THE DYNAMICS OF SELLER REPUTATION: EVIDENCE FROM EBAY*

Luís Cabral†‡
Ali Hortaçsu

We construct a panel of eBay seller histories and examine the importance of eBay's reputation mechanism. We find that, when a seller first receives negative feedback, his weekly sales rate drops from a positive 5% to a negative 8%; subsequent negative feedback ratings arrive 25% more rapidly than the first one and don't have nearly as much impact as the first one. We also find that a seller is more likely to exit the lower his reputation is; and that, just before exiting, sellers receive more negative feedback than their lifetime average.

Keywords: quality, reputation, auctions, eBay.
JEL Code Nos.: D44, L15, L86.

* We thank Kenny Ballendir, Tim Miller and Jeremy Shapiro for truly outstanding research assistance. We also thank Damien DeWalque, Anna Ingster, Svetlozar Nestorov, Mike Riordan, Anne Rogers, Steve Tadelis, the Editor, two referees, and seminar participants at Chicago, Essex, Copenhagen, SED meetings, NBER Summer Institute, Georgetown, Tufts, Rochester, NYU, Columbia, Dartmouth (Tuck), 2004 ASSET meeting, Arizona, 2005 EARIE meeting, Colorado and Harvard for helpful comments and suggestions. Hortaçsu acknowledges financial support from the National Science Foundation (Grants SES-0242031 and 0449625) and the Alfred P. Sloan Foundation. The usual disclaimer applies.

† Authors' affiliations: New York University and CEPR; University of Chicago and NBER.

‡ Corresponding author: Department of Economics, Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012, U.S.A.
email: lcabral@stern.nyu.edu

I. INTRODUCTION

Electronic commerce presents the theoretical and the empirical economist with a number of interesting research questions. Traditional markets rely significantly on the trust created by repeated interaction and personal relationships. Electronic markets, by contrast, tend to be rather more anonymous. Can the same level of trust and efficiency be obtained in these markets?

One possible solution, exemplified by eBay auctions, is to create reputation mechanisms that allow traders to identify and monitor each other. In this paper, we focus on the workings of the eBay reputation mechanism. We present some empirical evidence regarding the dynamics of eBay seller reputations; and discuss possible interpretations of these results.

Our focus on eBay's reputation mechanism is justified for two reasons. First, electronic commerce in general and eBay in particular are a significant economic phenomenon: in 2004, more than $34.1bn were transacted on eBay by more than one hundred million users.[1] Second, with its well defined rules and available information, eBay presents the researcher with a fairly controlled environment for theory testing. Specifically, a reasonable assumption on eBay is that the information one trader has about other traders is the same as the researcher's. Essentially, this information consists of a series of positive and negative feedback comments given by past trading partners. In this context, we can make sharper predictions about agent behavior than in other markets, in particular in markets where buyers and sellers share information that is not observed by the researcher.

A number of authors have conducted empirical studies of eBay's reputation mechanism. Almost all of these prior studies focus on the buyer's response to published feedback aggregates. In particular, a large number of studies estimate cross-sectional regressions of sale prices on seller feedback characteristics: Dewan and Hsu [2004], Eaton [2005], Ederington and Dewally [2003], Houser and Wooders [2005], Kalyanam and McIntyre [2001], Livingston [2005], Lucking-Reiley, Bryan, Prasad and Reeves [2006], McDonald and Slawson [2002], Melnik and Alm [2002], Resnick and Zeckhauser [2002].[2] Resnick, Zeckhauser, Swanson and Lockwood [2006] point out the potential for a significant omitted variable bias in these cross-sectional regressions, and conduct a controlled field experiment in which a seasoned seller sells identical postcards using his real name and an assumed name. They find an 8% premium to having 2000 positive feedbacks and 1 negative over a feedback profile with 10 positive comments and no negatives. Ba and Pavlou [2002] conduct a laboratory experiment in which subjects are asked to declare their valuations for experimenter generated profiles, and find a positive response to better profiles. Jin and Kato [2005] assess whether the reputation mechanism is able to combat fraud by purchasing ungraded baseball cards with seller-reported grades, and having them evaluated by the official grading agency. They report that while having a better seller reputation is a positive indicator of honesty, reputation premia or discounts in the market do not fully compensate for expected losses due to seller dishonesty.

We start our empirical investigation by estimating a cross-section regression of the impact of reputation on price. We find that a 1% level increase in the fraction of negative feedback is correlated with a 7.5% decrease in price. However, we find the estimates have a relatively low level of statistical significance. These results are comparable with previous research, both in terms of coefficient size and in terms of statistical significance.

Our next step is to go beyond cross-section regression and estimate the effects of reputation based on panel data. To do so, we assume that: (a) the frequency of buyer feedback is a good proxy for the frequency of actual transactions; (b) the nature of the feedback is a good proxy for the degree of buyer satisfaction. We provide statistical tests that suggest the likelihood of feedback is uncorrelated with a variety of seller characteristics, thus giving credence to our strategy of using feedback histories as proxies for transactions histories. We are thus able to construct a data panel of seller histories. These seller histories allow us to look not only at how buyers react to changes in reputation but also at how sellers potentially 'game' the system.

We find that, when a seller first receives negative feedback, his weekly sales growth rate drops from a positive 5% to a negative 8%. Moreover, subsequent negative feedback ratings arrive 25% more rapidly than the first one and don't have nearly as much impact as the first one. We also find that a seller is more likely to exit the lower his reputation is; and that, just before exiting, sellers receive more negative feedback than their lifetime average.[3] In sum, our data clearly suggests that reputation matters: buyers react to information about seller reputation; and sellers' actions, too, are influenced by reputation considerations.

Our main contribution to the study of online reputation mechanisms is twofold: First, we analyze panel data in addition to cross-section data. We believe that the difference between panel and cross-section data is important. In fact, consistently with previous literature, our cross-section results show weak statistical significance. By contrast, our results from panel data are typically much more significant, both economically and statistically. We thus agree with Resnick, Zeckhauser, Swanson and Lockwood's [2003] conjecture that there is significant unobservable seller heterogeneity.

Second, we analyze the impact of seller reputation on buyer *and* seller behavior. Our paper is one of the first empirical papers to directly address how reputation considerations influence both buyers and sellers' actions. In addition to the above mentioned Jin and Kato [2005], other papers addressing similar issues are Hubbard [2002]; Abbring, Chiappori, and Pinquet [2003]; and Jin and Leslie [2008].[4]

The paper is structured as follows. In Section II, we briefly describe the institutional setup of eBay, in particular the mechanics of its reputation mechanism. In Section III, we describe our dataset. The main empirical results are presented in Section IV. Section V concludes the paper.

II. THE EBAY REPUTATION MECHANISM

Since its launch in 1995, eBay has become the dominant online auction site, with millions of items changing hands every day.[5] eBay does not deliver goods: it acts purely as an intermediary through which sellers can post auctions and buyers bid. eBay obtains its revenue from seller fees, based on a complex schedule that include fees for starting an auction and fees on successfully completed auctions.[6] Most importantly, to enable reputation mechanisms to regulate trade, eBay uses an innovative feedback system.[7] After an auction is completed, both the buyer and the seller can give the other party a grade of $+1$ (positive), 0 (neutral), or -1 (negative), along with any textual comments.[8]

eBay displays several aggregates of the grades received by each seller and buyer, including (a) the difference between the number of positive and negative feedback ratings, (b), the percentage of positive feedback ratings, (c) the date when the seller registered with eBay, and (d) a summary of the most recent feedback received by the seller.[9] Finally, eBay provides a *complete record* of the comments received by each seller, starting with the most recent ones.

All of the information regarding each seller is publicly available. Hence, as claimed in the introduction, this is an environment where the economic analyst has the same information that a new buyer has about a seller.[10]

III. DATA DESCRIPTION

Our data was collected from eBay's website at monthly intervals between October 24, 2002 and March 16, 2003.[11] We focused our attention on auctions of (arguably) ex-ante homogenous goods to minimize the impact of object-level heterogeneity, but we also wanted to capture possible sources of variation across objects with different characteristics. Hence we collected transaction level information on the following objects:

1. IBM Thinkpad T23 PIII notebook computers ('Thinkpad' in the tables below). We chose this category because, according to the FBI's online fraud investigation unit, most customer complaints regarding online auction fraud arise from laptop auctions. We further chose this object because, while notebook computers tend to come in many different configurations (regarding memory, disk space, peripherals, screen size), this particular IBM model seemed to have relatively minor differences in configuration compared to other manufacturers. The average sale price of the Thinkpad T23's in our data set was $580.

2. Collectible coins. We chose this category because the collectible coin market is one of the most active segments on eBay and several previous studies of eBay auctions have looked at this market.[12] We selected two different kinds of coins: the 1/10 oz. 5 dollar gold coin of 2002 vintage (gold American Eagle; 'Eagle' in the tables below); and the 2001 silver proof set (ten coins of different denominations; 'Silver' in the tables below), both produced by the U.S. mint.[13] The average sale prices in our data set are $50 for the gold coin and $78 for the proof set.

3. 1998 Holiday Teddy Beanie Babies ('Teddy' in the tables below), produced by the Ty toy company. Beanie babies are a popular collectors' item on eBay, and according to the FBI's Internet Fraud unit comprise the second largest source of fraud complaints on online auctions. This is the least expensive item in our data set, with an average sale price of $10.7.

Along with transaction-level data, we collected data from each seller's feedback page, thus recording the seller's entire sequence of reviews. We should note that transaction-level data (price, object description, number of bidders, etc) is only available during 30 days. Therefore, while we had access to that data during the six-month period of data collection, our historical record for each seller only includes the feedback comments. Moreover, it is quite possible that the sellers we classify as beanie-baby sellers (because they sold beanie babies during the data collection period) actually sold different objects in the past.

A key assumption in our analysis is that the likelihood of buyer feedback is approximately constant (at least within object category). We discuss evidence supporting this assumption at the end of Section IV(i). Accordingly, we take the number of feedback comments as a proxy for the number of past sales and refer to a large seller as one with many feedback comments.

■ **Seller characteristics.** Table I provides some summary statistics on seller size. The average seller in our sample had 1625 total feedback responses. The median seller had 397. The largest seller has 52,298 feedback responses, the smallest 0 (i.e., the seller is yet to be rated, even though at least one sale took place). We found the distribution of seller sizes (proxied by number of feedback points they got) to be approximately lognormal. Sellers were largest in the market for Thinkpads, followed by teddies, gold coins and the proof sets.

[Place Table I approximately here.]

While the mean and median seller in our sample is quite large (in terms of transactions conducted), the number of negative comments is rather small. As can be seen from column (2) of Table I, the average seller in our sample has 4.9 negative feedback points, corresponding to 0.9% of all comments. The maximum number of negative feedbacks received by a seller is 819, but this seller took part in 52,298 transactions. Also notice that the median seller in our sample has only one negative; more than a quarter of the sellers have no negative comments.

One issue regarding the interpretation of comments is whether neutral comments are closer to positives or to negatives. Our subjective impression, after browsing through eBay community chatboards where users discuss issues regarding the feedback system, is that the information contained in a neutral rating is perceived by users to be much closer to negative feedback than positive.[14] Indeed, observe that in Table I the distributions of neutrals and negatives across sellers are extremely similar. The average seller received 7.2 neutral comments in her lifetime, with a median of 1 (as in the case of negative feedback). Given this evidence, we will henceforth lump negative and neutral comments together when referring to 'negative' feedback.

IV. EMPIRICAL RESULTS

In this section, we present our main empirical findings. They are divided into four subsections. In Section IV(i), we present the results from our cross-section regressions of price on reputation measures. The remaining subsections are based on our data panel. In Sections IV(ii) and IV(iii) we study the impact of the first few negative feedback ratings in a seller's history: impact on growth (Section IV(ii)) and impact on the frequency of negative feedback (Section IV(iii)). Finally, Section IV(iv) looks at seller exit: who is more likely to exit and what pattern of feedback do we observe near exit time.

IV(i). *Reputation and Price*

At the most basic level, we would expect a better seller reputation to influence the price paid for an otherwise identical object. To investigate this hypothesis, several papers in the prior empirical literature on eBay have ran regressions of the form:[15]

$$\text{price} = \beta\ (\text{reputation measure}) + \gamma\ (\text{other demand factors}) + \epsilon.$$

Since we have data for a series of auctions across four homogeneous product categories, we follow the literature by running similar cross-sectional regressions.

Table II reports our results from such cross-sectional regressions. In the first four regressions, the dependent variable is the log of the highest bid registered in the auction.[16] Hence the coefficient estimates can be interpreted (approximately) as percentage changes in price. The regression in column (1) allows for heteroskedasticity across object classes and controls for object dummies. The coefficient on the percentage of negatives in a seller's feedback history is negative and implies that a one point increase in this percentage (at the mean value, from 1% to 2%) leads to a 7.5% decline in sale price. The coefficient on the total number of transaction reviews (divided by 1000) received by the seller is positive (but not significant at conventional levels), and implies that a 1000 increase in the number of reviews is associated with a 5% increase in sale price.

[Place Table II approximately here.]

Observe that the magnitude of this estimate is close to the findings of several other cross-sectional studies. In particular, the 5% price premium implied by 1000 additional reviews is comparable to the 8% premium found by the field experiment of Resnick et al. [2003], which compared sales prices obtained by a seller ID with 2000 positive comments (and 1 negative), and a seller with about 15 positive comments (and zero negatives).

However, as first pointed out by Resnick et al. [2003], several unobservable confounding factors may render a 'causal' interpretation of the reputation measure difficult. For example, sellers with better reputation measures may also be much better at providing accurate and clear descriptions of the items they are selling; hence their writing ability, and not their reputation, may be underlying the higher prices they are receiving.

The next set of results reported in Table II enable us to get a feel for the importance of such confounding factors in cross-sectional price regressions. In column (2), we adjust the standard errors by allowing for correlation in the error term within a seller. This adjustment leads to the coefficient on the percentage of negatives being no longer statistically significant (though the coefficient on total number of reviews becomes significant). Column (3) provides even more clear evidence that unobservable factors may be at work. In this regression, we include a dummy variable for the auctions run by *hdoutlet*, the dominant seller (with close to 50% market share) in the Thinkpad market. This leads to the economic and statistical significance of the percentage of negatives and the length of the transaction record to disappear entirely, implying that the comparison of auctions of this seller vis-a-vis other, much smaller sellers, drives much of the findings in column (1).

The results in column (2) and column (3) suggest that factors other than differences across sellers transaction histories may affect the cross-sectional variation in prices; and it may be difficult for an econometrician to account for these factors since the econometrician is typically not a very knowledgeable buyer in these markets. In fact, a few of the other coefficient estimates in Table II also suggest that factors other than reputation scores play a larger role in the cross-sectional variation of prices. For example, the presence of the word 'refurbished,' or whether the seller allowed payment by a credit card, are both correlated with large variations in price.

In summary, the results in the first three columns of Table II suggest, at best, a rather weak cross-sectional correlation between sale price and the reputation measures that eBay publishes.

One way to strengthen the case for a causal connection between cross-sectional variation in reputation and sale price is to exploit an exogenous change in reputation measures which is not correlated with the way sellers prepare their listings. We exploit the following exogenous change in eBay's website format: before March 1st, 2003, bidders would only see the seller's overall (net positive) feedback points next to the seller's name. On March 1st, 2003, eBay began to display the percentage of positive comments received by the seller, as well as the date when the seller registered on eBay.[17]

In column (4) of Table II, we find that the interaction of the percentage of negatives with a dummy variable for the format change implies that the response of prices became more negative after the format change.[18] According to the regression results, the economic effect of a 1% increase in negative feedback was a 5% change in price before the format change (but insignificant), and a -10% change after the format change. This suggests that bidders did not utilize the 'percentage' information (presumably due to information acquisition and processing costs) before the format change, but began to utilize it after the information became freely available. As one might expect, the coefficient estimate on the total number of seller feedbacks remains unchanged when interacted with the format change, since eBay displayed this information before and after the format change.[19]

In Table II, we present results from two additional regressions. Column (5) is a linear probability regression of a completed sale indicator, and Column (6) is a regression with log of (number of bidders+1) as dependent variable. The percentage of the negative feedback variable has the expected negative sign but is only marginally statistically significant (on the completion probability) and not significant (on the number of bids). The sign of the variable 'total number of feedbacks' is actually the opposite of what we would expect (and is statistically significant). Moreover, unlike sale price, the change in display format seems to have had no effect on the extent to which negative feedback or the number of feedback comments affects the completion probability or the number of bids. Finally, the remaining covariates have statistically significant effects on both the completion probability and the number of bids.

Overall, the results from Table II suggest that it is difficult to get clear results using cross sectional data.

■ **Is feedback an exogenous process?** The remainder of our empirical analysis in this section will be founded on the assumption that the frequency of feedback is a good proxy for the frequency of transactions. One way to test this assumption is to uncover the determinants of feedback giving, namely whether there are systematic patterns related to the seller's type.

Specifically, to test whether feedback-giving is an exogenous event, we took the transactions we used in our price regressions, and matched them with our feedback data. We found that 40.7% of these transactions resulted in a feedback,[20] with 3 negatives and 3 neutrals (i.e., 1.4% of feedbacks were non-positive). We then ran a regression of the binary outcome of receiving a feedback on seller characteristics, along with dummies for object types.

Our results show that seller characteristics such as seller's total transactions, percentage of negative feedback and percentage of negative feedback in the most recent 6 month period do not have a statistically significant correlation with feedback reception.[21] There are differences across object categories in the frequency of feedback reception (the Eagle and Silver coins are more likely to receive feedback than Teddys and Thinkpads), but this most likely reflects different social norms across categories.

In sum, we are fairly confident that frequency of feedback provides a good proxy for frequency of transactions. In the next subsections, we will use this device to create a panel data of transactions. The idea is that at any moment in time when seller i makes a trade we have access to all of his or her feedback history, and by approximation to all of his or her transactions history. One disadvantage of this approach is that we lose all price data regarding past transactions: all we have regarding past transactions is feedback (if it was given), nothing else. The main advantage is that we may correct for seller specific effects and obtain stronger correlations. The panel data approach also allows us to study the lifetime patterns of sellers, in particular, when and why they exit.

IV(ii). *Negative Feedback and Sales*

We now use our panel data on sellers' feedback records. We begin in this subsection by examining the impact of negative feedback on the seller's sales rate. Our typical seller receives his first negative during the early stages of his career. During this period, sales rates are typically increasing over time. To account for the possibility of growth rates varying with age, we first regress each seller's weekly growth rates on the seller's 'age,' as measured by total transactions completed on eBay until that time.[22] We also use as regressors age squared and indicators for different object categories. In what follows, when we refer to weekly growth rates, we mean age-detrended weakly growth rates, that is, the residual from the above regression.[23]

We then averaged the weekly sales growth rates over a four week window before and after the week in which the seller got his first, second, third, fourth and fifth negative feedback;[24] and conducted paired t-tests of the null hypothesis of equality of growth rates before and after the negative feedback event.[25]

The results, reported in Table III, are striking: For the Thinkpad, for example, the impact of the first negative feedback comment is to slow growth by 13% a week, from a positive growth rate of about 5% to a negative growth rate of about -8%. The values for the other products are of similar magnitude. Moreover, we find these differences are highly statistically significant. By contrast, the difference in growth rates before an after the second negative feedback is positive. However, except for Eagle, the difference is not statistically significant. The impact of the third negative feedback also does not appear to be statistically significant.[26]

[Place Table III approximately here.]

Several notes are in order. First, our exercise depends importantly on the assumption that the probability of feedback is the same before and after negative feedback is received.[27] Second, our strategy for collecting seller histories retrospectively may imply a sample bias (we only have data for surviving sellers). In particular, there may be sellers who exited after receiving the first negative feedback and are thus excluded from our sample. But intuition suggests that, if anything, this reinforces the point that the first negative feedback has a negative impact on sales.

More importantly, one possible objection is that of endogeneity. For example, it may be that expectations of future sale declines result in less service by the seller, which in turn may lead to worse feedback. Alternatively, there can be changes in seller quality over time that affect both variables (feedback and sales) simultaneously.

In order to address the possibility of correlation without causality, we considered the natural experiment of the effect of 'mistaken' feedback comments. In a small percentage of cases, buyers mistakenly give a negative rating when their comment is clearly positive.[28] We repeated our analysis of the impact of negative feedback for this subsample. We considered both first and second negative feedbacks.[29]

The results are reported in Table IV. Due to the small size of our 'mistakes' sample, we pool together all four object categories.[30] Although the exact values are different from the full sample, we still have very large differences, significant both economically and statistically.

[Place Table IV approximately here.]

In summary, there is significant evidence that the first negative feedback has a strong negative impact on the seller's growth rate; and that subsequent negative feedback comments have lower or no impact on the sales rate.[31]

IV(iii). *Frequency of Negative Feedback*

Our second result relates to the frequency of arrival of negative feedback. We measure 'time' in number of sales transactions. As mentioned above, negative comments often came in the context of a 'war of words' between seller and buyer. To prevent such incidents from biasing our results, we excluded consecutive negative comments by the same buyer. We also excluded any negative comments that were left within a two-day period after another negative.[32] Finally, we excluded those negative/neutral comments that were received as a 'buyer.'[33]

Table V displays three magnitudes of a seller's record, all measured in number of transactions: T1, 'time' to the first negative; T2, 'time' between the first and the second negative; and ET, the estimated interval between negatives if they are uniformly distributed across a seller's history. Under the null hypothesis that negative feedback is generated by a stationary process, we would expect all three to be equal.

[Place Table V approximately here.]

The results suggest that both T1 and T2 are greater than ET, and moreover T1 is greater than T2. These differences are not uniformly significant. While the difference T1−ET is significant at the 5% level for every product, the difference T1−T2 is not significant for Sliver; and the difference T2−ET is not significant for Eagle or Teddy. Notice however that, for the most expensive item, Thinkpad, all three differences are significant.

The differences are also economically significant. For example, it takes an average Thinkpad seller 93 sales before the first negative is received; but it only takes an additional 58 sales (38% less) before the second negative arrives.

One potential problem with the results in Table V is the possibility of sample selection bias. Specifically, we can think of two possible biases. First, there may be sellers who were born before we started collecting data and who have exited after an early negative feedback. By excluding these, we may overestimate the value of T1. Second, by excluding sellers with one negative only we may also be biasing our estimate of T2. In order to estimate the potential for bias from our sampling strategy, we performed a series of additional calculations, shown in Table VI. First, we redid the calculations from Table V by restricting the sample to sellers born after October 24, 2002, the date at which we started sampling from eBay. We get different values of T1, T2, but the difference between T1 and T2 remains significant. This can be seen in the first panel of Table VI. In particular, when we pool all object categories, T1>T2 at the 3% significance level. For individual products, we get no significant difference for the Thinkpad and the Silver coins; but the number of observations at this level is rather small.

[Place Table VI approximately here.]

The problem of excluding sellers with one negative only is particularly troubling if they got their negative early on during their lives. Then clearly T2>T1 for those sellers, and their exclusion would bias our test of T1>T2 against the null T1=T2. In our sample of sellers born after October 24, 2002 who received negative feedback, 8 out of 28 had one negative feedback comment only. Their average T1 is equal to 410; the average number of transactions after the first negative is 171. These numbers suggest that the exclusion of one-negative sellers does not imply any significant upward bias in our evaluation of the difference T1−T2. In the second panel of Table VI, we repeat the calculation in the first panel by including all sellers with some negative feedback. For the sellers with one negative comment only we assume T2=T1, consistently with our null hypothesis. The overall results still suggest that T2>T1 (at the 3.5% level, when pooling all objects).

In sum, the empirical evidence suggests that T1>T2: it takes fewer transactions to get the second negative than it takes to get the first one. This result is intriguing. One is naturally led to ask if the change in negative feedback frequency results from a change in seller behavior or simply from a change in the buyers' propensity to give negative feedback. We therefore next consider a series of results to test the hypothesis of buyer behavior.

Suppose that buyers have a threshold of dissatisfaction above which they give negative feedback. Suppose moreover that this threshold drops after the first negative. There are several behavioral mechanisms through which this can happen, and we consider these in turn.

One way in which such a 'threshold decline' may occur is through a decrease in the cost of writing a negative comment. As we noted above, many negative comments are followed by a 'war of words' between buyer and seller. Seller retaliation might impose an economic cost on the complaining buyer, especially if the buyer is also a seller. Such an effect would confound our results if the probability of retaliation by a seller in reaction to her first negative is higher than retaliation to her second negative, an explanation proposed by several eBay users we talked to.[34]

To investigate this possibility, we first checked, for every negative or neutral comment-giver in our sample, whether their particular negative comment was accompanied by a retaliatory negative left by the seller. The result was striking: of the almost 10,000 negative/neutral instances in our data, 2462 resulted in a retaliatory comment by the seller. It is also interesting to note that sellers were less likely to retaliate against neutral comments, as opposed to negatives: we found that a buyer leaving a negative comment has a 40% chance of being hit back, while a buyer leaving a neutral comment only has a 10% chance of being retaliated upon by the seller.

However, our data indicates that sellers are *not* more likely to retaliate upon their first negative, as opposed to subsequent negatives. In Table VII, we regress an indicator for retaliation by the seller following a particular negative/neutral comment on dummy variables for the second through sixth occurrence of such a comment. As displayed in columns (1) and (2), the dummy variables do not enter significantly — the seller is not more likely to retaliate against the first negative comment, as opposed to subsequent negatives. Interestingly, in the first regression, we find that sellers with higher ex-post percentage of negatives are more likely to retaliate (the regression coefficient can be interpreted as saying that a seller with 1% higher percentage of negatives is 4% more likely to retaliate). However, it does not appear that 'fear of retaliation' is a significant driver of the difference in inter-arrival times of negative comments.

[Place Table VII approximately here.]

A second variation on the 'threshold' story is that, in addition to time variation, there is also buyer variation in propensity to give negative feedback. In particular, one can imagine that first negatives are primarily given by negative-prone buyers, whereas subsequent negatives originate in a wider set of buyers. To test this possibility, we looked at the string of feedbacks that were left by every negative/neutral comment giver in our data set.[35] We then computed the percentage of negative comments that each of these reviewers left about others, a measure of each reviewer's 'critical attitude.' In Table VII, columns (3) and (4), we regress the critical attitude of the reviewer leaving a particular negative/neutral comment on dummy variables for the second through sixth occurrence of a negative/neutral. The regression result tells us that buyers who left the first negative were not systematically more 'critical' than buyers who left subsequent negative feedback.[36]

To conclude our test of the 'threshold' story, we directly tested the hypothesis that second negatives have a lower threshold than first negatives. We constructed a series of pairs of first and second negative comments. We then asked a third party (a student) to make a subjective evaluation as to which of the two remarks was more negative.[37] The results show that 51% of the second negatives were considered 'nastier' than the corresponding first negative, a split that is not statistically different from 50/50.

Finally, we consider the possibility that buyers are influenced by other buyers' behavior (herding, conformism, etc).[38] Faced with poor performance by a seller with a perfect record, a buyer might be inclined to think that there is no ground for a negative feedback. For example, if there is a communication problem between buyer and seller, the former may attribute this to a problem with him or herself, not with the seller. However, if the seller has already received negative feedback, especially regarding the same problem that the buyer is now facing, then the buyer may have a greater inclination to attribute this to a problem with the seller and give negative feedback. This is especially true for aspects of the transaction that are more subjective and difficult to input (e.g., communication problems).

To consider this possibility we classified the first and second negative remarks according to their nature. The breakdown of the reasons for negative feedback is presented in Table VIII. The buyer influence story should imply an increase in the relative importance of 'subjective' problems in second negatives. However, the results suggest a very similar pattern for first and second negative (correlation greater than 0.92). Moreover, 'item never sent,' arguably the most objective reason for negative feedback, actually increases in relative importance (though by a small amount). At the opposite extreme, 'bad communication,' arguably the most subjective reason for negative feedback, also increases in importance (though by an even smaller amount).

[Place Table VIII approximately here.]

In sum, the empirical evidence does not suggest any change in buyer feedback behavior following the first negative. Accordingly, we argue the relevant change is in seller behavior. Specifically, we believe the data provides evidence of moral hazard on the seller side: upon receiving the first negative feedback, sellers put less effort into providing good sales service.

IV(iv). *Reputation and Exit*

In the previous section, we focused on an important dimension of the seller's strategy: the effort put into each transaction. Another important dimension of the seller's strategy is exit, both in terms of changing one's identity or in terms of leaving eBay altogether. In this section, we analyze seller exit behavior. To do so, we supplemented our data set by revisiting our sample of sellers in the first week of January, 2004, and checking whether they were still in business. There was considerable attrition in our sample: of the 819 sellers originally sampled in our sweep of the transaction-level data, we found that 152 had not conducted any transactions within the last 45 days.[39] We also could not locate the feedback records for 104 sellers in our sample, since eBay's database claimed that these seller ID's were no longer valid. These two events (not conducting any recent transactions, and not having a valid eBay ID) constitute our definition of 'exit.'

We then ran logit regressions of an 'exit' outcome on seller's observable reputational statistics as of May 2003 (at the end of our initial sampling period). As explanatory variables, we consider (a) the log number of negatives and neutrals and (b) the log number of positives.

The regression results are reported in the upper-panel of Table IX, both for the pooled sample of all sellers and by object category. The signs of the reputational variables appear to conform with intuition — sellers with fewer negatives (more positives) are more (less) likely to exit, though the statistical significance of the number of positives is higher. To get a sense of the economic significance of the results, in Figure 1 we plot the predicted exit probability as a function of the log total number of positives that a seller had in May 2003. As can be seen, a variation from 1 to 4 log points in the number of positives is associated with a decline in exit probability from 60% to about 20%, implying an economically significant correlation.

[Place Table IX approximately here.]

[Place Figure 1 approximately here.]

Next, we investigate whether the 'exits' we see in our data set are accompanied by a concentration of negatives just before exit, a situation we refer to as 'opportunistic' exit. Note that there are two very different interpretaions of what a late accumulation of negatives means. One is that, anticipating exit, a seller decides to take advantage of his reputation. An alternative interpretation is that, following a (possibly exogenously caused) string of negatives, a seller decides to exit. We make no attempt to distinguish between these alternatives, and so the term 'opportunistic' must be understood with the above caveat.

We looked at the last 25 sale transactions conducted by exiting sellers, and counted the number of negative comments for these last 25 sale transactions. Some of the examples were quite striking: one of the sellers in our sample, who had 22755 positives, racked up 11 negatives in his/her last 25 transactions; whereas he/she had a total of 54 negatives in his/her previous transactions (in other words, the percentage of negatives and neutrals over his/her overall history was 0.6%, versus 44% in the last 25 transactions). On average, the percentage of negatives in the last 25 comments of exiting sellers (excluding those who remained as buyers and those sellers whose ID's became invalid, and thus we could not get data) was 4.38%, as opposed to an average 1.61% over their entire histories. This difference is statistically significant at the 1% level.

To see if reputational statistics as of May 2003 have any predictive power over such 'opportunistic' exits, we repeated the logit regressions in Table IX, where we now defined the dependent variable to be 1 if the percentage of negatives within the last 25 transactions of a seller was more than twice the percentage of negatives during the seller's entire history. The results of these regressions are reported in the bottom panel of Table IX. Notice that although the number of positives that a seller has is still negatively correlated with the probability of exit, the number of negatives enters into this regression much more significantly.

Once again, to assess the economic significance of the results, we plot the predicted probability of 'opportunistic' exit, but this time using the log number of negatives as the independent variable. Figure 2 shows that an increase from 1 log point of May 2003 negatives to 2 log points is associated with a 10% increase in opportunistic exit probability, once again pointing to an economically significant relationship.

[Place Figure 2 approximately here.]

V. CONCLUDING REMARKS

We may briefly summarize our empirical findings as follows. We observe a positive growth rate of sales until the first negative feedback is received; the growth rate of sales drops substantially, and indeed becomes negative after the first negative feedback is received. After the first negative feedback we also observe an increase in the rate of negative feedback. Finally, a typical eBay seller is more likely to exit the worse his record is; and the last few trades are likely to include more negative feedback than an average trade during his lifetime.

What do these facts have to say about the economic theory of reputation? Over the past twenty-five years or so, a number of economic theories of reputation have been developed. For all their variety, these theories can be classified into two main frameworks.[40] One, pioneered by the work of Klein and Leffler [1981] and Shapiro [1983], sees reputation as a coordination, or bootstrap, equilibrium in a repeated game context. Here, buyers play an active role in 'punishing' sellers when it is perceived that the latter have not lived up to expectations. A second framework, pioneered by the work of Kreps, Milgrom, Roberts and Wilson [1982], models reputation as a Bayesian updating process: based on the observation of past transactions, sellers form a belief about the type of seller they interact with.[41]

In general, it is difficult to distinguish between the 'bootstrap' and 'Bayesian' reputation models. Note in particular that both reputation mechanisms are consistent with a positive correlation between reputation and incentives to invest in reputation, as our data suggests. Specifically, with a perfect record, reputation is high *and* the incentives to invest on reputation are high. Once the first negative arrives, reputation drops significantly, and so do the incentives to invest on reputation. Since, by assumption, the probability of a positive or negative transaction is a function of effort, the implication is that the likelihood of negative feedback is much lower before the first negative is received than after, as our empirical evidence suggests.

Clearly, we need more detailed data before we can unequivocally select a particular model. However, considering the nature of the eBay market, and based on our own experience of interacting and talking to eBay traders, we believe a Kreps-Milgrom-Roberts-Wilson type of model, combining adverse selection and moral hazard on the seller's side and Bayesian updating on the buyer's side, explains the data best. In an appendix posted on the authors' websites, we provide a more detailed description of the relevant theory, as well as a model that we believe does a good job at explaining our empirical evidence.

Regardless of which theoretical model best explains the data, an important conclusion of our paper is that the eBay reputation system gives way to noticeable strategic responses from both buyers and sellers. Obviously, this does not imply that its current structure is optimal. In fact, we believe an exciting area for future research is precisely the design of an efficient reputation mechanism.

REFERENCES

Abbring, J.; Chiappori, P.-A. and Pinquet, J., 2003, 'Moral Hazard and Dynamic Insurance Data', *Journal of the European Economic Association*, 1, pp. 767–820.

Asch, S., 1946, 'Forming impressions of personality', *Journal of Social and Abnormal Psychology*, 41, pp. 258–290.

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

Bajari, P. and Hortaçsu, A., 2003, 'Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay', *Rand Journal of Economics*, 34, pp. 329–356.

Bajari, P. and Hortaçsu, A., 2004, 'Economic Insight from Internet Auctions', *Journal of Economic Literature*, 42, pp. 457–486.

Cabral, L., 2005, 'The Economics of Reputation: A Primer', New York University.

Cabral, L. and Hortaçsu, Ali, 2006, 'The Dynamics of Seller Reputation: Theory and Evidence from eBay', NBER Working Paper No. 10363.

Cohen, A., 2002, *The Perfect Store: Inside eBay* ((Little Brown and Company, Boston)).

Dellarocas, C., 2003, 'The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms', *Management Science*, 49, pp. 1407–1424.

Dewan, S. and Hsu, V., 2004, 'Adverse Selection in Electronic Markets: Evidence from Online Stamp Auctions', *The Journal of Industrial Economics*, 52, pp. 497–516.

Eaton, D., 2005, 'Reputation Effects in Online Auction Markets', Murray State University.

Ederington, L. and Dewally, M., 2003, 'A Comparison of Reputation, Certification, Warranties, and Information Disclosure as Remedies for Information Asymmetries: Lessons from the On-Line Comic Book Market', mimeo.

Hoch, S. and Ha, Y., 1986, 'Consumer Learning : Advertising and the Ambiguity of Product Experience', *Journal of Consumer Research*, 13, pp. 221–233.

Houser, D. and Wooders, J., 2005, 'Reputation in Auctions: Theory and Evidence from eBay', *Journal of Economics and Management Strategy*, 15, pp. 353–369.

Hubbard, T., 2002, 'How Do Consumers Motivate Experts? Reputational Incentives in an Auto Repair Market', *Journal of Law and Economics*, 45, pp. 437–468.

Jin, G. and Kato, A., 2005, 'Price, Quality and Reputation: Evidence From An Online Field Experiment', forthcoming in *Rand Journal of Economics*.

Jin, G. and Leslie, P., 2008, 'Reputation Incentives for Restaurant Hygiene', forthcoming in *American Economic Journal (Microeconomics)*.

Kalyanam, K. and McIntyre, S., 2001, 'Returns to Reputation in Online Auction Markets', mimeo.

Klein, B. and Leffler, K., 1981, 'The Role of Market Forces in Assuring Contractual Performance', *Journal of Political Economy*, 89, pp. 615–641.

Kreps, D.; Milgrom, P.; Roberts, J. and Wilson, R., 1982, 'Rational Cooperation in the Finitely Repeated Prisoners' Dilemma', *Journal of Economic Theory*, 27, pp. 245–252.

Livingston, J., 2005, 'How Valuable is a Good Reputation? A Sample Selection Model of Internet Auctions', Review of Economics and Statistics.87453–465

Lucking-Reiley, D.; Bryan, D.;Prasad, N. and D. Reeves, 2006, 'Pennies from eBay: the Determinants of Price in Online Auctions', forthcoming, *The Journal of Industrial Economics*.

McDonald, C. and Slawson, V., 2002, 'Reputation in an Internet Auction Market', *Economic Inquiry*, 40, pp. 633-650.

Melnik, M. and Alm, J., 2002, 'Does A Seller's eCommerce Reputation Matter? Evidence from eBay Auctions', *The Journal of Industrial Economics*, 50, pp. 337-350.

Ockenfels, A. and Roth, A., 2006, 'Late and Multiple Bidding in Second Price Internet Auctions: Theory and Evidence Concerning Different Rules for Ending an Auction', Games and Economic Behavior.55297–320

Resnick, P. and Zeckhauser, R., 2002, 'Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System', in Michael R. Baye (Ed), *Advances in Applied Microeconomics*, Vol. 11 (Elsevier, Amsterdam).

Resnick, P.; Zeckhauser, R.; Swanson, J. and Lockwood, K., 2006, 'The Value of Reputation on eBay: A Controlled Experiment', *Experimental Economics*, 9, pp. 79–101.

Roth, A. and Ockenfels, A., 2002, 'Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet', *American Economic Review*, 92, pp. 1093–1103.

Snyder, M. and Cantor, N., 1979, 'Testing Hypotheses about other people: The use of historical knowledge', *in Journal of Experimental Social Psychology*, 15, pp. 330–342.

Notes

[1]Although eBay started in the U.S., it is rapidly becoming a European and worldwide phenomenon. In the second quarter of 2005, 46% of eBay's revenue originated from non-U.S. operations. According to Nielsen, eBay is the leading e-commerce site in Germany, UK, France and Italy.

[2]See Dellarocas [2003] and Bajari and Hortaçsu [2004] for surveys of these results.

[3]In Cabral and Hortaçsu [2006] we go a bit further and look at how a seller's activity evolves over his lifetime. We show that a typical seller starts his career with a substantially higher fraction of transactions as a buyer relative to later stages of his career as an eBay trader. This suggests that sellers invest in building a reputation as a buyer and then use that reputation as a seller.

[4]Hubbard [2002] examines the California vehicle emissions inspection market and shows that consumers are 30 percent more likely to return to a firm at which they previously passed an inspection than to one at which they previously failed. A more detailed exam of the data rejects a pure 'bootstrap' model in favor of one where consumers learn about the seller's type.

Jin and Leslie [2008] study L.A. restaurants' incentives for hygiene. They show the incentives are greater in chain restaurants and restaurants frequented by repeat customers. This observation is consistent with a dynamic reputation story.

Finally, Abbring, Chiappori, and Pinquet [2003] develop a test similar to ours in the context of auto insurance. In the French auto insurance market, an accident increases the cost of future accidents. An implication of moral hazard is that the first accident should *decrease* the arrival rate of future accidents.

Abbring, Chiappori, and Pinquet [2003] fail to find evidence of such decrease in accident rate.

[5]We will not attempt a detailed account of how eBay has evolved and what its trading rules are; the interested reader may find this in a number of survey articles such as Dellarocas [2003], Bajari and Hortaçsu [2004]; and in the popular press (e.g., Cohen, 2002).

[6]Success is defined as a bid above the minimum bid or a secret reserve price set by the seller. eBay collects its fee even if the physical transaction does not take place.

[7]eBay also offers an escrow service, but this service is used for only a small fraction of the transactions.

[8]There have been several changes on eBay regarding how these ratings can be given by the users. Since 1999, each grade/comment has to be linked to a particular transaction on eBay. Typically, eBay stores transactions data (including price) only for 90 days; hence, this restricts the extent of 'historical research' that a buyer can conduct.

[9]Indicators (b) and (c) have only been presented since March 1, 2003.

[10]Of course, 'old' buyers may know about private transactions that they did not comment on.

[11]eBay stores data on completed auctions for 30 days. We attempted to get data from all completed auctions in the above period.

[12]Bajari and Hortaçsu [2003], Melnik and Alm [2002] and Lucking-Reiley, Prasad and Reeves [2006].

[13]An important difference between these two types of coins is that, while the

proof set is in mint condition (and preserved in a plastic container), the gold coin may come in various grades. In our data, we found three different ones: MS–70, MS–69 and MS–67, in decreasing order of value.

[14]We repeated some of our regressions below using both negatives and negatives+neutrals as a measure of bad reputation. We obtained qualitatively similar results.

[15]For surveys of these papers, see Bajari and Hortaçsu [2003], Resnick et al. [2003]

[16]According to eBay rules this is equal to the second highest bid plus the bid increment.

[17]We found out about this policy change by accident. We should point out that before March 1st, 2003, this information was already available to bidders. However, in order to see the fraction of the seller's negative comments, the bidder would have to click on the seller's highlighted username (which would take the bidder to a new 'feedback profile' page) and manually compute the ratio of negative to total feedback comments.

[18]This regression corrects standard errors by allowing for heteroskedasticity at the seller level. We also added a dummy variable for *hdoutlet*. Omission of either of these features lead to significance of the coefficient at higher levels.

[19]Strictly speaking, eBay reports the overall positive feedback score, which is slightly different from the total number of feedbacks. However, the correlation between the two measures is 0.96.

[20]Resnick and Zeckhauser [2001], using a different data sample in which they match transactions with feedback data, estimate this probability at approximately 50%.

[21]Detailed results are available upon request.

[22]The alternative age measure, based on days since first registering, leads to similar results.

[23]In a previous draft, we used raw growth rates instead of age-detrended growth rates. The results were qualitatively similar.

[24]For many sellers, longer evaluation periods would include subsequent negative feedback. We believe a four-week window is a good balance between avoiding loss of data and statistical significance.

[25]Many times, when an eBay seller receives a negative comment, there is a 'war of words' between the seller and the buyer who places the negative. During this 'war of words,' the two parties can give several negatives to each other within a period of two or three days. We excluded the negative comments that sellers received during such episodes, and concentrated on the timing between *de novo* negative feedback comments.

[26]We redid Table III with negative comments only (that is, not counting neutrals). We obtain the same signs and about the same coefficient values, but do lose a little bit on statistical significance: the test for equality was rejected at 1% for Thinkpad and at 10% for Eagle and Silver; the p-value for Teddy was 15%.

[27]As mentioned above, we did not find any significant patterns in buyer feedback rates. Unfortunately, the portion of our sample for which we observe first negatives is very small, so we cannot really directly test our assumption.

[28]Some examples of comments associated to negative ratings include: 'Great to deal with... very fast... excellent communication too!' 'Good transaction.' 'Excellent person to work with and I would highly recommend! Thanks!'

'Received in great shape. Thank you.' 'I received fast and friendly.'

[29]As for Table V, we consider age-detrended growth rates. In a previous version of the paper, we considered raw growth rates and obtained similar results.

[30]The first negative results, when broken down by category, are still strongly significant, and with the correct signs. However, sample sizes are very small in the 'second negative' case.

[31]As Footnote 2 of Table III states, we computed growth rates as differences in logs. When computed as the ratio $(x_{t+1} - x_t)/x_t$, we obtained different values but the same qualitative patterns.

[32]We also experimented with 1 day and 5 day periods. Our results are robust to the choice of window length.

[33]There were only four instances of this in our sample.

[34]We should note that it is not at all clear whether this would play out in an equilibrium setting. However, since eBay users suggested this as an alternative explanation, we decided to evaluate its merits.

[35]On eBay one can also observe what each user wrote about each other.

[36]Interestingly, our data suggests a lower critical threshold for giving negatives in the Teddy market than in the Thinkpad market: the average negative comment-giver in the Thinkpad market gave negatives 10% of the time, whereas the average complainant in the Teddy market complained only 3% of the time. We speculate that this result may very loosely be attributed to our observation that the Teddy market on eBay can be seen as a 'community of collectors' with frequent repeated interactions, where wrong doings are less tolerated, whereas transactions in the Thinkpad market are not typically

repeated.

[37]We randomly mixed the order of the comments so that the student could not tell which was the first, which was the second negative. We also allowed for the following possibilities: 'repeat' (remarks are literally identical), 'mistake' (remarks are clearly positive even though a negative was given), and 'difficult to tell.'

[38]There is an extensive psychology literature on this, including Asch [1946], Snyder and Cantor [1979] and Hoch and Ha [1986].

[39]This 45 day period is particularly significant as it includes the pre- and post-Christmas seasons, the busiest on eBay.

[40]While these frameworks can be applied to a variety of situations, we will focus here on the issue of seller reputation.

[41]See Cabral [2005] for further discussion of these two alternative approaches.

Table I

Distribution of Feedback Aggregates Across Sellers

	Number of Positives	Number of Negatives	Number of Neutrals	$N/(N+P)$ (entire history)
Mean	1,625	4.9	7.2	0.009
Std. Dev.	3,840	25.1	33.5	0.038
Min.	0	0	0	0
Max.	52,298	651	654	1
1%	0	0	0	0
5%	5	0	0	0
10%	18	0	0	0
25%	99	0	0	0
50%	397	1	1	0.0028
75%	1,458	3	4	0.0092
90%	4,361	9	13	0.021
95%	7,134	19	29	0.034
99%	15,005	52	86	0.068
N	819	819	819	795

Table II

Cross Sectional Regressions

Model #	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	log(p)	log(p)	log(p)	log(p)	completed sale	log(# bids)
% negative comments	−7.54 (2.51)*	−7.54 (9.88)	0.68 (6.81)	5.16 (7.75)	−1.96 (1.09)*	−5.35 (3.31)
Total # of feedbacks	0.05 (0.04)	0.05 (0.03)*	0.00 (0.00)	0.00 (0.00)	−0.003 (0.001)**	−0.011 (0.004)**
% negative comments after format change				−15.80 (7.83)**	0.01 (1.92)	1.90 (3.65)
Total # of feedbacks after format change				0.00 (0.01)	−0.001 (0.001)	−0.002 (0.003)
Indicator for *hdoutlet*			4.81 (0.43)***	4.80 (0.43)***	0.44 (0.07)***	2.68 (0.29)***
Listing includes photo	−0.18 (0.05)**	−0.18 (0.14)	−0.04 (0.10)	−0.04 (0.10)	−0.18 (0.03)***	−0.36 (0.07)***
Refurbished item	−0.62 (0.91)	−0.61 (1.06)	−2.43 (0.66)***	−2.45 (0.64)***	−0.16 (0.07)**	−0.88 (0.27)***
Paypal accepted	0.17 (0.21)	0.17 (0.18)	−0.05 (0.09)	−0.06 (0.09)	−0.30 (0.05)***	−0.33 (0.09)***
Credit cards accepted	0.36 (0.23)	0.36 (0.10)***	0.29 (0.10)***	0.28 (0.10)***	0.67 (0.03)***	0.99 (0.07)***
Auction duration (days)	0.04 (0.03)	0.04 (0.02)**	0.04 (0.02)***	0.04 (0.02)***	0.12 (0.005)**	0.11 (0.01)***
Peak hour	0.10 (0.08)	0.10 (0.11)	0.05 (0.10)	0.03 (0.10)	−0.01 (0.02)	0.12 (0.05)**
Eagle	0.52 (0.08)***	0.52 (0.51)	0.91 (0.50)*	0.94 (0.50)*	−0.09 (0.05)*	−0.66 (0.21)***
Proof Set	0.84 (0.05)***	0.84 (0.49)*	1.21 (0.48)**	1.23 (0.48)**	−0.07 (0.05)	−0.68 (0.20)***
Teddy	−1.04 (0.10)***	−1.04 (0.53)*	−0.50 (0.50)	−0.48 (0.50)	−0.09 (0.05)*	−1.22 (0.20)***
log(Minimum Bid)	0.003 (0.00)***	0.00 (0.00)***	0.00 (0.00)***	0.00 (0.00)***	−2.8E−04 (1.1E−04)**	−0.0016 (3E−04)***
Indicator for new format				−0.26 (0.13)**	0.22 (0.07)***	0.31 (0.14)**
Constant	2.468 (0.664)**	2.47 (0.63)***	2.05 (0.62)***	2.64 (0.54)***	0.39 (0.10)***	1.02 0.28
Observations	1053	1053	1053	1053	1053	1053
R−squared	0.42	0.42	0.52	0.53	0.74	0.67

Notes:
1. Day of week and calendar week controls are added in all specifications.
2. In columns 2–4, robust standard errors (clustered by seller id) are reported in parentheses.
3. Significance levels: 10, 5, 1 percent (one to three stars).

Table III

Impact of Negatives on Sales Growth (%)

	Avg. Week. Growth R.	Object Thinkpad	Eagle	Silver	Teddy
First Negat.	Before	5.17	6.88	5.07	12.06
	After	−7.56	−4.67	−8.25	−5.28
	Difference	−12.74***	−11.56***	−13.32***	−17.34***
	Std. Error	4.89	3.56	3.44	3.69
	N	66	95	130	136
Second Negat.	Before	2.57	−1.67	3.41	6.41
	After	9.53	9.00	7.61	7.51
	Difference	+6.96	+10.67**	+4.20	+1.10
	Std. Error	5.03	4.82	5.96	6.12
	N	37	70	78	83
Third Negat.	Before	8.14	2.75	2.81	1.00
	After	4.91	−2.53	2.13	9.70
	Difference	−3.23	−5.28	−0.68	+8.70
	Std. Error	6.14	7.47	3.21	6.22
	N	28	52	57	64

Notes:

1. Standard errors in parentheses. Significance levels 10, 5, 1 percent (one to three stars).
2. Weekly detrended growth rates are based on the number of sales-related feedbacks received by the seller.
3. Growth rate in week t = ln(no. feedbacks in week t) - ln(no. feedbacks in week $t-1$).
4. Weekly growth rates are averaged over 4 week periods taken before and after the reception of a negative.

TABLE IV

IMPACT OF NEGATIVES ON SALES GROWTH (%).
SUBSAMPLE OF 'MISTAKEN' NEGATIVE FEEDBACK RATINGS.

		Average age-detrended weekly growth rate
First Negative	Before	10.97
	After	−13.21
	Difference	−24.18***
	Std. Error	7.95
	N	41
Second Negative	Before	3.56
	After	16.92
	Difference	+13.36
	Std. Error	11.84
	N	19

Notes: See Table III.

Table V

Frequency of Negative Feedback

T1: Sale-Related Feedbacks to First Negative.
T2: Sale-Related Feedbacks Between 1st and 2nd Negative.
ET: Average Number of Sale-Related Feedbacks Between Negatives.

	All Cat.	Thinkpad	Eagle	Silver	Teddy
T1	240.88	93.24	339.66	267.71	226.99
T2	188.76	58.59	199.24	261.26	199.86
ET	162.39	50.8	216.1	189.61	163.5
T1 − T2	52.12	34.66	140.41	6.45	27.13
T1 > T2 : p-val	0.021	0.036	0.017	0.452	0.27
T1 − ET	78.48	42.44	123.56	78.09	63.49
T1 > ET: p-val	0.0002	0.0083	0.02	0.025	0.044
T2 − ET	26.36	7.79	-16.86	71.64	36.36
T2 > ET: p-val	0.032	0.176	0.73	0.027	0.089
N	311	58	79	78	96

Notes:
1. Sample includes all sellers with more than 2 negatives received on sales.
2. T-tests are conducted using within seller differences.
3. ET calculated as total feedback/(# negatives & neutrals), where we count only sales transactions.

Table VI

Frequency of Negative Feedback: Correcting for Selection Bias

T1: Sale-related feedbacks to first negative.
T2: Sale-related feedbacks between 1st and 2nd negative.

	All Cat.	Thinkpad	Eagle	Silver	Teddy
Sample: sellers with 2+ negatives born after Oct 24, 2002					
	All sellers	Thinkpad	Eagle	Silver	Teddy
T1	196	26.5	175	238	174
T2	80	23.2	37	501	64
T1−T2	116	3.2	138	263	110
T1>T2: p-val	0.03	0.36	0.03	0.22	0.03
N	20	6	5	4	5
Sample: sellers with 1+ negatives born after Oct 24, 2002					
	All sellers	Thinkpad	Eagle	Silver	Teddy
T1	257	26.5	206	403	346
T2*	174	23.2	107	253	277
T1−T2	83	3.2	99	150	69
T1>T2: p-val	0.035	0.36	0.03	0.2	0.04
N	28	6	7	7	8

* T2=T1 for sellers with only 1 negative.

Table VII

Alternative Explanations for Differences in Arrival Times. Dependent variable for (1) and (2): buyer's negative comment was followed by seller's negative comment. Dependent variable for (3) and (4): frequency of negative comments by the buyer who gave a particular negative comment

	Dependent variable			
	(1) Retaliation	(2) Retaliation	(3) Profile	(4) Profile
2nd Negative	0.016 (0.055)	0.025 (0.063)	0.011 (0.013)	0.011 (0.015)
3rd Negative	0.030 (0.059)	0.043 (0.068)	0.003 (0.015)	-0.003 (0.016)
4th Negative	-0.005 (0.064)	0.000 (0.069)	0.020 (0.020)	0.020 (0.021)
5th Negative	0.044 (0.068)	0.118 (0.074)	0.015 (0.018)	0.011 (0.018)
6th Negative	0.053 (0.071)	0.107 (0.073)	0.045 (0.023)*	0.040 (0.024)
Percentage of Negatives	4.664 (1.907)**		-0.053 (0.372)	
Number of transactions	0.000 (0.000)		-0.000 (0.000)	
eagle dummy	0.100 (0.120)	(seller f.e.)	-0.079 (0.038)**	(seller f.e.)
mint dummy	0.000 (0.094)		-0.087 (0.037)**	
teddy dummy	0.091 (0.089)		-0.071 (0.039)*	
Constant	0.115 (0.098)	0.239 (0.045)***	0.105 (0.043)**	0.038 (0.012)***
Observations	558	567	575	584
R-squared	0.03	0.38	0.06	0.38

Robust standard errors in parentheses. Significance levels 10, 5, 1 percent (one to three stars).

Table VIII

Reasons for Negative Feedback (%)

	First Negative	Second Negative
Misrepresented item	22	16
Bad communication	19	20
Item damaged	15	17
Item not received	10	13
Backed out	7	4
Angry / upset	7	7
Overcharged shipping	6	4
Slow shipping	6	10
Bad packaging	4	6
Feedback issues	3	3
Bid on own item	1	1
Total	100	100

Table IX

Can Reputational Variables Predict Seller Exits?
Dependent variable: seller exited by January 4, 2004

	All Exits				
	All sellers	Thinkpad sellers	Eagle sellers	Silver sellers	Teddy sellers
Log. number negat. May 03	0.34 (0.22)	0.36 (0.49)	0.54 (0.42)	-0.12 (0.71)	0.72 (0.40)*
Log. number posit. May 03	-0.58 (0.11)***	-0.57 (0.23)***	-0.53 (0.20)***	-0.87 (0.39)***	-0.71 (0.19)
Observations	819	199	255	115	250

	Opportunistic Exits				
	All sellers	Laptop sellers	Gold sellers	Silver sellers	Beanie sellers
Log. number negat. May 03	1.02 (0.40)**	1.98 (1.09)*	3.44 (0.99)***	-0.92 (0.96)	-0.09 (0.73)*
Log. number posit. May 03	-0.33 (0.22)	-0.96 (0.53)*	-1.09 (0.47)**	0.36 (0.54)	0.29 (0.38)
Observations	715	174	219	102	220

Notes: 1. Standard errors in parentheses. Significance levels 10, 5, 1 percent (one to three stars).

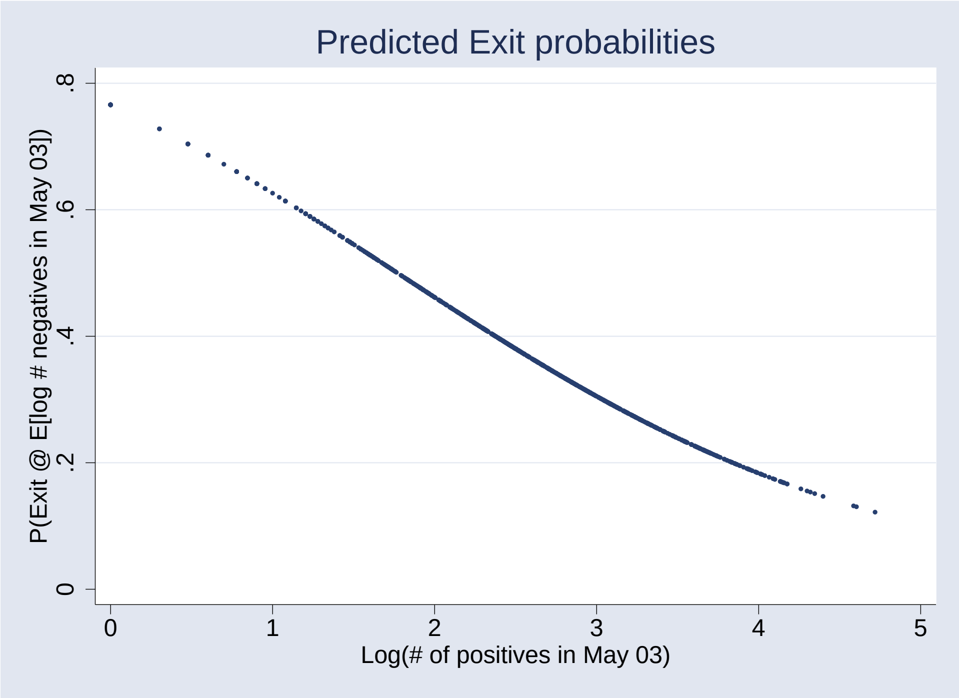

Figure 1
Predicted Exit Probabilities

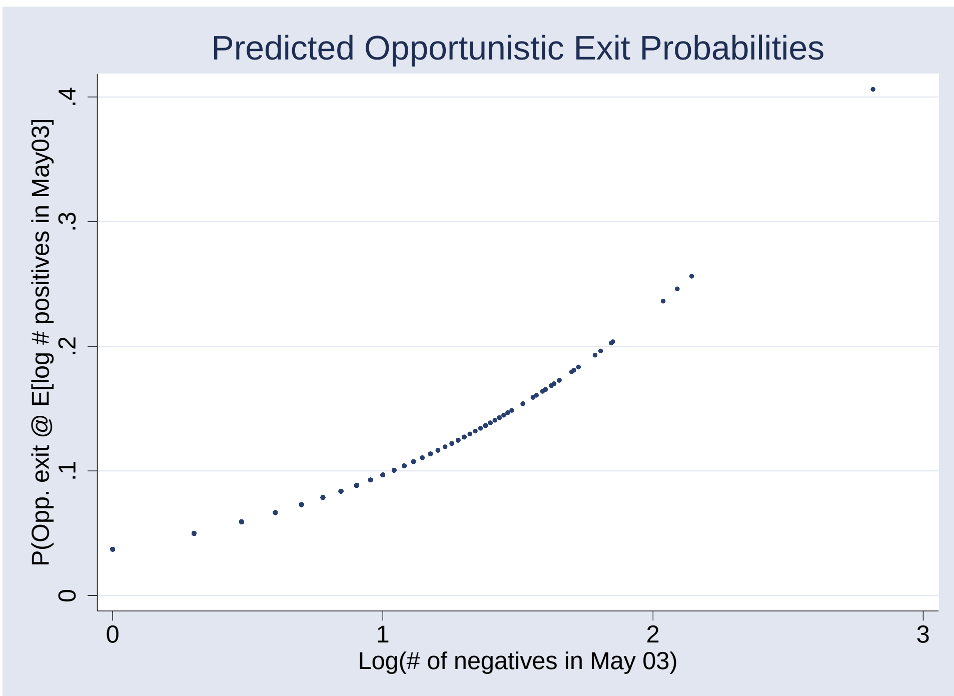

Figure 2
Predicted Opportunistic Exit Probabilities