The Dynamics of Seller Reputation:
Appendix

Luís Cabral  
*New York University and CEPR

Ali Hortaçsu*  
University of Chicago and NBER  

July 2008

Abstract

This appendix complements our paper “The Dynamics of Seller Reputation: Evidence from eBay.”
1 Introduction

This appendix complements our paper, Cabral and Hortaçsu (2008). In that paper, we provide a series of empirical observations on the dynamics of eBay sellers’ reputations. We briefly discuss some possible interpretations of the results. In particular, we propose that a model combining adverse selection and moral hazard best explains some of the key features of the data. In the present appendix, we present a detailed description of such model. We also consider alternative theoretical model from the economics literature on reputation and discuss the extent to which they fit the data.

2 A model of eBay seller reputation

In this section, we present a theoretical framework that we think fits our empirical observations relatively well. The section is organized as follows. In Section 2.1, we present a general discussion of the model and its properties. Next, in Section 2.2, we present a simple model of auction entry which we use as a module in our general model. Finally, Section 2.3 develops the model in greater detail and presents some testable results (which are consistent with the evidence in Cabral and Hortaçsu, 2008).

2.1 General description

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, $\theta$, 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 $\alpha$ 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 $\mu$; and if they believe that, the seller being a bad type, high effort is chosen with probability $\rho$; 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 Section 2.3, we show that, if the prior that seller is good, $\mu_0$, is sufficiently high; and if the value 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, $\mu$ 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 Section 2.2, 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 Cabral and Hortaçsu (2008). 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 in Cabral and Hortaçsu (2008) 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.

2.2 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 whether to enter the auction, paying a cost $c$ if they decide to enter and bid.\(^1\) 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 $v$.\(^2\) The entry probability $p$ is thus given by the indifference condition $(1 - p)^{B-1} v = c$, or simply

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

The seller’s expected payoff is given by

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

\(^1\)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.

\(^2\)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.
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 $v$.

2.3 A model with adverse selection and moral hazard

We now present in greater detail our 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 $\omega$ and $\omega'$, respectively. However, with no loss of generality we assume $\omega = 1$ and $\omega' = 0$. 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 of 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. 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 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 $\mu_0$ 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 $\delta$. In each period, the seller auctions one unit with a second price auction with no secret reserve price or minimum bid.

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

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

where $\mu$ is the posterior belief that the seller is good and $\rho$ 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 $\mu$, 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 $\alpha$ or $\beta$, depending on whether the seller exerts effort (probability $\rho$) or not (probability $1 - \rho$).

As in Section 2.2, 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 Section 2.2 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**

$$\frac{e}{\beta \varepsilon + (\alpha - \beta)(\pi(1) - \pi(\beta))} < \delta < \frac{e}{\beta \varepsilon + (\alpha - \beta)(\pi(\alpha) - \pi(\beta))}.$$
In words, we assume that the value of the discount factor, $\delta$, is intermediate. A very high value of the $\delta$ implies that there is a multiplicity of equilibria. In fact, for $\delta$ sufficiently close to one any feasible, individually rational payoff profile is attainable as a perfect Bayesian equilibrium of the game. A very low value of $\delta$, 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 $\mu$ 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 $\beta$ 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

\[
V^H = \pi(\alpha) - e + \alpha \delta V^H + (1 - \alpha) \delta \pi(\beta)/(1 - \delta)
\]
\[
V^L = \pi(\alpha) + \beta \delta V^H + (1 - \beta) \delta \pi(\beta)/(1 - \delta).
\]

Straightforward computation shows that the condition $V^L > V^H$ is equivalent to $\delta < \frac{e}{\beta e + (\alpha - \beta) (\pi(\alpha) - \pi(\beta))}$. 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 $\mu > \mu_0 \approx 1$. In the limit when $\mu = 1$, the seller’s expected payoff from high and low effort is given by

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

The above calculations imply that $\rho$ declines as the first $N$ appears. Moreover, Bayesian updating implies that $\mu$ 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 $\gamma$ lower than, but close to, one. In fact, below we consider an extension of the basic model where $\gamma$ is strictly less than one.

Having said that, we should restate that the result depends crucially on the particular values of $\delta$ we consider. If $\delta$ 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 $\delta$. In other words, if the discount factor is high enough, then equilibrium theory has no predictive power. At the other extreme, if $\delta$ 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.

In the next section, we consider alternative theoretical models and the extent to which they fit the stylized facts uncovered in Cabral and Hortacsu (2008).

3 Alternative theories

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, 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.\(^5\)

**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.\(^6\) Once a low quality sale is observed, buyers “punish” the seller by lowering their expectations regarding quality and accordingly lowering their demand.\(^7\)

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.\(^8\) 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.\(^9\)

\(^5\)See Cabral (2005) for further discussion of these two alternative approaches.
\(^6\)See Friedman (1971), Telser (1980) for earlier related work.
\(^7\)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.
\(^8\)On the issue of changing names at low cost, see Friedman and Resnick (2001).
\(^9\)In fact, sellers can set minimum bids and “buy now” options. In the limit, a very low
would be difficult for buyers to coordinate on setting low bids for an object 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.

A third way of creating an entry cost is 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 is 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

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.

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.
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 empirical 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 $t$, $\theta_t$, is governed by a dynamic stochastic process.\textsuperscript{11} In particular, suppose that a seller starts off with a high value of $\theta$ (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 ??.

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.\textsuperscript{12} 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

\textsuperscript{11}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).

\textsuperscript{12}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.
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 of a decrease in the frequency of negative feedback arrival, as the switch from low $\theta$ to high $\theta$ 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.\textsuperscript{13} 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.

\section*{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

\textsuperscript{13}We are grateful to Steve Tadelis for this interpretation.
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 negative 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.
References


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


