Otherwise, it tends to remain a dominant strategy to always leave negative feedbacks in order to be able to renegotiate after ratings have been revealed.

Finally, it can be argued that the performance of buyers, if asked to pay first, is subject to little uncertainty. Either the full payment arrives in time, which is verifiable by the seller, or it does not. Then, opportunism on the buyer's side does not play a role. To the contrary, sellers can misbehave on a variety of aspects of their performance. Therefore, it may be worthwhile to limit feedback to buyers rating sellers as in Amazon auctions (Dini and Spagnolo, 2006; Dellarocas et al., 2006).²⁸ Such a change of rules is likely to induce less feedback giving. However, by removing potentially very substantive biases, it should increase the informational content of feedback records and therefore lead to an unequivocal improvement in the allocation decisions taken in this increasingly important market.

6 Future research

This paper provides a detailed discussion of the institutional details of eBay's reputation mechanism. We have focussed on the formal incentives that are provided to the trading partners to communicate experiences truthfully to the community. More generally, North (1990) points out that it is the *combination* of such formal incentives and other informal constraints that induce agents' behaviours. The same formal rules will result in different outcomes when applied in different countries which differ in their culture. As a global marketplace, eBay provides the opportunity to study the effects of its (same) mechanism in different environments. An analysis in this direction would certainly be valuable.

Besides, and more closely related to what has been done in this paper, it would be interesting to conduct an analysis similar to ours with 2008–2009 data. One of the empirical findings in this paper is that average detailed ratings for big sellers are very positive, indicating that the reputation mechanism is effective in providing the right incentives to those sellers. It would be interesting to see whether this also holds true for small sellers for which those detailed ratings were not available in the data that we collected as statistics are only displayed once 10 detailed ratings have been received.

Moreover, more detailed data could be used to investigate how many classic ratings a seller has received *as a seller* and compare this to the number of detailed ratings that she can only receive as a seller. By construction, there is always a classic rating for every detailed rating, but not *vice versa*. One could infer from this how willing buyers are to provide detailed ratings. This in turn would indirectly reveal how useful buyers find detailed ratings for evaluating the performance of sellers.

A more direct way to infer this would be to study the effect of detailed ratings on the probability of selling a good and the selling price since this is informative about the way potential buyers interpret those summary measures. For this, data on the auction level have to be collected, which is a somewhat tedious but worthwhile endeavour.

Acknowledgements

We would like to thank Heski Bar-Isaac, Chris Harris, Ali Hortaçsu, Johannes Koenen, Florian Müller, Axel Ockenfels, Klaus Schmidt, Henry Schneider, and Mike Ward for stimulating discussions and comments on related work. Florian Hauber and Johannes Koenen provided outstanding research assistance. We would also like to thank seminar participants at the ESMT in Berlin and the University of Siena as well as conference participants of the EARIE 2008 in Toulouse for their comments. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged. Finally, we would like to thank the editor and two referees for their helpful suggestions.

References

- Anderson, S.T., Friedman, D., Milam, G. and Singh, N. (2007) 'Seller strategies on eBay: Does size matter?', *International Journal of Electronic Business*, Vol. 5, No. 6, pp.643–669.
- Bajari, P. and Hortaçsu, A. (2003) 'The winner's curse, reserve prices, and endogenous entry: empirical insights from eBay auctions', *RAND Journal of Economics*, Vol. 34, No. 2, pp.329–355.
- Bajari, P. and Hortaçsu, A. (2004) 'Economic insights from internet auctions', *Journal of Economic Literature*, Vol. 42, No. 2, pp.457–489.
- Cabral, L.M.B. and Hortaçsu, A. (2006) *The Dynamics of Seller Reputation: Theory and Evidence from eBay*, Mimeograph, New York University, New York, NY.
- Dellarocas, C. (2006) 'Reputation mechanisms', in Hendershott, T. (Ed.): *Handbook on Information Systems and Economics*, Elsevier, Amsterdam, NL.
- Dellarocas, C. (2005) 'Reputation mechanisms design in online trading environments with pure moral hazard', *Information Systems Research*, Vol. 16, No. 2, pp.209–230.
- Dellarocas, C. and Wood, C.A. (2008) 'The sound of silence in online feedback: estimating trading risks in the presence of reporting bias', *Management Science*, Vol. 54, No. 3, pp.460–476.
- Dellarocas, C., Dini, F. and Spagnolo, G. (2006) 'Designing feedback mechanisms for e-procurement platforms', in Dimitri, N., Piga, G. and Spagnolo, G. (Eds.): *Handbook* of Procurement, Chap. 18, Cambridge University Press, Cambridge, UK, pp.446–482.
- Dini, F. and Spagnolo, G. (2006) 'Reputation mechanisms and electronic markets: economic issues and proposals for public procurement', in Thai, K.V., Araujo, A., Carter, R.Y., Callender, G., Drabkin, D., Grimm, R., Ejlskov Jensen, K.R., Lloyd, R.E., McCue, C.P. and Telgen, J. (Eds.): *Challenges in Public Procurement: An International Perspective*, Academic Press, Gent, BE.
- Fehr, E. and Schmidt, K. (1999) 'A theory of fairness, competition and cooperation', *Quarterly Journal of Economics*, Vol. 114, No. 3, pp.817–868.
- Houser, D. and Wooders, J. (2006) 'Reputation in auctions: theory, and evidence from eBay', Journal of Economics and Management Strategy, Vol. 15, No. 2, pp.353–370.
- Jin, G.Z. and Kato, A. (2007) 'Dividing online and offline: a case study', *Review of Economic Studies*, Vol. 74, No. 3, pp.981–1004.
- Lin, Z., Li, D. and Huang, W.W. (2007) 'Traders beware: an examination of the distribution of eBay sellers' online reputation', *International Journal of Electronic Business*, Vol. 5, No. 5, pp.499–517.
- Lippert, S. and Spagnolo, G. (2006) *SSE/EFI Working Paper in Economics and Finance No. 570*, Stockholm School of Economics, Stockholm, SE, available at http://swopec.hhs.se
- Livingston, J.A. and Evans, W.N. (2004) Do Bidders in Internet Auctions Trust Sellers? A Structural Model of Bidder Behaviour on eBay, Working Paper, Bentley College.

- Lucking-Reiley, D., Bryan, D., Prasad, N. and Reeves, D. (2007) 'Pennies from eBay: the determinants of price in online auctions', *Journal of Industrial Economics*, Vol. 55, No. 2, pp.223–233.
- Mailath, G.J. and Samuelson, L. (2006) *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press, Oxford, UK.
- Mcdonald, C.G. and Slawson, V.C. (2002) 'Reputation in an internet auction market', *Economic Inquiry*, Vol. 40, No. 3, pp.633–650.
- Melnik, M.I. and Alm, J. (2002) 'Does a seller's reputation matter? Evidence from eBay auctions', *Journal of Industrial Economics*, Vol. 50, No. 3, pp.337–349.
- Milgrom, P.R., North, D.C. and Weingast, B.R. (1990) 'The role of institutions in the revival of trade: the law merchant, private judges, and the champagne fairs', *Economics and Politics*, Vol. 2, No. 1, pp.1–23.
- North, D.C. (1990) Institutions, Institutional Change and Economic Performance, Cambridge University Press, Cambridge, UK.
- North, D.C. (1991) 'Institutions', Journal of Economic Perspectives, Vol. 5, No. 1, pp.97-112.
- Reichling, F. (2004) *Effects of Reputation Mechanisms on Fraud Prevention in eBay Auctions*, Thesis, Stanford University, Stanford, California.
- Resnick, P. and Zeckhauser, R. (2002) 'Trust among strangers in internet transactions: empirical analysis of eBay's reputation system. The economics of the internet and e-commerce', in Baye, M.R. (Ed.): *Advances in Applied Microeconomics*, Elsevier Science, Amsterdam, Vol. 11, pp.127–157.
- Resnick, P., Zeckhauser, R., Swanson, J. and Lockwood, K. (2004) *The Value of Reputation* on eBay: A Controlled Experiment, Working Paper, Harvard Kennedy School of Business.
- Roth, A.E. and Ockenfels, A. (2002) 'Last-minute bidding and the rules for ending second-price auctions: evidence from eBay and Amazon on the internet', *American Economic Review*, Vol. 92, No. 4, pp.1093–1103.

Notes

¹See http://investor.ebay.com/common/download/download.cfm?companyid=ebay&fileid=192977 &filekey=08eaa22c-c31c-41ad-b4d3-0a954354566f&filename=ar2007.pdf (January 2009).

²According to the Internet Crime Complaint Center (IC3) 2005 Internet Fraud Crime Report "internet auction fraud was by far the most reported offence, comprising 62.7% of [97,076] referred complaints". See http://www.ic3.gov/media/annualreports.aspx (September 2006). Likewise, the FTC reports that "internet auction fraud is on the rise, with an increasing number of consumers complaining about sellers who deliver their advertised goods late or not at all, or deliver something far less valuable than promised". See the FTC's "Top Ten Dot Cons" on http://www.ftc.gov/bcp/conline/edcams/dotcon/auction.htm (February 2006). Lin *et al.* (2007) find that on eBay, most complaints from buyers did in fact not refer to online fraud. Moreover, they find that buyers could be responsible for online disputes more often than sellers.

³The service is provided by a company that collects the money from the buyer and the good from the seller and only then sends the money to the seller and the good to the buyer. See http://pages.ebay.com/help/pay/escrow.html (January 2009).

⁴Typically, if the seller accepts payments via Paypal, the buyer has to pay the extra fees. See http://pages.ebay.com/help/buy/paypal-buyer-protection.html (January 2009) for Paypal's buyer protection.

⁵eBay states that the feedback "comments and ratings are valuable indicators of your reputation as a buyer or seller on eBay", see http://pages.ebay.com/help/feedback/questions/feedback.html (February 2006). Moreover, in the founder's letter posted on February 26, 1996, Pierre Omidyar claims that "some people are dishonest. Or deceptive ... But here, those people can't hide. We'll drive them away". See http://pages.ebay.com/services/forum/feedback-foundersnote.html (February 2006).

⁶This tendency to reciprocate may be due to behavioural components in agents' decision making processes, similar to the ones found by Fehr and Schmidt (1999), the attempt to build up a reputation as a 'reciprocator' or 'impersonator' in order to discourage future negative ratings and encourage positive ones – "the high courtesy equilibrium" of Resnick and Zeckhauser (2002) –, or the combination of both motives.

⁷The price for an online seller tool which includes this service is currently \$15.99 a month, see http://pages.ebay.com/sell/automation.html (February 2006) for a description.

⁸Quote taken from http://ideas.4brad.com/archives/000018.html (February 2006). See the newsgroup discussions on http://www.the-gas-station.com/messages.cfm?type=normal &thread_id =49933&lastdays=2000& and http://community.auctionsniper.com/groupee/forums/a/tpc/f/ 785608021/m/308108399/r/3721016131, for example (February 2006).

⁹Quote from http://www.the-gas-station.com/messages.cfm?type=normal&thread_id=49933 &lastdays=2000& (February 2006).

¹⁰See http://auctionbytes.com/cab/abn/y04/m08/i10/s01 (February 2006). A free reminder service for "last minute feedback" is offered by UK Auction Watch at http://www.ukauctionhelp. co.uk/remindme.php (February 2006).

¹¹On the German site, for example, users have the additional option to display only ratings from the past month, or the past 6 or 12 months. For these periods, users can then choose to display feedback by type, i.e., only positive, only neutral, or only negative – or only withdrawn – ratings. This feature is not available on USA eBay; it is on other sites.

¹²From a strategic viewpoint, this closely resembles "last minute bidding" in English auctions with fixed ending time (Roth and Ockenfels, 2002), with a "last minute action" being exploited in both cases in order to prevent opponents from reacting.

¹³There, a 'last minute' bid prolongs the auction period automatically. See http://www. amazon.com/gp/help/customer/display.html?ie=UTF8&nodeId=1161360 (September 2006).

¹⁴eBay states that

"[a]fter both parties have agreed to withdraw the feedback, both parties will have their feedback scores adjusted at the same time ... eBay will add a note to the feedback comment, saying that the feedback was mutually withdrawn ... If you haven't left feedback for your trading partner and you go through the Mutual Feedback Withdrawal process, you will no longer be able to leave feedback for that transaction ... You may only request Mutual Feedback Withdrawal once for every feedback left ... Members may initiate a request to mutually withdraw feedback within 30 days of either person leaving feedback or within 90 days of the transaction end date, whichever is later."

Taken from http://pages.ebay.com/help/feedback/questions/mutual-withdrawal.html (September 2006).

- ¹⁵Jin and Kato (2007) find in a field experiment that "at least some buyers" overestimate the informational content of feedback score and "drastically underestimate the risk of trading online". Likewise, Resnick *et al.* (2004) question whether price premia, which they find, reflect a reputation equilibrium, and should in fact not be observed in the data.
- ¹⁶For ease of the exposition, we make the simplifying assumption that we can always enter the feedback withdrawal process after at least one feedback has been left. In reality, every player may initiate this only once, see footnote 14 for details. Note that only a subset of the users on eBay is likely to be aware of the possibility of withdrawal.
- ¹⁷Buyers cannot rate on 'shipping time' and 'shipping and handling charges' in motor vehicle categories.
- ¹⁸http://pages.ebay.com/help/feedback/detailed-seller-ratings.html (September 2007).
- ¹⁹Note that there are (rare) situations in which the rating can be inferred. This is because the average rating is presented accurate to a tenth of a star. For example, if a user has an average

rating of 5 stars from ten ratings and receives an 11th rating he can infer it if it changes his average rating, i.e., if it is at most 4 stars. However, in typical situations this will not be possible as the typical rating is between 3 and 5 stars and the average will be calculated from substantially more than ten ratings. Moreover, for this to be possible the user needs to watch in real time how his average rating changes in order to associate the rating with a transaction through the classic rating that is left at the same time and then shown explicitly in the feedback record.

- ²⁰For this we downloaded the listings of auctions that ended in the past. This can be done using the 'advanced search' option on eBay. 59 users have listed an item in two categories and one user in three categories. These multiple listings were treated as multiple observations in our data set.
- ²¹In particular, we use the information on the number of ratings, by type, that were received within the last month, respectively. They are added over the four waves. For the ratings that were received in the last five months prior to the change we use the number of ratings received in the "past six months" that was published on June 1 and subtract the number of ratings received in the 'last month'.
- ²²Table 4 shows percentiles of the distribution of feedback scores by availability of summary statistics for detailed ratings. It reflects that these are only available if at least ten such ratings have been left, i.e., for the more active sellers with higher feedback scores.
- ²³Figure 2 shows that the third quartile of the average rating per user is well below 5 stars.
- ²⁴Dellarocas and Wood (2008) investigate the consequences of non-random missing feedbacks on feedback scores. They find that dissatisfied traders are more likely not to give feedback.
- ²⁵An additional advantage of these features is that separate reputation measures with respect to four key criteria are calculated so that more detailed information is communicated to the community.
- ²⁶One might object that such a delay is costly because in the meantime an untrustworthy user may deceive more eBay users. However, this would also happen if unsatisfied trading partners would abstain from leaving a negative rating because they fear retaliation.
- ²⁷That is, both whether a feedback was left and the type of feedback should both be concealed to opponents. This is somewhat different from the suggestion of concealing only the type of feedback left which has been made in independent work by Reichling (2004).
- ²⁸This is also suitable for e-procurement platforms. See Dini and Spagnolo (2006) and Dellarocas *et al.* (2006) for further details.