The Dynamics of Seller Reputation: Evidence from eBay

Luís Cabral*
New York University and *CEPR*

Ali Hortaçsu*
University of Chicago and *NBER*

March 2006

Abstract

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 7% to a negative 7%; 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.

We consider a series of theoretical models and measure them against these empirical results. Regardless of which theoretical model best explains the data, an important conclusion of our paper is that eBay's reputation system gives way to noticeable strategic responses from both buyers and sellers.

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

---

1   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 theory implications.

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 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 (2001), Eaton (2002), Ederington and Dewally (2003), Houser and Wooders (2003), Kalyanam and McIntyre (2003), Livingston (2002), Lucking-Reiley, Bryan, Prasad and Reeves (2000), McDonald and Slawson (2002), Melnik and Alm (2002), Resnick and Zeckhauser (2001).[2] Resnick, Zeckhauser, Swanson and Lockwood (2003) point out the

---

[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 (2002), Resnick, Zeckhauser, Swanson and Lockwood (2003), and Bajari and Hortaçsu (2004) for surveys of these results.

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 (2004) 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. Finally, there are also a number of papers on reputation in markets other than eBay, including Jin and Leslie (2004). We will return to this later.

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 9% 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 rate drops from a positive 7% to a negative 7%. 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. Finally, a typical seller starts his career as a buyer (that is, will make more transactions as a

buyer than as a seller).

Our data clearly suggests that reputation matters: buyers react to information about seller reputation; and sellers' actions, too, are influenced by reputation considerations.

We then propose a theoretical framework that squares with key empirical observations in our dataset. We consider a model in the tradition of Kreps et al (1982), featuring adverse selection and moral hazard. In equilibrium, a seller with a perfect record puts a lot of effort into maintaining that record. Once a negative is received, the marginal return to investing in reputation decreases dramatically. As a result, the seller (if a low type) decreases effort. Lower effort leads to lower quality transactions; and this in turn leads to a higher likelihood of negative transactions (and negative feedback). In this way, our model is consistent with our observations regarding sales rate and frequency of negative feedback.

Our main contribution to the study of online reputation mechanisms is twofold: First, we analyze panel data in addition to cross-section data. Second, we analyze the impact of seller reputation on buyer *and* seller behavior. We next discuss our perceived relevance of these contributions.

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.

To the best of our knowledge, together with Jin and Kato (2004) on baseball card sellers and Jin and Leslie (2004) on L.A. restaurants, ours is one of the first empirical papers to study the relation between seller's reputation and seller's actions. In fact, we believe ours is the first study to look at the lifetime dynamics of sellers' reputation: how a reputation is built and how it is used; how sellers make choices to influence their reputation; and how sellers react to changes in their reputation.

The paper is structured as follows. In Section 2, we briefly describe the institutional setup of eBay, in particular the mechanics of its reputation mechanism. In Section 3, we describe our dataset. The main empirical results are presented in Section 4, whereas Section 5 develops a series os possible theoretical interpretations. Section 6 concludes the paper.

2 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. 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 and in the popular press.[3] For the purposes of our analysis, it will not be too inaccurate to characterize eBay's auction mechanism as a variant of the second-price auction.[4]

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.[5] Most importantly, to enable reputation mechanisms to regulate trade, eBay uses an innovative feedback system.[6] 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.[7]

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 (since March, 2001), (c) the date when the seller registered with eBay, and (d) a summary of the most recent feedback received by the seller.[8] 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

---

[3]See Cohen (2002) for an entertaining historical account of eBay. Survey articles on Internet auctions include Lucking-Reiley (1999), Dellarocas (2003), and Bajari and Hortaçsu (2004).

[4]In reality eBay auctions are dynamic auctions in which bidders place (possibly multiple) "proxy bids" indicating their maximum willingness-to-pay. See Roth and Ockenfels (2002), Ockenfels and Roth (2003), and Bajari and Hortaçsu (2003) for detailed analyzes of dynamic bidding behavior on eBay.

[5]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.

[6]eBay does offer an escrow service for use with especially valuable goods, though this service is used for only a small fraction of the transactions.

[7]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 (in particular price) only for 90 days; hence, this restricts the extent of "historical research" that a buyer can conduct.

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

claimed in the introduction, this is an environment where the economic analyst has the same information that a new buyer has about a seller.[9] We will thus take this informational equivalence as given when formulating our theoretical model and its empirical implications.

3 Data description

Our data was collected from eBay's website at monthly intervals between October 24, 2002 and March 16, 2003. 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 (displayed in Figure 1):[10]

1. Collectible coins. We chose this category since 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.[11] We selected at two different kinds of coins: the 1/10 oz. 5 dollar gold coin of 2002 vintage (gold American Eagle); and the 2001 silver proof set (ten coins of different denominations), both produced by the U.S. mint.[12] The average sale prices in our data set are \$50 for the gold coin \$78 for the proof set.

2. IBM Thinkpad T23 PIII notebook computers. 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.

---

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

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

[11]Bajari and Hortaçsu (2003), Melnik and Alm (2002) and Lucking-Reiley, Prasad and Reeves (2000).

[12]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.

3. 1998 Holiday Teddy Beanie Babies, 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, as shown in Figure 2, 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 4.1. 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 1 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., 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.

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 1, 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.[13]

---

[13]Some negative comments for sellers have the following textual content: "THIS PERSON RIPPED ME OFF, SENT SHODDY ITEM INSTEAD OF ITEM LISTED," "Sold product

Figure 1: Pictures of auctioned objects considered in this study: the 1/10 oz. 5 dollar gold coin of 2002 vintage (gold American Eagle) and the 2001 silver proof set (ten coins of different denominations), both produced by the U.S. mint; the IBM Thinkpad T23 PIII notebook computer; and the 1998 Holiday Teddy Beanie Baby.

Feedback Summary
218 positives. **128** are from unique users.

0 neutrals.

1 negatives. **1** are from unique users.

See all feedback reviews for wsb5.

ebaY ID card   wsb5(127) ⭐
Member since: Saturday, Jun 19, 1999 Location: United States

Summary of Most Recent Reviews

	Past 7 days	Past month	Past 6 mo.
Positive	12	51	116
Neutral	0	0	0
Negative	0	0	0
Total	**12**	**51**	**116**
Bid Retractions	0	0	0

View wsb5 's Items for Sale | ID History | Feedback About Others

Feedback Reviews for wsb5		Feedback Help	FAQ

leave feedback
for wsb5

If you are wsb5 :
Respond to comments

wsb5 was the **Seller = S**
wsb5 was the **Buyer = B**

Left by	Date	Item#	S/B
rattman50(11) ⭐	Apr-29-03 14:05:51 PDT	3019804072	S
Praise : Nice coin! Fast shipment!			
silverpeacedollar(26) ⭐	Apr-29-03 09:09:31 PDT	3018674118	S
Praise : hi great job nice coin and good service thanks!!!!!!			
z3forefun(351) ⭐	Apr-29-03 06:39:59 PDT	3018676358	S
Praise : very nice coin, accurately represented, fast shipping			
patrag40(161) ⭐	Apr-28-03 17:41:37 PDT	3018673349	S
Praise : The coin has been cleaned but a great deal			
bernardtreeman(62) ⭐	Apr-25-03 18:11:09 PDT	3014810862	S
Praise : thanks for a nice coin. +++++++++++++AAAAAAA			
kucak1(114) ⭐	Apr-25-03 06:07:31 PDT	3013485158	S
Praise : HIGHLY RECOMMEND THIS GENTLEMAN!!! Thanks, Willard!!!			
rdt9819 (73) ⭐	Apr-24-03 14:37:12 PDT	3018676926	S
Praise : GOOD TRANSACTION WOULD BUY AGAIN A+++++++			
bfjfkman (24) ⭐	Apr-23-03 15:03:21 PDT	3018675234	S
Praise : Fast Delivery, Good Packaging, Great Deal. (Very Nice Coins, too.)			
bfjfkman (24) ⭐	Apr-23-03 15:02:00 PDT	3018677589	S

Figure 2: Feedback page of user *wsb5*.

8

Table 1: 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

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. Indeed, observe that in Table 1 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 striking similarity, we will henceforth lump negative and neutral comments together when referring to "negative" feedback.

---

he didn't have! Will not send refund! I am filing charges! No ansr," "Overgraded junk. Does not respond to emails. An irresponsible seller. Avoid him." On the other hand, we found that more than 40% of the positive comments contain the expression "A+". Some more colorful positive comments were: "Heaven must be missing an angel! Transaction couldn't be better! Thank U!!!" and "Mega cool mad phat deal nasty crazy cool even. Thanks."

4 Empirical results

In this section, we present our main empirical findings. They are divided into five subsections. In Section 4.1, 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 4.2 and 4.3 we study the impact of the first few negative feedback ratings in a seller's history: impact on growth (Section 4.2) and impact on the frequency of negative feedback (Section 4.3). Section 4.4 looks at seller exit: who is more likely to exit and what pattern of feedback do we observe near exit time. Finally, Section 4.5 examines the relative pattern of purchases and sales over a seller's lifetime.

4.1 Reputation and price

At the most basic level, we would expect a better seller reputation to influence the price paid by buyers for an otherwise identical object. To investigate this hypothesis, several papers in the prior empirical literature on eBay have run regressions of the form:[14]

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

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 2 reports our results from such cross-sectional regressions. In these regressions, the dependent variable is the log of the highest bid registered in the auction.[15] 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 1000 additional reviews is associated with a 5% increase in sale price.

Observe that the magnitude of this estimate is close to the findings of several other cross-sectional studies. In particular, the 5% price premium

---

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

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

Table 2: Cross sectional regressions. Dependent variable: log of highest bid.

	(1)	(2)	(3)	(4)
% Negative comments	-7.54	-7.54	0.68	5.16
	(2.51)*	(9.88)	(6.81)	(7.75)
# Feedbacks	0.05	0.05	0.00	0.00
	(0.04)	(0.03)*	(0.00)	(0.00)
% Neg. comm. after format change				-15.80
				(7.83)**
# Feedbacks after format change				0.00
				(0.01)
Indicator for new format				-0.26
				(0.13)**
Indicator for "hdoutlet"			4.81	4.80
			(0.43)***	(0.43)***
Listing includes photo	-0.18	-0.18	-0.04	-0.04
	(0.05)**	(0.14)	(0.10)	(0.10)
Refurbished item	-0.62	-0.61	-2.43	-2.45
	(0.91)	(1.06)	(0.66)***	(0.64)***
Paypal accepted	0.17	0.17	-0.05	-0.06
	(0.21)	(0.18)	(0.09)	(0.09)
Credit cards accepted	0.36	0.36	0.29	0.28
	(0.23)	(0.10)***	(0.10)***	(0.10)***
Auction duration (days)	0.04	0.04	0.04	0.04
	(0.03)	(0.02)**	(0.02)***	(0.02)***
Peak hour	0.10	0.10	0.05	0.03
	(0.08)	(0.11)	(0.10)	(0.10)
American Eagle	0.52	0.52	0.91	0.94
	(0.08)***	(0.51)	(0.50)*	(0.50)*
Mint Set	0.84	0.84	1.21	1.23
	(0.05)***	(0.49)*	(0.48)**	(0.48)**
Beanie Baby	-1.04	-1.04	-0.50	-0.48
	(0.10)***	(0.53)*	(0.50)	(0.50)
log(Minimum Bid)	0.003	0.00	0.00	0.00
	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Constant	2.468	2.47	2.05	2.64
	(0.664)**	(0.63)***	(0.62)***	(0.54)***
Observations	1053	1053	1053	1053
R-squared	0.42	0.42	0.52	0.53

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).

11

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 2 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 finding 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 2 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 is correlated with large variations in price.

In summary, the results in the first three columns of Table 2 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

12

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 (see Figure 2).[16]

In column (4) of Table 2, 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.[17] 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 -15% 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 feedbacks of the seller is no different when interacted with the format change, since eBay displayed this information before and after the format change.[18]

■ **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 patters 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,[19] with 3 negatives and 3 neutrals (i.e. 1.4% of feedbacks were

---

[16]We found out about this policy change by accident. We should point out that before March 1st, 2003, the information shown in Figure 2 was already available to bidders. However, in order to see the fraction of 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 $N/(N+P)$.

[17]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.

[18]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.

[19]Resnick and Zeckhauser (2001), using a different data sample in which they match

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 most recent 6 month period do not have a statistically significant correlation with feedback reception.[20] There are differences across object categories in the frequency of feedback reception (the eagle and mintset coins are more likely to receive feedback than beanies and Thinkpads), but this most likely reflects different social norms across different 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: 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: when and why they exit and how the relative importance of purchases and sales evolves over time.

4.2   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. Accordingly, we averaged the weekly sales 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.[21] We then conducted paired t-tests of the null hypothesis of equality of growth rates before and after the negative feedback event.[22]

transactions with feedback data, estimate this probability at approximately 50%.

[20]Detailed results are available upon request.

[21]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.

[22]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

Table 3: Impact of negatives on sales growth (%).

	Avg. Week. Growth R.	Object			
		Thinkpad	Proof set	G. Eagle	B. Baby
First Negat.	Before	7.12	6.85	9.04	14.19
	After	−6.76	−7.51	−3.89	−4.28
	Difference	−13.88***	−14.36***	−12.92***	−18.47***
	Std. Error	4.88	3.45	3.58	3.69
	N	66	130	95	136
Second Negat.	Before	3.96	4.50	−0.22	7.68
	After	9.93	8.00	9.47	8.03
	Difference	+5.97	+3.50	+9.69**	+0.36
	Std. Error	5.00	5.96	4.82	6.12
	N	37	78	70	83
Third Negat.	Before	9.19	3.80	3.58	2.00
	After	5.28	2.48	−2.09	10.25
	Difference	−3.90	−1.32	−5.68	+8.24
	Std. Error	6.14	3.22	7.44	6.23
	N	28	57	52	64

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

The results, reported in Table 3, are striking: For all four object categories, the impact of the first negative feedback comment is to slow growth by 14% a week, from a positive growth rate of about 7% to a negative growth rate of −7%.[23] Moreover, we find this difference is highly statistically significant. The difference in growth rates before an after the second negative feedback is positive. However, except for Golden American Eagle, the difference is not statistically significant. The impact of the third negative feedback also does not appear to be statistically significant.

Several notes are in order. First, our exercise depends importantly on the assumption that the probability of feedback is the same before and after a Negative. However, this is only a problem if buyers are somewhat reluctant to give *positive* comments about a seller after the seller has received her first (or second or third) Negative. Intuition suggests that the opposite is more likely to be true. 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 and are thus excluded from our sample. But intuition suggests that, if anything, this reinforces the point that the first Negative 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 this possibility, 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.[24] We repeated our analysis of the impact of negative feedback for this subsample. We considered both first and second negative feedbacks.

The results are reported in Table 4. Due to the small size of our "mistakes" sample, we pool together all four object categories.[25] Although the exact

---

excluded the negative comments that sellers received during such episodes, and concentrated on the timing between *de novo* negative feedback comments.

[23]These are the numbers for the Thinkpad. The values for the other products are of similar magnitude.

[24]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."

[25]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"

Table 4: Impact of negatives on sales growth (%). Subsample of "mistaken" negative feedback ratings.

Average weekly growth rate		
First	Before	13.86
Negative	After	−11.87
	Difference	−25.74***
	Std. Error	7.93
	N	41
Second	Before	6.34
Negative	After	18.03
	Difference	−11.70
	Std. Error	11.76
	N	19

Notes: See Table 3.

values are different from the full sample, we still have very large differences, significant both economically and statistically.

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.[26] Notice that the drop in sales rate might correspond either to buyer behavior (fewer bidders and thus more objects going unsold) or to seller behavior (anticipating fewer bidders, seller puts fewer objects up for sale). In Section 5 we discuss various implications of this and other pieces of empirical evidence.

4.3   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 consec-

---

case.

[26]As Footnote 2 of Table 3 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.

Table 5: 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	Mint	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.

utive negative comments by the same buyer. We also excluded any negative comments that were left within a two-day period after another negative.[27] Finally, we excluded those negative/neutral comments that were received as a "buyer."[28]

Table 5 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.

The results suggest that both T1 and T2 are greater than ET, and moreover

---

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

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

Table 6: 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	Mint	Teddy
Sample: sellers with 2+ negatives born after Oct 24, 2002					
	All sellers	Thinkpad	Eagle	Mint	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	Mint	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.

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 mint coins; and the difference T2−ET is not significant for the American Eagle coin or the Beanie Baby. Notice however that, for the most expensive item, the IBM 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 5 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 6. First, we redid the calculations from Table 5 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 6. 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 Mint coins; but the number of observations at this level is rather small.

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 only 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 6, 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 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.[29]

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 7, 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 n 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.

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. So first negatives would primarily be given by negative-prone buyers, whereas subsequent negatives would 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.[30] We then computed the percentage of negative comments that each of these reviewers left about others, as a measure of each reviewer's "critical attitude." In Table 7, columns (3) and (4), we regress the critical attitude of the reviewer leaving a particular

---

[29]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.

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

Table 7: 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 8: 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

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 the buyers who left subsequent negative feedback.[31]

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.[32] 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.

---

[31]Interestingly, our data suggests a lower critical threshold for giving negatives in the Beanie Babies market than in the laptop market: the average negative comment-giver in the laptop market gave negatives 10% of the time, whereas the average complainant in the Beanie Babies market complained only 3% of the time. We speculate that this result may very loosely be attributed to our observation that the Beanie Babies 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 laptop market are not typically repeated.

[32]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."

Finally, we consider the possibility that buyers are influenced by other buyers' behavior (herding, conformism, etc).[33] 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 8. 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).

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.

4.4   Reputation and exit

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 (pre- and post-Christmas are the busiest seasons on eBay). 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

---

[33]There is an extensive psychology literature on this, including Asch (1946), Snyder and Canto (1979) and Hoch and Ha (1986).
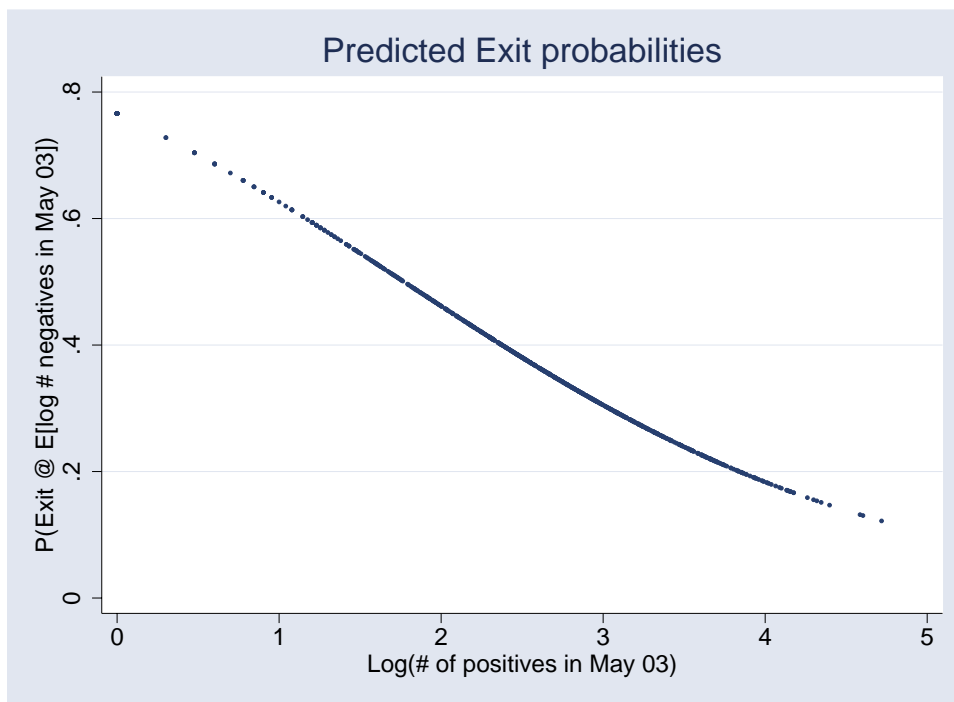
Figure 3: Predicted exit probabilities.

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 9, where we reported our results for the pooled sample of sellers, and also 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 3 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.

Next, we investigate whether the "exits" we see in our data set are accom-

Table 9: Can reputational variables predict seller exits? Dependent variable: seller exited by January 4, 2004.

	All Exits				
	All sellers	Laptop sellers	Gold sellers	Silver sellers	Beanie 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).

26

panied by "opportunistic" profit-taking by sellers, and whether reputational variables can predict such behavior.[34] In particular, 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 her last 25 transactions; whereas she had a total of 54 negatives in her previous transactions (the percentage of negatives and neutrals over 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 9, where we now defined the dependent variable to be an equal to 1 if the percentage of negatives within the last 25 transactions of a seller was more than twice the percentage of negatives within the seller's entire history. The results of these regressions are reported in the bottom panel of Table 9. Notice that although the number of positives that a seller has is still negatively correlated with the probability of exit, the number of negatives enter 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 4 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 out an economically significant relationship.

4.5   Traders' lifecycle

Casual observation of feedback histories suggests that many sellers appear to start out as "buyers," completing a string of purchases before attempting their first sale. As an example, Figure 4 plots the percentage of sell vs. buy transactions by user *bearsylvania*, an established Beanie Baby dealer, as a function of the number of weeks he has been active on eBay. As can be seen,

---

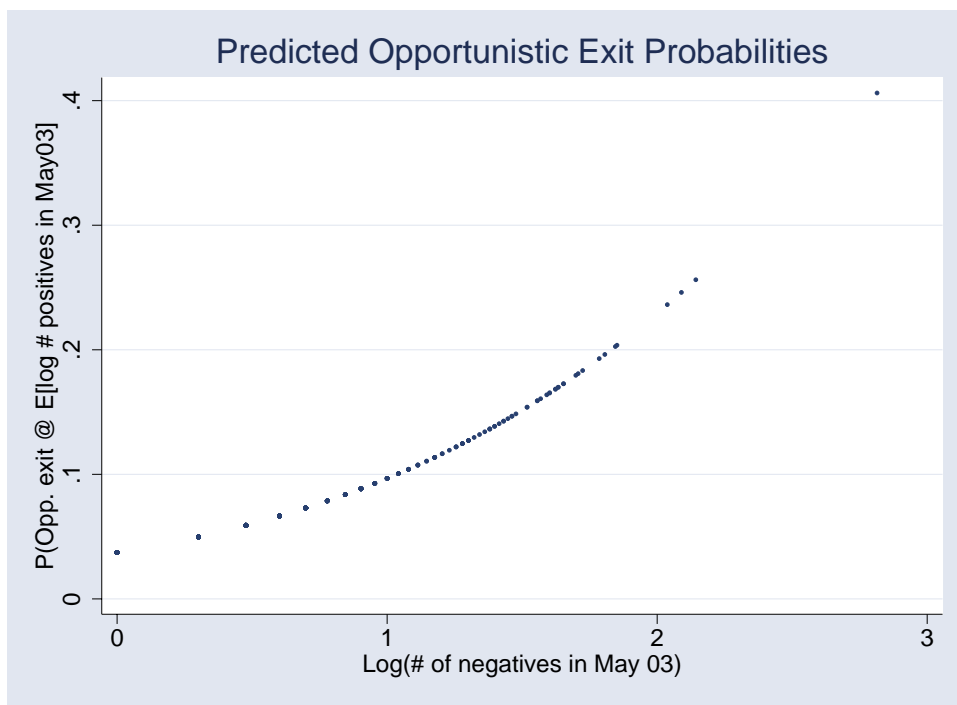[34]For a model of opportunistic use of reputation, see Phelan (2001). See also Gale and Rosenthal (1994).

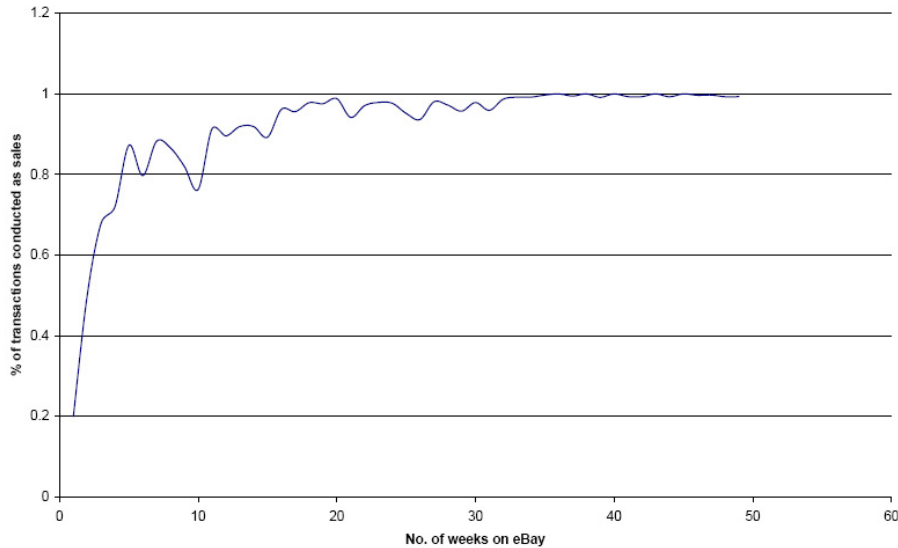Figure 4: Predicted opportunistic exit probabilities.

Figure 5: Becoming a seller: user *bearsylvania*'s percentage of transactions as a seller over time.

bearsylvania started out as a buyer first, and quickly changed the pattern of his transactions from purchases to sales.

To estimate the prevalence of this phenomenon, we looked at the first and last twenty comments received by each seller. We then defined a seller as having switched from being a buyer to being a seller if more than 50% of the first 20 comments referred to purchases, and more than 70% of the last 20 comments referred to sales.[35]

An important difficulty with implementing the above coding scheme with our data is that eBay does not report a buyer/seller classification for feedback comments received prior to June 16, 2001. Since about two-thirds of our sample sellers began their eBay careers prior to this date, we made our own assignment as buyer vs. seller based on the textual content of the comments.[36]

---

[35]To check the robustness of this definition of a "switch," we defined a second indicator with thresholds 40% and 80%, respectively.

[36]We automated the classification procedure by first calculating the empirical frequencies of word-stems such as "buy," "sell," "pay," "ship," "pack" across buyer/seller categories in a subsample of the post-June 16,2001 data. We then compared the likelihood of a given comment to be a "buyer" or "seller" comment based on the presence of these keywords. The accuracy of our classification of "seller" comments was remarkable: for the post-June 16,2001 data (for which we have eBay's classifications) we were able to classify all but 117 of 12952 comments correctly. Our classification of "buyer" comments was less accurate, since

Given the assumptions that go into our classification scheme, we will report some of our results for these two subsamples of sellers separately.

We found that 38% of Beanie Baby sellers, 22% of laptop sellers, 31% of gold coin sellers, and 31% of proof set sellers followed the "buy first, sell later" strategy (as defined above). We also found that, on average, 81% of a seller's last 20 transactions were sales, compared to 46% of the first 20 transactions. A paired t-test of equality of the two percentages reveals a strongly statistically significant increase in the percentage of sales (t-statistic equal to 25).[37]

These results show that "buying first and selling later" is a widespread phenomenon on eBay, and is somewhat more prominent in some object categories than others. For example, eBay is widely known as one of the main trading venues for Beanie Babies. It is conceivable that Beanie Baby enthusiasts first start out as buyers in this market, and switch to selling once they accumulate an inventory. On the other hand, laptop sellers are more likely to sell items they have acquired through other channels.

Next, we investigate the correlation of the "buy first sell later" indicator variable with the (log) number of positives and negatives in the seller's record. Column (1) of Table 10 reports the results of a logit regression using all sellers in the sample for whom we could ascertain the buy/sell breakdown. Column (2) uses the sample of sellers who joined eBay after June 16, 2001 (i.e., the set of sellers for whom we have direct data from eBay). Finally, in column (3) of this table, we look at the subsample of sellers who started their career before eBay began to report buyer/seller classifications of received feedback (thus we had to impute these classifications).

The regression yields very similar results across the samples, and appears to indicate that better sellers (those with many positives and few negatives) are less likely to have displayed this switching behavior. As can be seen in

---

most of these buyer comments contain very little information (we checked to see if human classification performed better in a subsample of comments; the improvement was marginal, precisely due to the lack of informative content). In particular, we classified 1934 of 5035 "buyer" comments as a "seller" comment, a 60% accuracy rate. Hence, our classification scheme is biased towards finding "sellers" as opposed to "buyers." To address this problem, when computing the percentage of a sales-related comments that a user gets within a given time period, we add 17% (the average bias in the control sample) on top of the percentage computed using our classification scheme.

[37]To make sure that these results were not driven by the assumptions needed to construct the buyer/seller classification for sellers with missing data, we repeated the same exercise using the post-June 16, 2001 sample of sellers. We found that, on average, 77% of last 20 transactions were sales, as opposed to 46% of the first 20 transactions. Once again the paired t-test strongly rejects equality.

Table 10: Seller's lifecycle. Dependent variable: seller switched from buying to selling.

	(1)	(2)	(3)
Log number negatives	-0.39 (0.10)***	-0.46 (0.20)**	-0.36 (0.11)**
Log number positives	0.33 (0.07)***	0.35 (0.13)***	0.27 (0.10)***
Observations	636	240	396

Sample:
1. Pooled sample.
2. Sellers who joined eBay after June 16, 2001. All transactions classified as "buy" or "sell" by eBay.
3. Sellers who joined eBay before June 16, 2001. Transactions were classified using the method described in Section 5.5.

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

Figure 6, a change in the log number of negatives from 2 to 5 is associated with a drop in this switching probability from 30% to 10%, thus the effect appears economically significant.[38]

5   Theory implications

Over the past twenty five years or so, a number of economic theories of reputation have been developed. While these theories can be applied to a variety of situations, we will focus here on the issue of seller reputation. Our goal is to determine the extent to which the empirical evidence presented in the previous section can shed light on the validity of economic theory.

For all its variety, the economic theory of reputation can be classified into a few possible frameworks. One, pioneered by the work of Klein and Leffler (1981) and Shapiro (1983), sees reputation as a coordination, or bootstrap,

---

[38]We also conducted a robustness check on our dependent variable by changing the threshold of being a "buyer" to having less than 40% of transactions as sales, and the threshold of being a "seller" to having more than 80% of transactions as sales. This modification does not appear to have an important effect on the coefficient estimates.
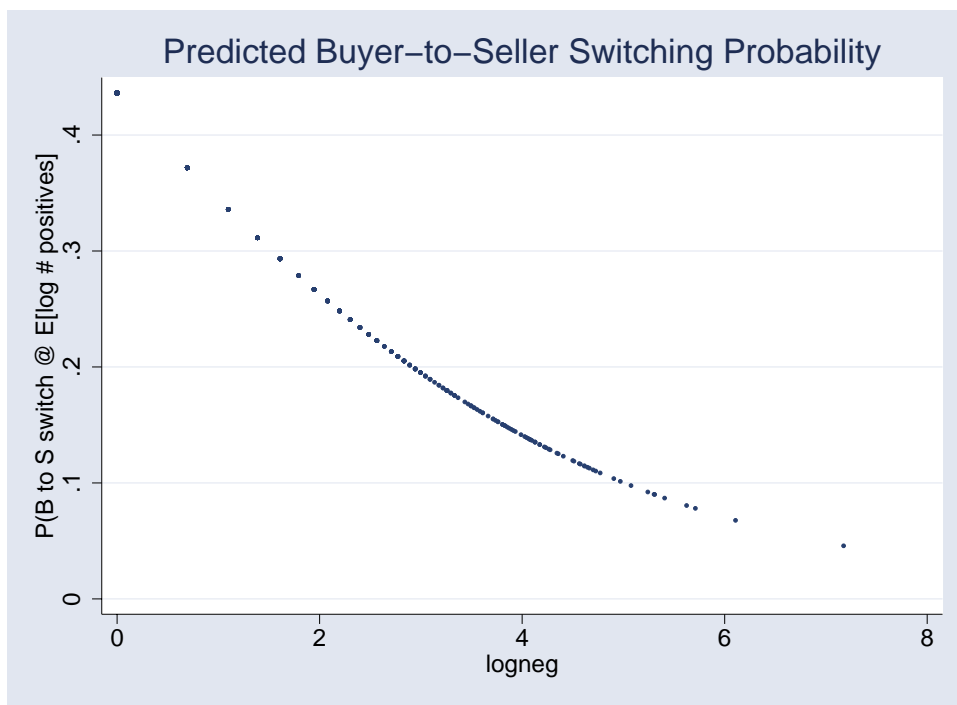
Figure 6: Buyer to seller switching probability.

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.[39]

In this section, we first present a theoretical framework that we think fits the empirical observations relatively well. We then consider other alternative models and discuss their advantages and disadvantages.

■ **A model of eBay reputation dynamics.** Our proposed theoretical framework is in the tradition of Kreps, Milgrom, Roberts and Wilson's (1982) reputation model. Following in part Diamond (1989), we adapt the model to consider some of the specific features of eBay.

Consider a seller living for T periods and a series of one-period-lived buyers. In each period, a transaction takes place, the outcome of which can be either positive or negative. Whichever is the case, the transaction's outcome is truthfully reported by the buyer and as such becomes public knowledge. So, in each period a buyer has access to the full history of the seller's transactions.

Buyers receive a utility 1 when the transaction's outcome is positive and zero when it is negative. Buyers are assumed to be risk neutral, and so their willingness to pay is given by their belief that the transaction's outcome will be positive.

The probability that a transaction's outcome is positive is a function of seller type, θ, as well as seller effort. For simplicity, suppose the seller can only be one of two types. A good seller produces a positive transaction with probability close to one regardless of effort. A bad seller, in turn, has a positive transaction with probability α if it makes effort (and a cost e) and $\beta < \alpha$ if it does not. So, if buyers believe the seller is a good type with probability μ; and if they believe that, the seller being a bad type, high effort is chosen with probability ρ; then the buyer is willing to pay

$$v(\mu, \rho) = \mu + (1 - \mu)\left(\rho\alpha + (1 - \rho)\beta\right).$$

Assume that the seller's payoff is an increasing function of the buyer's willingness to pay (below we discuss this assumption). In the Appendix, we show that, if the prior that seller is good, μ_0, is sufficiently high; and if the value

[39]See Cabral (2005) for further discussion of these two alternative approaches.

of the discount factor is intermediate; then there is a unique Markov equilibrium. The equilibrium strategy for a bad seller is to choose high effort while its record is perfect and switch to low effort once the first negative feedback is received. The intuition is that, while the seller's record is perfect, a bad seller type has a high incentive to pool with the good seller type. Therefore, the bad seller type makes high effort. Once a first negative feedback is received, however, the marginal benefit from high effort significantly decreases, as buyers know almost for sure that the seller is bad. Moreover, during the perfect record phase, buyers keep increasing their belief the seller is good, that is, μ is increasing in the length of the (perfect) record.

The assumption that the discount factor has an intermediate value is important. If the discount factor is too small, then there exists no equilibrium where the low type tries to establish a reputation, that is, effort is always low. If, on the other hand, the discount factor is very close to one, then a folk-theorem-type result ensues, that is, there exist many different equilibria.

By contrast, various other simplifying assumptions are not important for our main qualitative results to hold. In particular, the equilibrium we consider is not knife-edged, that is, it is robust to a variety of perturbations. For example, we assume buyers correctly report the outcome of each transaction, but we could have assumed some probability that a reporting mistake takes place. Also, we assume only two types of seller, but could have considered a wider set of types.

Finally, we make the assumption that the seller's payoff is an increasing function of the buyer's willingness to pay. In the Appendix, we present a simple model of endogenous entry into an auction that has this feature. One important feature of this model is that, not only is the seller's payoff increasing in his reputation, but so is the probability of a sale.

In summary, our model makes predictions that seems to square with key empirical observations in our dataset. A seller with a perfect record puts a lot of effort into maintaining that record. Once a negative is received, the marginal return to investing in reputation decreases dramatically. As a result, the seller (if a low type) decreases effort. Lower effort leads to lower quality transactions; and this in turn leads to a higher likelihood of negative transactions (and negative feedback). In this way, our model is consistent with the following observations: (a) increasing sales rate during the phase when a seller's record is perfect; (b) decrease in price and sales rate following the first negative; (c) increase in the frequency of negative feedback ensuing the first negative feedback. Although we don't observe a decrease in price following the first

negative (due to lack of data) all other predictions are borne out by our panel.

In the remainder of this section, we consider alternative theoretical models and the extent to which they fit the stylized facts uncovered in Section 4.

■ **Pure moral hazard.** Klein and Leffler (1981) and others have considered bootstrap equilibria whereby buyers trust sellers will provide high quality so long as they have done it in the past.[40] Once a low quality sale is observed, buyers "punish" the seller by lowering their expectations regarding quality and accordingly lowering their demand.[41]

A pure moral hazard model does a good job at explaining the breakdown of sales after the first negative. In fact, once a low quality sale is observed, buyers "punish" the seller by lowering their expectations regarding quality. The model also explains the increase in frequency of negative feedback following the first negative feedback. Knowing that buyers' expectations are low, the seller accordingly makes low effort; and as a result negative feedback occurs more often.

The pure moral hazard model has two potential problems (which are related). First, since the cost of creating an identity on eBay is very low, one would expect sellers to exit following the first negative feedback and start selling under a different name.[42] Second, buyers' willingness to pay should be the same before the first negative feedback is received, presumably leading to a constant sales rate before the first negative arrives.

The issue of free entry can be resolved if there is a significant cost of entering and starting a new reputation. Klein and Leffler (1981) and Shapiro (1983) propose money burning in the form of advertising. But advertising does not play a big role on eBay. Alternatively, low prices can serve as a way to burn money. We don't have data on prices, but research by Resnik et al. (2003) suggests that new sellers receive a significant negative price premium. But on eBay prices are typically set by buyers, who submit bids, not by sellers.[43] It would be difficult for buyers to coordinate on setting low bids for an object

---

[40]See Friedman (1971), Telser (1980) for earlier related work.

[41]Klein and Leffler (1981) assume quality is a deterministic function of effort, and that quality is perfectly observable, though ex-post, by buyers. Dellarocas (2003), following earlier work on collusion by Green and Porter (1984) and Porter (1983) on the case of imperfect observability, suggests a stationary mechanism where poor performance is "punished" by buyers for a period of time.

[42]On the issue of changing names at low cost, see Friedman and Resnick (2001).

[43]In fact, sellers can set minimum bids and "buy now" options. In the limit, a very low "buy now" option would essentially amount to a low posted price. Still, we believe it is fair to say that prices are largely determined by buyers.

they know is worth a high bid. So, in order for buyers to bid less when the seller is new, it must be that they expect lower quality. In a pure moral hazard model, this means that a new seller must start with low effort and gradually increase it over time. We should therefore observe a non-monotonic path in negative feedback hazard rates: first decreasing, then increasing. Our data shows no decrease in hazard rates, though this may result from sample bias.[44]

A third way of creating an entry cost if for new sellers to start with a low sales rate and then gradually increase it over time. This is similar to the idea of "building trust" (see Ghosh and Ray, 1996; Kranton, 1996; Datta, 1997; Watson, 1999 and 2002; Ray, 2002). In an efficient equilibrium, the seller is indifferent in each period between offering high quality products or cheat and start a new reputation. This is consistent with the data: initially increasing output, reduction in sales rate and higher likelihood of exit following low quality events. The problem with this view is that it requires a tremendous amount of coordination and monitoring by buyers.

A fourth way of creating an entry cost if for new sellers to start off as buyers. In fact, our empirical evidence shows that this is indeed the case. Suppose that buyers lose money by making purchases (including the opportunity cost of doing it). Suppose that buyers don't buy from a seller until the latter paid his "dues" as a new agent. Then restarting a reputation implies a significant cost. One problem with this view is, again, that it requires a high level of monitoring on the buyers' side. In fact anecdotal evidence suggests that buyers do not distinguish in a seller's history trades as a buyer and trades as a seller.

■ **Word of mouth effects.**   In a recent elegant model, Fishman and Rob (2004) consider an infinite series of one-period lived buyers with word-of-mouth effects: with some probability, a new buyer meets an old buyer and learns about the latter's recent quality experience. The outcome of a transaction is random and positively related to the seller's investment. If a buyer has a good experience with the seller, then the new buyer visits the same seller; if the old buyer had a bad experience, then the new buyer searches a seller randomly. In this context, a seller with a series of good transaction outcomes increases his customer base over time (through word of mouth) *and* increases his incentives to invest in quality. In fact, a low quality transaction breaks down the word of mouth chain and puts the seller back in square zero.

Fishman and Rob's (2004) model is consistent with our first two empiri-

---

[44]A simple non-parametric analysis suggests a distribution of timing of first negative close to log-normal, certainly not bimodal as would be implied by a non-monotonic hazard rate.

cal facts: a seller's transactions rate increases with the length of his (perfect) record and drastically declines upon receiving a first negative feedback. Notice the model is also consistent with an increasing price: since the seller's incentives to invest in quality are increasing in size, which is correlated to age, sale price starts off at a low level and gradually increases. One important difference between Fishman and Rob's (2004) theoretical model and the reality of eBay is that word-of-mouth effects don't seem very important on the latter. However, one can presumably think of a variation of their model that is closer to the institutional details of eBay and features the same basic intuition.

■ **Pure adverse selection.** Consider now the case of pure adverse selection. Suppose that a seller's type at time t, θ_t, is governed by a dynamic stochastic process.[45] In particular, suppose that a seller starts off with a high value of θ (good type) and that, with some probability, type switches to low permanently. For example, the seller initially lives close to a UPS shipping station and for some exogenous reason has to move to a location that is farther away from a UPS shipping station. Such a shift in seller's type would lead not only to the first negative but also to an increased frequency of negative feedback, consistently with the results in Section 4.3.

The model introduced at the beginning of this section (and formally developed in the the Appendix), implies that the likelihood an item is sold is an increasing function of seller reputation.[46] Then simple Bayesian updating implies a series of results consistent with the empirical evidence: a seller starts off selling few products and at a low price; conditional on keeping a good record, a seller increases both the sales rate and price. Once a negative feedback is received, both price and sales rate drop. Moreover, subsequent negatives bring no new information, and so we would expect, as the data suggests, that the sales rate does not change any further. The rate of negative feedback arrival, however, increases, since the seller is now of low type.

One problem with the pure adverse selection story is that it would require that shifts from high type to low type be absorbing. In fact, our results show the frequency of negative feedback arrival increases in a permanent way following the first instance of negative feedback. We do not observe instances

---

[45]The literature on firm growth and industry evolution frequently considers the possibility of firm efficiency evolving according to a Markov process. See Hopenhayn (1992), Ericson and Pakes (1995). More recently, some reputation models have explicitly considered the possibility of changing types. See Mailath and Samuelson (1998), Phelan (2001).

[46]The model we present features a buyer's decision to bid on a product, paying a cost for doing so. Alternatively, we could consider a seller cost of putting an object up for bid.

of a decrease in the frequency of negative feedback arrival, as the switch from low θ to high θ would require.

One justification for this asymmetry is that sellers have a fixed amount of effort which they must divide among all units they sell. And thus, as their sales rate increases, effort decreases, leading to the observed increase in negative feedback frequency.[47] One problem with this view is as follows: we observe that, after the initial shock in sales rate following the first negative, the sales rate picks up again as the seller becomes older. However, we do not observe a corresponding increase in the frequency of negative feedback.

■ **Opportunistic exit.** One of the most popular stylized facts regarding the dynamics of reputation is that reputation is slow to build but can be destroyed very rapidly. In a recent paper, Phelan (2001) proposes a specific formal model of this phenomenon. He considers a model with adverse selection and moral hazard. While the model is applied to the issue of government trust, some of its features can be adapted to the case of eBay. Suppose a seller can be of two different types: good or bad. A good seller only produces good transactions; a bad seller produces a good transaction at the cost of exerting effort. Suppose moreover that a seller's type evolves stochastically over time: in each period, a seller changes type with some given probability. For some parameter values, the optimal strategy for a seller is to produce good transaction outcomes—even if the seller is bad and must exert effort. As time goes by, buyers' beliefs are that the seller is good increase—and so does price. At some point, however, a bad seller gains more from "cheating" buyers than by exerting effort. In Phelan's (2001) model, there is no possibility of "exit" by changing one's identity. Exit, however, provides an additional motivation for the type of opportunistic behavior described in Phelan's model.

The evidence from exits at eBay seems broadly consistent with the possibility of opportunistic behavior of this kind. Notice however that, once again, there are two different stories that fit the data. One is adverse selection and moral hazard, in the line of Phelan (2001). Alternatively, pure adverse selection with changing types would also do the job. In other words, the string on negative feedback comments we observe at the end of exiters' lives can be interpreted in two different ways. One is that sellers strategically build up a reputation and then, planning to exit soon, decide to cheat on buyers. Another interpretation is that, with some probability, a seller's type changes to such an extent that (independently of the seller's actions) a series of neg-

---

[47]We are grateful to Steve Tadelis for this interpretation.

ative transactions takes place; and, as a result of such unfortunate sequence of exogenous events, the seller decides to exit. Anecdotal evidence suggests that opportunistic seller behavior is behind many of the cheat-them-and-leave episodes; but more research is required in order to tease out the two alternative interpretations.

■ **Buying a reputation.** One of our more robust findings is that traders typically start off as buyers and then gradually turn into sellers. Anecdotal evidence suggests that, in many cases such as the IBM Thinkpad, sellers obtain their products from manufacturers rather than eBay purchases. Consequently, starting off as a buyer does not seem to be a requirement for the business of being a seller. Anecdotal evidence also suggests that it is easier to receive positive feedback as a buyer than as a seller. In other words, it is easier (and cheaper) to create a good reputation as a buyer than as a seller. Putting these considerations together, we come to an alternative interpretation of our life cycle empirical observation, namely that sellers start off as buyers with the intent of building a reputation that they can then use to sell their product.

The economic theory of reputation does not directly address this issue. However, Tadelis (1999), Mailath and Samuelson (2001), and others have considered the problem of buying names (and the associated reputations). Name trades do not take place at eBay (to the best of our knowledge). Still, the question addressed by Tadelis (1999) and Mailath and Samuelson (2001), "Who wants to buy a reputation?," seems to apply here as well: what type of seller has an incentive to start off by investing (as a buyer) on an initial reputation history? Is it low-type sellers or high-type sellers?

Tadelis (1999) considers two different effects (see Proposition 3). The "reputation maintenance effect" suggests that good types are more likely to buy a good reputation. The idea is that they can more easily maintain that reputation and therefore can get more out of it. Suppose however that only good types buy good reputations. Then a bad type would get a great (signalling) benefit from buying such a reputation. This second effect, which Tadelis (1999) refers to as the "reputation start-up effect," works the other way around. Mailath and Samuelson's (2001) Proposition 4 suggests that the value of a reputation is highest for intermediate seller types. In sum, the theory does not seem to have a very clear prediction regarding the coefficients on the first two rows of Table 10.

6   Concluding remarks

We may briefly summarize our empirical findings as follows: a typical eBay seller starts off his record mainly by buying items. Restricting to sales trades, we observe an increasing sales rate until the first negative is received; and a substantial drop thereafter. After the first negative 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.

In general, it is difficult to distinguish between various theoretical models, including in particular "bootstrap" and "Bayesian" reputation models.[48] However, in the particular case of eBay we believe the Bayesian model does a better job. This still leaves a variety of possibilities, including pure adverse selection or a combination of adverse selection and moral hazard. More data may be required to tease these out.

Regardless of which theoretical model best explains the data, an important conclusion of our paper is that eBay's reputation system gives way to noticeable strategic responses from both buyers and sellers. That is, the mechanism has "bite". Previous research on eBay has highlighted the impact of reputation on consumer demand (see references in the introduction). Our results are broadly consistent with this literature but go one step further by stressing the impact of the feedback mechanism on seller behavior. To the best of our knowledge, the only other documented evidence of a seller's response to a reputation mechanism is Jin and Leslie (2004), who study restaurants' incentives for hygiene. Specifically, they show the incentives are greater in chain restaurants and restaurants frequented by repeat customers, which is consistent with a dynamic reputation story. Jin and Leslie (2004) and in particular Jin and Leslie (2003) present evidence that moral hazard plays an important role in restaurant hygiene: the introduction of hygiene grade cards led to a decrease in hospitalizations which can be attributed to higher effort by restaurants.[49]

---

[48]Both these reputation mechanisms are consistent with a positive correlation between reputation and incentives to invest in reputation, as the data suggests. That is, 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.

[49]Also related is the work by Abbring, Chiappori, and Pinquet (2003), who 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

Obviously, the fact the reputation system has bite 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.

To conclude, we should mention that our theoretical analysis is based on a fundamental assumption, namely that buyers offer feedback in a non-strategic way. A natural next step is thus to study the strategic motives underlying various agents' feedback behavior. This we plan to do in a new empirical project (Cabral, Hortaçsu and Yin, 2005).

---

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. By contrast, in eBay feedback context, our preferred theory predicts that the marginal benefit of effort decreases when an "accident" (negative feedback) is received; and accordingly the rate of future "accidents" goes up, as we empirically find to be the case.

Appendices

A   A simple auction model with endogenous bidder entry

Suppose there are B potential identical bidders, each with valuation v for the product being auctioned. All B potential bidders simultaneously decide wether to enter the auction, paying a cost c if they decide to enter and bid.[50] We assume bidders play the unique symmetric entry equilibrium. In this equilibrium, each bidder enters the auction with probability p, where p is determined by the indifference condition between entering and not entering the auction. Finally, the bidders that decided to enter simultaneously set their bids and payoffs are paid.

There are three relevant possible outcomes of the bidder entry game. If two or more bidders enter, then the seller gets v and each bidder gets zero. If one or zero bidders enter, then the seller gets zero and the bidder (if there is one) gets v.[51] The entry probability p is thus given by the indifference condition $(1-p)^{B-1}\, v = c$, or simply

$$p(v) = 1 - \sqrt[B-1]{\frac{c}{v}}. \tag{1}$$

The seller's expected payoff is given by

$$\pi(v) = \left(1 - B p(v) \left(1 - p(v)\right)^{B-1} - \left(1 - p(v)\right)^{B}\right) v. \tag{2}$$

(The expression in parentheses on the right-hand side is the probability that there is more than one bidder, the only case when the seller makes a profit.) Finally, the probability of a sale is given by

$$\sigma(v) = 1 - \left(1 - p(v)\right)^{B}.$$

Notice that $p(v), \pi(v)$ and $\rho(v)$ are all increasing in v.

---

[50] See Levin and Smith (1994), Bajari and Hortaçsu (2003). On eBay, this is best thought of as an opportunity cost. If a bidder is only interested in buying one object she will avoid bidding for several objects at the same time.

[51] We implicitly assume that the seller's object is perishable. A possible extension is to assume that an unsold object has value v_U to the seller.

B A model with adverse selection and moral hazard

In this appendix, we present a simple model of seller reputation dynamics. This model includes some of the features of the models in Kreps, Milgrom, Roberts and Wilson (1982), Diamond (1989). First, we make the following assumptions regarding buyer behavior:

Assumption 1 *A transaction has two possible outcomes: successful or unsuccessful, with consumer benefit equal to 1 and 0, respectively.*

More generally, we could assume that consumer benefit is given by $\overline{\omega}$ and $\underline{\omega}$, respectively. However, with no loss of generality we assume $\overline{\omega} = 1$ and $\underline{\omega} = 0$. Another possible extension is that the outcome is continuous and the transaction considered successful if the outcome is above some critical value.

Assumption 2 *A successful transaction is reported with probability one as a successful transaction. An unsuccessful transaction is reported with probability one as an unsuccessful transaction.*

All of the relevant results can be extended to the case when there is a small probability of error in feedback or a less than 100% feedback rate. However, the analysis becomes substantially more complicated. A more crucial assumption we need is that the probability and accuracy the feedback be independent of the seller and of the seller's history. Following eBay's terminology, we refer to a successful transaction as a "positive," or simply P; and an unsuccessful transaction as "negative," or simply N.

Assumption 3 *Buyers are risk neutral.*

Given Assumption 1, Assumption 3 implies that willingness to pay is simply the expected probability of a P transaction.

Our model combines elements of adverse selection and moral hazard. It is similar in structure to Diamond's (1989) model of reputation acquisition in credit markets.[52] Although the context in which we apply it is quite different, the basic mechanism is the same. In his model, the informed party is a firm who knows its type and must choose effort level. The uninformed parties

---

[52]Diamond's model, in turn, builds on the earlier work of Kreps, Milgrom, Roberts and Wilson (1982), Kreps and Wilson (1982), Milgrom and Roberts (1982). See also Hölmstrom (1999) for a related model featuring similar dynamics.

are lenders, who must determine the interest rate. In our formulation, the informed party is a seller who knows her type and must choose effort level. The uninformed parties are the buyers, who must determine whether to bid and, if so, how much to bid.

Specifically, we assume that each seller can be of two types. A good seller always produces P transactions. A bad seller produces a P transaction with probability $\alpha < 1$ at an effort cost e or with probability $\beta < \alpha$ at no effort cost. Let μ_0 be the buyers' prior belief that the seller is good. Each seller lives for an infinite number of periods and discounts the future according to the discount factor δ. In each period, the seller auctions one unit with a second price auction with no secret reserve price or minimum bid.[53]

On the buyer's side, we assume that, in each period, there are B potential identical bidders who live for one period. Each bidder has a valuation given by

$$v(\mu, \rho) = \mu + (1 - \mu) \left(\rho\alpha + (1 - \rho)\beta \right), \tag{3}$$

where μ is the posterior belief that the seller is good and ρ is the belief that the seller, being bad, will make an effort to improve transaction quality. Basically, $v(\mu, \rho)$ is the buyers' expected probability of a P transaction: with probability μ, the seller is good, in which case P happens with probability one; with probability $1 - \mu$, the seller is bad, in which case the outcome is P with probability α or β, depending on whether the seller exerts effort (probability ρ) or not (probability $1 - \rho$).

As in Appendix A, we assume potential buyers must pay a cost c in order to bid. We further assume that bidders play the unique symmetric entry equilibrium. The results from Appendix A apply, where v is now given by (3).

We now turn to the characterization of the seller's equilibrium strategy. We do so in the context of the following important assumption, which we will maintain throughout:

Assumption 4 $\quad \frac{e}{\beta e + (\alpha - \beta)\left(\pi(1) - \pi(\beta)\right)} < \delta < \frac{e}{\beta e + (\alpha - \beta)\left(\pi(\alpha) - \pi(\beta)\right)}.$

In words, we assume that the value of the discount factor, δ, is intermediate. A very high value of the δ implies that there is a multiplicity of equilibria. In fact, for δ sufficiently close to one any feasible, individually rational payoff profile is attainable as a perfect Bayesian equilibrium of the game. A very low

---

[53]A straightforward extension is to assume the seller puts an object up for auction at an exogenously given rate, independent of its type.

value of δ, in turn, implies that there is only one equilibrium, one where the (bad) seller never exerts effort.

The following result characterizes a Perfect Bayesian equilibrium of this game. This result is different from Diamond's (1989), who considers a finitely lived seller. However, the basic intuition is the same, namely, the idea that reputation and effort are "correlated" in equilibrium.

Proposition 1 *In a perfect Bayesian equilibrium,*

1. *After the first N, the buyers' willingness to pay decreases.*

2. *After the first N, the seller chooses low effort.*

3. *There exists a t' such that the seller chooses high effort if he has a perfect record longer than t'.*

Proof: Consider first the case when the seller's history includes an N. Bayesian updating implies $\mu = 0$, where μ is the posterior that the seller is good. The only possibility of an equilibrium where the seller chooses high effort is one where an N is punished by never believing the seller will choose high effort again, $\rho = 0$. Such a punishment implies a discounted profit of $\pi(\beta)/(1-\delta)$, where β is the buyer's willingness to pay a bad seller who does not exert effort.

If instead buyers expect the seller to choose high effort, that is $\rho = 1$, then the seller's expected payoff from high and low effort, assuming maximal punishment, is given by

$$
\begin{aligned}
V^H &= \pi(\alpha) - e + \alpha\delta V^H + (1-\alpha)\delta\pi(\beta)/(1-\delta) \\
V^L &= \pi(\alpha) + \beta\delta V^H + (1-\beta)\delta\pi(\beta)/(1-\delta).
\end{aligned}
$$

Straightforward computation shows that the condition $V^L > V^H$ is equivalent to $\delta < \frac{e}{\beta e + (\alpha-\beta)\left(\pi(\alpha)-\pi(\beta)\right)}$. It follows that the only equilibrium following an N is low effort.

Consider now the case of a bad seller with a perfect record. Bayesian updating implies that $\mu > \mu_0 \approx\to 1$. In the limit when $\mu = 1$, the seller's expected payoff from high and low effort is given by

$$
\begin{aligned}
\tilde{V}^H &= \pi(1) - e + \alpha\delta V^H + (1-\alpha)\delta\pi(\beta)/(1-\delta) \\
\tilde{V}^L &= \pi(1) + \beta\delta V^H + (1-\beta)\delta\pi(\beta)/(1-\delta).
\end{aligned}
$$

Straightforward computation shows that the condition $V^H > V^L$ is equivalent to $\delta > \frac{e}{\beta e + (\alpha-\beta)\left(\pi(1)-\pi(\beta)\right)}$.

The above calculations imply that ρ declines as the first N appears. Moreover, Bayesian updating implies that μ drops from a positive value to zero. We thus conclude that v decreases as the first N is given. ∎

We should note that Proposition 1 is not a knife-edged result: following the steps of the proof, one can see that continuity arguments apply if we assume that a good type produces a P with probability γ lower than, but close to, one. In fact, below we consider an extension of the basic model where γ is strictly less than one.

Having said that, we should restate that the result depends crucially on the particular values of δ we consider. If δ is very high, then the folk theorem applies: any equilibrium path that is feasible and individually rational is the result of a perfect Bayesian equilibrium for a high enough value of the discount factor δ. In other words, if the discount factor is high enough, then equilibrium theory has no predictive power. At the other extreme, if δ is very low then there is a unique equilibrium where the seller chooses low effort in every period. Points 1 and 2 in Proposition 1 still hold true, but not Point 3.

The results above have various empirical implications. In particular, they imply that (a) the sales rate decreases after the first negative; (b) the negative feedback arrival rate increases upon the first negative (c) both the sales rate and the frequency of negative feedback remain constant after the first negative.

References

ABBRING, JAPP H., PIERRE-ANDRÉ CHIAPPORI, AND JEAN PINQUET (2003), "Moral Hazard and Dynamic Insurance Data," *Journal of the European Economic Association* **1**, 767–820.

ASCH, SOLOMON (1946), "Forming impressions of personality," *Journal of Social and Abnormal Psychology* **41**, 258–290.

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

BAJARI, PAT, AND ALI HORTAÇSU (2003), "Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay," *Rand Journal of Economics* **34**, 329–356.

BAJARI, PAT, AND ALI HORTAÇSU (2004), "Economic Insight from Internet Auctions," *Journal of Economic Literature* **42**, 457–486.

BAR-ISAAC, HESKI (2002), "Reputation and Survival: Learning in a Dynamic Signalling Model," *Review of Economic Studies* **70**, 1–21.

CABRAL, LUÍS M B (2000), "Stretching Firm and Brand Reputation," *Rand Journal of Economics* **31**, 658-673.

CABRAL, LUÍS M B (2005), "The Economics of Reputation: A Primer," New York University.

CABRAL, LUÍS M B, ALI HORTAÇSU, AND PAI-LING YIN (2003), "The Strategic Dynamics of eBay Buyer and Seller Feedback," Work in Progress (note: title is provisional).

COHEN, ADAM (2002), *The Perfect Store: Inside eBay*, Little Brown and Company.

DATTA, SAIKAT (1997), "Building Trust," London School of Economics, April.

DELLAROCAS, CHRYSANTHOS (2003), "The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms," *Management Science* **49**, 1407–1424.

DELLAROCAS, CHRYSANTHOS (2003), "Efficiency and Robustness of eBay-like Online Reputation Mechanisms in Environments with Moral Hazard," MIT.

DEWAN, S. AND V. HSU (2001), "Trust in Electronic Markets: Price Discovery in Generalist Versus Specialty Online Auctions," mimeo.

DIAMOND, DOUGLAS W (1989), "Reputation Acquisition in Debt Markets," *Journal of Political Economy* **97**, 828–862.

EATON, D. H. (2002), "Valuing Information: Evidence from Guitar Auction on eBay," mimeo.

EDERINGTON, L. H. AND M. DEWALLY (2003), "A Comparison of Reputation, Certification, Warranties, and Information Disclosure as Remedies for Information Asymmetries: Lessons from the On-Line Comic Book Market," mimeo.

ERICSON, RICHARD, AND ARIEL PAKES (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economics Studies* **62**, 53–82.

FRIEDMAN, ERIC J., AND PAUL RESNICK (2001), "The Social Cost of Cheap Pseudonyms," *Journal of Economics and Management Strategy* **10**, 173–199.

FRIEDMAN, JAMES (1971), "A Noncooperative Equilibrium for Supergames," *Review of Economic Studies* **28**, 1–12.

GALE, DOUGLAS, AND ROBERT W ROSENTHAL (1994), "Price and Quality Cycles for Experience Goods," *Rand Journal of Economics* **25**, 590–607.

GHOSH, P., AND DEBRAJ RAY (1996), "Cooperation and Community Interactions without Information Flows," *Review of Economic Studies* **63**, 491–519.

GREEN, ED AND ROBERT PORTER (1984), "Noncooperative Collusion Under Imperfect Price Information," *Econometrica* **52**, 87–100.

HOCH, S., AND Y. HA (1986), "Consumer learning : advertising and the ambiguity of product experience," *Journal of Consumer Research* **13**, 221–233.

HOLMSTRÖM, BENGT (1999), "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies* **66**, 169–182.

HOPENHAYN, HUGO (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica* **60**, 1127–1150.

HOUSER, D. AND J. WOODERS (2003), "Reputation in Auctions: Theory and Evidence from eBay," forthcoming in *Journal of Economics and Management Strategy*.

JIN, GIN Z., AND A. KATO (2004), "Blind Trust Online: Experimental Evidence from Baseball Cards," University of Maryland Working Paper.

JIN, GIN Z, AND PHILIP LESLIE (2004), "Reputational Incentives for Restaurant Hygiene," University of Maryland and Stanford Univrsity.

JOVANOVIC, BOYAN (1982), "Selection and Evolution of Industry," *Econometrica* **50**, 649–670.

KALYANAM, K. AND S. MCINTYRE (2001), "Returns to Reputation in Online Auction Markets," mimeo.

KLEIN, B, AND K LEFFLER (1981), "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy* **89**, 615–641.

KRANTON, RACHEL (1996), "The Formation of Cooperative Relationships," *Journal of Law, Economics and Organization* **12**, 214–233.

KREPS, DAVID M., PAUL MILGROM, JOHN ROBERTS AND ROBERT WILSON (1982), "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory* **27**, 245–252.

KREPS, DAVID M., AND ROBERT WILSON (1982), "Reputation and Imperfect Information," *Journal of Economic Theory* **27**, 253–279.

LEVIN, DAN, AND JAMES SMITH (1994), "Equilibrium in Auctions with Entry," *American Economic Review* **84**, 585–599.

LIVINGSTON, J. (2002), "How Valuable is a Good Reputation? A Sample Selection Model of Internet Auctions," mimeo.

LUCKING-REILEY, D., D. BRYAN, N. PRASAD, N., AND D. REEVES (2000), "Pennies from eBay: the Determinants of Price in Online Auctions," University of Arizona Working Paper.

MAILATH, GEORGE J, AND LARRY SAMUELSON (2001), "Who Wants a Good Reputation?," *Review of Economic Studies* **68**, 415–441.

MAILATH, GEORGE J, AND LARRY SAMUELSON (1998), "Your Reputation Is Who You're Not, Not Who You'd Like To Be," University of Pennsylvania and University of Wisconsin.

MCDONALD, C. G., AND V. C. SLAWSON (2002), "Reputation in an Internet Auction Market," *Economic Inquiry* **40**, 633-650.

MELNIK, M. I. AND J. ALM (2002), "Does A Seller's eCommerce Reputation Matter? Evidence from eBay Auctions," *Journal of Industrial Economics* **50**, 337-350.

MILGROM, PAUL, AND JOHN ROBERTS (1982), "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory* **27**, 280–312.

OCKENFELS, AXEL, AND ALVIN E. ROTH (2003), "Late and Multiple Bidding in Second Price Internet Auctions: Theory and Evidence Concerning Different Rules for Ending an Auction," forthcoming in *Games and Economic Behavior.*

PHELAN, CHRISTOPHER (2001), "Public Trust and Government Betrayal," forthcoming in *Journal of Economic Theory.*

PORTER, ROBERT H (1983), "Optimal Cartel Trigger Price Strategies," *Journal of Economic Theory* **29**, 313–338.

RAY, DEBRAJ (2002), "The Time Structure of Self-Enforcing Agreements," *Econometrica* **70**, 547–582.

RESNICK, PAUL, AND RICHARD ZECKHAUSER (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, Amsterdam: Elsevier.

RESNICK, PAUL, RICHARD ZECKHAUSER, J. SWANSON, AND K. LOCKWOOD. (2003), "The Value of Reputation on eBay: A Controlled Experiment," Harvard Kennedy School Working Paper.

ROTH, ALVIN E., AND AXEL OCKENFELS (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**, 1093–1103.

SHAPIRO, CARL (1983), "Premiums for High Quality Products as Rents to Reputation," *Quarterly Journal of Economics* **98**, 659–680.

SNYDER AND CANTO (1979), "Testing Hypotheses about other people: The use of historical knowledge," *in Journal of Experimental Social Psychology* **15**, 330–342.

TADELIS, S. (1999), "What's in a Name? Reputation as a Tradeable Asset," *American Economic Review* **89**, 548–563.

TADELIS, S. (2002), "The Market for Reputations as an Incentive Mechanism," *Journal of Political Economy* **92**, 854–882.

TELSER, L G (1980), "A Theory of Self-enforcing Agreements," *Journal of Business* **53**, 27–44.

WATSON, JOEL (1999), "Starting Small and Renegotiation," *Journal of Economic Theory* **85**, 52–90.

WATSON, JOEL (2002), "Starting Small and Commitment," *Games and Economic Behavior* **38**, 176–199.