SELLER RATING, PRICE, AND DEFAULT IN ONLINE AUCTIONS

NORRIS BRUCE, ERNAN HARUVY, AND RAM RAO

Whereas in brick and mortar stores quality and seller reliability can be effectively communicated through a variety of quality cues, in the online world different mechanisms help sellers establish reputation and trust. Among the commonly used mechanisms by infomediaries that screen and rate sellers are customer-feedback reputation mechanisms where buyers provide seller ratings and feedback. The success of such mechanisms and their impact on prices in electronic markets are described and empirically investigated in this work. Specifically, we investigate both theoretically and empirically (1) whether seller ratings affect consumer valuations, particularly in the presence of insurance; (2) whether seller ratings are indicative of future default likelihood; and (3) whether a seller who is terminal (about to go out of business) is more likely to default prior to exiting.

© 2004 Wiley Periodicals, Inc. and Direct Marketing Educational Foundation, Inc.

JOURNAL OF INTERACTIVE MARKETING VOLUME 18 / NUMBER 4 / AUTUMN 2004
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/dir.20021
INTRODUCTION

The need to establish reputation and trust is critical for retailers without known brand names. In fact, it is possible that reputation matters more on the Internet than in comparable brick and mortar businesses. A firm with a plush office or expensive décor might be able to signal quality and thus extract higher revenues than a competitor with a lower quality that is unable to afford a lavish office. On the Internet, however, the most elaborate of Web sites can be imitated at low cost. As such, consumers are and should be wary of quality cues and the establishment of trust is far more elusive.

The issue of trust online has received considerable attention in recent years. Hoffman, Novak, and Peralta (1999a, 1999b) found that customers expect Internet exchange to be based on a relationship of trust and cooperation. Urban and Sultan (1999) found that abundant information as well as virtual advisors with helpful algorithms to assist in decision making can go a long way toward establishing trust.

The need for trust applies to both price and nonprice attributes. On nonprice attributes, the buyer's trust centers on the items' quality and delivery as represented by the seller. On price, the buyer seeks reassurance that the seller's price is competitive. In the absence of a better alternative, a consumer might engage in ex ante investigation prior to each transaction. That is, the consumer could go out and sample the suppliers' prices (e.g., Stigler, 1961; Diamond, 1971; MacMinn, 1980) as well as investigate their credibility on nonprice attributes via consumer reports and other information sources. Alternatively, infomediaries can provide that service for the consumer (Chen, Iyer, & Padmanabhan, 2002). Infomediaries can provide information about prices (Salop & Stiglitz, 1977) as well as vouch for quality or provide implicit or explicit insurance to customers as an attractive alternative to costly research (Allen & Gale, 1999).

Infomediaries can provide both price comparisons and quality assessments. For quality assessments, some infomediaries provide their own research; others rely on customer feedback. Research suggests that buyer feedback can reduce customers' perceptions of risk and lead them to trust the seller (e.g., Kollock, 1999). BizRate.com, Epinions.com, and Imandi.com, for example, have adopted a customer-rating feature with customer-rated choices of online products and stores to buy from. In addition to these feedback-based quality assurances, these sites also provide price comparisons, sorting by price and finding the lowest price for each product. Sellers pay for this valuable service and vie for good customer feedback and price ranking.

Though feedback-based quality ratings are gaining in prominence, nowhere are they more visible than in online auctions. With online auction sites, the auction site itself serves as the infomediary. Online auction sites—aware of the trust and reputation barriers—have established reputation-building mechanisms for sellers. Such mechanisms allow buyers to post positive, neutral, or negative comments following a completed transaction. A positive feedback to the seller is an indication by the buyer that the item has arrived in a timely manner, was undamaged, and was correctly represented by the seller. Negative feedback predominantly includes one of two types of complaints: (1) The seller has sent a lower quality item than described, or (2) no item was sent at all. The three major Internet auction firms, eBay, Yahoo, and Amazon have adopted this feedback mechanism, with minimal cosmetic variations. Though these firms spend some resources on policing auctions, they mainly rely on feedback-based community policing (e.g., Bradley, 2001). As such, feedback ratings are buyers' main source of information on seller reliability.

The widespread use of customer feedback and its apparent success is puzzling: Buyers have no discernible incentive to furnish any feedback, and have no nonnegligible incentive to avoid giving negative feedback since the seller could retaliate. Retaliation could be carried out by not replacing the defective or misrepresented product. The seller is also able to leave a negative feedback in return. A negative feedback could harm the buyer if he or she ever attempts to sell items using the same login name. The academic literature has been trying to confirm or refute the effectiveness of the feedback mechanism with limited success. The literature offers mixed evidence on whether seller rating has an effect on price. Some find a clear effect for both positive and negative
feedback, whereas others find mixed or no effect. For example, Kauffman and Wood (2004) found that feedback ratings had no effect on the prices of collectible coins traded in online auctions. Melnick and Alm (2002), on the other hand, found that both positive and negative feedback had an impact on the prices of collectible coins in auctions. Bajari and Hortacsu (2000), however, found that only positive feedback affected the price of collectible coins. Finally, Lucking-Reiley, Bryan, Prasad, and Reeves (2000) found that only negative feedback had an effect. To the uninitiated this may seem like a futile task when four careful groups of researchers arrive at four different results on virtually the same commodity. However, with commodities reputation may not have the same impact as with more complex and expensive merchandise. In fact, auction studies with Pentium chips (Houser & Wooders, 2000), Palm Pilots and PDAs (Kalyanam & McIntyre, 2001), and used computer monitors and printers (Lee, Im, & Lee, 2000) find significant impact of feedback on price. On the other hand, the price of electric guitars (Eaton, 2002) was not found to be significantly affected by feedback, nor were new computer monitors and printers (Lee, Im, & Lee, 2000), or MP3 players (Resnick & Zeckhauser, 2002), all of which are expensive and complex. A detailed survey of the literature is provided in Resnick, Zeckhauser, Swanson, and Lockwood (2002).

The volume of mixed evidence leaves the main set of questions unanswered: (1) Is feedback an effective signal of future quality? (2) If so, should a buyer make use of this signal? If the infomediary provides an implicit or explicit insurance, it is not at all clear that the buyer needs to worry about customer feedback. Insurance is an important component of auction infomediaries. A buyer can often file an insurance claim in the case of undelivered items or fraud. PayPal or PayDirect in many cases are able to recover the payment directly from the seller. In addition, the buyer could file a fraud complaint form with eBay or Yahoo and recover his costs through the insurance program. With all these measures, why should the buyer be concerned with the seller’s feedback?

First, complaints have both lower and upper bounds. Items such as books and CDs, commonly traded on eBay, are not covered by eBay’s insurance policy (which has a $25 processing cost). Second, the most the person can recover under the insurance policy is $200 ($400 if he had used PayPal). Furthermore, complaints are time sensitive. A complaint with PayPal cannot be filed more than 30 days after payment. A complaint with eBay cannot be filed less than 30 days after end of auction. Ironically, to recover insurance from PayPal, one must first recover from eBay. The aggrieved party should be careful to navigate around these complex deadlines and restrictions.

This paper provides explanations and elaborates on theories of intermediation that explain the role of the infomediary on the critical dimension of providing information or signals on sellers’ qualities. We lay out the issues of importance in this area and present some evidence that offers insight into the value of feedback systems as a signal of seller’s reliability.

We offer a plausible explanation for the mixed results on the effect of reputation on price in research to date. Specifically we argue that the relationship between price and reputation is more likely to be significant for low price items due to the nonnegligible costs of obtaining refunds for unsatisfactory transactions. We provide empirical evidence on the relationship between the average price in the category and the impact of feedback.

We then show that feedback is indeed an effective signal when the relative frequency of default is the dependent variable. We also assess whether the past feedback history of the seller is a significant indicator of future failure. Finally, we discuss the implications of feedback mechanisms on price both with respect to recommendations to sellers and directions for future research.

**DOES (AND SHOULD) SELLER REPUTATION AFFECT PRICE IN THE PRESENCE OF INSURANCE?**

As remarked in the Introduction, the empirical question of whether reputation should and does affect price has been extensively studied in the extant literature. However, the issue has not received a rigorous theoretical treatment, and the evidence has been
mixed. We propose that (1) seller rating significantly impacts the price and that (2) the relative impact shrinks the higher the value of the product. The model that follows presents the rationale for these assertions. Following the presentation of the model we present some empirical evidence that supports these assertions.

**The Model**

The model allows for insurance of the kind eBay and Yahoo offer. That is, the insurance has a deductible. Another critical assumption of the model is that the cost of complaining and pursuing remedies is non-negligible. As such, small inexpensive items that are not delivered or are faulty will not be compensated for, whereas default on expensive items will always result in the buyer actively seeking compensation. If the buyer knows in advance that he is likely to seek compensation on expensive items but not on inexpensive items, the seller ratings become relatively more important for inexpensive items.

The model is one of a two-stage game. In stage 1, the bidder observes the signal $S$ (seller rating) and the promised quality $Q$ and decides on a bid. In stage 2, the bidder either wins the item or does not win the item. A winning bidder pays and observes the product quality. If the product quality is lower than promised, he may then file a dispute which will result in compensation, minus a deductible at a later point in time. The dispute process is not costless and the award is delayed and hence discounted.

**Parameter Definitions:**

$Q =$ quality of product,

$V(Q) =$ product valuation of quality $Q$,

$S =$ seller rating,

$P =$ bid amount.

We assume product valuation to be linear in quality $Q$, so that $V(Q) = Q$ and $V(Q, S) = Q - \varepsilon/S$, where $Q$ is the promised quality. The deviation of true quality from the promised quality, $\varepsilon/S$, is distributed $\varepsilon/S \sim U[0, 2mQ/S]$ and $Q \geq 0$ (i.e., $S \geq 2m$). Thus, $E(Q) = \bar{Q}(1 - m/S)$. We denote the distribution of $\varepsilon$ by its probability density function $g(\varepsilon) = 1/2m\bar{Q}$.

Due to the deductible, the cost of dispute, and the discounting, disputes are costly. Thus, if a buyer disputes and if the dispute is successful, the buyer receives $\delta(P - d) - c$, where $\delta$ is the discounting of future payoff parameter, $c$ is the cost of dispute (in monetary equivalent of the time lost and energy expended); and $d$ is a deductible. Both $c$ and $d$ are independent of $P$.

In stage 2, a buyer disputes iff $\delta(P - d) - c > Q > 0$ and does not dispute otherwise. In stage 1, the buyer determines a bid for the product. A standard proof in second price auction format is that the bidder bids his true valuation of the object being auctioned. If the buyer is not expected to dispute in stage 2, then the buyer should bid $P = \bar{Q}(1 - m/S)$. If $\delta = 1$ and $c = d = 0$, then the buyer would always dispute in stage 2 and would bid $P = \bar{Q}$. Otherwise, the buyer bids the expected value. In this case, the bid price is

$$P = \frac{S(\theta^2 + \bar{Q}^2) - 2\bar{Q}\theta(S - 2m)}{4m\bar{Q}},$$

(1)

where $\theta = \delta(P - d) - c$. It is straightforward to see that price is increasing in seller quality. We refer to this finding as claim 1 (see proof of claim 1 in the Appendix for more detail). The insight that follows is that a buyer may choose not to dispute an item that was misrepresented. For low priced items, buyers will almost never incur the expense of pursuing a defaulting seller. In addition, the deductible may make such an attempt futile from the start. However, for high priced items, the buyers will always seek remedy from a defaulting seller, making seller ratings less critical. This observation leads us to the first hypothesis.

**H1:** The impact of seller ratings on bid prices will decline as price increases.

Hypothesis 1 implies that seller ratings are more critical to price determination for low-priced items. A mathematical derivation to back up this result is included in the Appendix. If this hypothesis is confirmed, it would mean that more costly items are less likely to be sensitive to seller ratings. In the next section, we provide some empirical evidence that appears to confirm this assertion.
Empirical Evidence

The data were collected from publicly available completed eBay auctions in May 2002. Auctions for four items were selected:

1. Laptop: IBM Thinkpad laptops (33 items, 162 total bids)
2. Dell Dimension PCs (82 items, 625 total bids)
3. Shrek DVD (273 items, 1035 total bids)
4. Harry Potter books (82 items, 282 total bids)

Note that the laptop and PC are large expensive items whereas Shrek DVDs and Harry Potter books are relatively inexpensive. For each auction, the following characteristics were recorded:

1. seller’s ID (eBay name)
2. seller rating
3. final sale price
4. shipping fee
5. seller’s reservation price (lowest bid)
6. number of bids
7. number of items sold (usually one)
8. bid time
9. auction duration
10. additional item-specific characteristics
   (new/used, edition, model, features).

Consider the following notation for the empirical model:

\[ b_{jt} : \text{Bid at time } t \text{ in auction } j, \]
\[ Z_{jt} : \text{Explanatory variables}, \]
\[ \varepsilon_{jt} \sim N(0, \sigma^2). \]

Explanatory variables include variables that measure whether the version is new or used, edition or model variables, and accessories. The number of bids is not used as an explanatory variable, since it is an endogenous indicator of reputation and not an independent variable (Lucking-Reiley et al., 2000; Resnick et al., 2002). Most importantly, the explanatory variable of interest is the seller rating. The coefficient for seller rating is multiplied by the natural log of seller rating. Since bids must be positive, we assume that the relationship between bids, \( b_{jt} \), and the explanatory variables, \( Z_{jt} \), is of the form:

\[ b_{jt} = \varepsilon_{jt} \exp(\beta Z_{jt}). \]  

Taking logs, the relationship becomes

\[ \hat{b}_{jt} = \ln(b_{jt}) = \beta Z_{jt} + \varepsilon_{jt}. \]  

We estimate equation (2) by maximum likelihood (ML). Note that the contribution of the \( j \)th auction to the likelihood is

\[ L_j = \prod_{t=1}^{T_j} \frac{f(\hat{b}_{jt})}{1 - F(\hat{b}_{jt})}, \]  

where \( \hat{b}_{jt} \) is the highest current bid in auction \( j \) at time \( t \). Since buyers will only bid if they can make an offer that exceeds the current highest bid, we divide the likelihood by \( (1 - F(\hat{b}_{jt})) \). Thus, \( (1 - F(\hat{b}_{jt})) \) is the probability that the buyer’s bid, \( \hat{b}_{jt} \), exceeds the current highest bid \( \hat{b}_{jt} \), where \( F \) is the normal cumulative distribution function. Hence, all observed bids are drawn from a truncated normal distribution. Table 1 displays ML estimates and significance levels.

We can see that the PC and laptop estimates for seller rating are not significant, whereas the corresponding estimates on Shrek DVDs and Harry Potter books are highly significant. This is consistent with what the theory prescribes. However, notice that the interaction term between the “used” dummy and seller rating for IBM laptops is significant. That is, the prices for used laptops are affected by seller rating (but not new laptops). Conceivably, this is because used laptops are less of a commodity than new laptops, and so ratings may convey more information. Note that this is consistent with the finding by Lee et al. (2000).

As another check, we examine rating elasticity of price (% change in price over % change in seller rating)\(^1\)

\(^1\) Since rating is transformed to \( \ln(R) \) in the regression, an alternative measure of rating elasticity might look at the percentage change in price over the percentage change in \( \ln(R) \). We verified that the results remain qualitatively the same with the alternative measure.
using the reported parameter estimates. We define rating elasticity of price as $E = (\partial b / \partial R)b$. Using the parameter estimates from Table 1, we obtain the rating elasticity estimates for each level of seller rating and each product. The estimates are given in Table 2. The first four rows show the rating elasticity for each category for seller ratings of 10, 100, 1,000, and 10,000, holding price, and all other explanatory variables at their average levels in the category. The next four rows show the rating elasticity for the average seller rating in each category holding price and all other explanatory variables at their average levels: Specifically, the average seller rating is 769 for Potter books, 1,177 for Shrek DVDs, 3,433 for IBM laptops, and 483 for Dell PCs. We see that in each row the highest rating elasticity is observed for the least expensive categories (Potter books and Shrek DVDs).

The empirical findings of this section demonstrated that seller ratings matter. That is, higher ratings

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|}
\hline
\textbf{VARIABLE} & \textbf{COEFFICIENT} & \textbf{STD. ERROR} & \textbf{t-STATISTIC} & \textbf{p-VALUE} \\
\hline
Dell PC & & & & \\
Constant & 6.00 & 0.07 & 85.84 & 0.000 \\
Seller rating & -0.01 & 0.01 & -0.47 & 0.637 \\
XP Dummy & 0.06 & 0.14 & 0.42 & 0.674 \\
Model 4100 & -0.44 & 0.13 & -3.28 & 0.001 \\
Model 4300 & -0.62 & 0.34 & -1.82 & 0.069 \\
Model 4400 & -0.21 & 0.06 & -3.68 & 0.000 \\
Model 8100 & 0.15 & 0.15 & 1.03 & 0.303 \\
$\sigma$ & 0.55 & 0.02 & 25.03 & 0.000 \\
\hline
IBM & & & & \\
Constant & 5.92 & 0.60 & 9.88 & 0.000 \\
Seller rating & -0.13 & 0.10 & -1.31 & 0.191 \\
Used & -1.20 & 0.56 & -2.09 & 0.037 \\
Used*rating & 0.17 & 0.10 & 1.65 & 0.099 \\
300 models & 0.11 & 0.12 & 0.88 & 0.378 \\
600 models & 0.66 & 0.10 & 6.65 & 0.000 \\
DVD device & 0.19 & 0.12 & 1.63 & 0.104 \\
$\sigma$ & 0.44 & 0.02 & 17.71 & 0.000 \\
\hline
Shrek DVD & & & & \\
Constant & 2.22 & 0.08 & 26.31 & 0.000 \\
Seller rating & 0.03 & 0.01 & 4.24 & 0.000 \\
Used & -0.17 & 0.08 & -2.14 & 0.033 \\
Edition & 0.10 & 0.03 & 2.96 & 0.003 \\
$\sigma$ & 0.48 & 0.01 & 45.04 & 0.000 \\
\hline
Potter Book & & & & \\
Constant & 1.63 & 0.26 & 6.21 & 0.000 \\
Seller rating & 0.09 & 0.03 & 2.62 & 0.009 \\
Used & -0.27 & 0.24 & -1.13 & 0.260 \\
Hard Cover & 0.25 & 0.11 & 2.26 & 0.025 \\
Audio & 0.78 & 0.11 & 7.00 & 0.000 \\
One/Lot & 0.45 & 0.16 & 2.72 & 0.007 \\
$\sigma$ & 0.84 & 0.04 & 23.43 & 0.000 \\
\hline
\end{tabular}
\caption{Regression of Bid on Seller Rating and Item Characteristics}
\end{table}

\textit{Note.} In the Dell data, for the models we studied, there were no used sales in that period.
translate into higher prices. We also showed theory and evidence that suggest they might matter more for lower-priced items. Given the relationship between rating and price, it may be possible for a lower-quality seller to increase his rating for a period of time, only to default at a later point in time. Such practices could spell the end of the rating system as an effective monitoring device. The success of the rating system crucially depends on negative ratings being an effective signal of future default. The next section examines this point theoretically and empirically.

**DOES SELLER RATING SIGNAL FUTURE PERFORMANCE?**

The issue we attempt to tackle in this section is the effectiveness of the seller rating on eBay as a signal of future default probability. Recall from the Introduction that the buyer has no well-defined incentive to report negative feedback or any feedback for that matter. First, the buyer receives no reward or compensation for giving any feedback, other than his own self-gratification (no formal argument exists for why giving feedback result in gratification). Second, the seller could retaliate for negative feedback in multiple ways, from not delivering the product to reciprocating with negative feedback of his own. Third, the ability to provide feedback is time-sensitive: Whereas a positive feedback tends to be relatively quick (once the item was delivered in good condition), negative feedback can take longer as the buyer may first attempt to work out a resolution with the seller. Fourth, buyers may be reluctant to give negative feedback. Once given, feedback cannot be changed, and, as long as there is a chance that the seller was not at fault, a buyer may not feel obliged to retaliate with negative feedback. As such, seller rating need not be correlated with future expected performance. Moreover, even if buyers provide meaningful feedback, negative feedback may be indicative of isolated failure incidents. All these reasons and others could reduce the predictive usefulness of the seller’s rating. In this section, we present a model and empirical evidence that suggest that seller default in the present is indicative of seller default in the future.

**A Model of Seller Default and Termination**

We present an infinite horizon model showing that likelihood of termination can be signaled by prior default. We assume that every seller exits the system at some point due to some rare event such as bankruptcy or death. The event leading to exit is assumed to be exogenous. Sellers find out one period in advance that they are about to exit. That is, the seller finds out he is “terminal” and is forced to exit the system in one period. Seller’s exit is modeled as a Bernoulli variable with probability \( \rho \) of exit in each period.

There are two types of products, High quality and Low quality (H and L). We denote consumer valuations for the two product qualities by \( U_H \) and \( U_L \).
respectively, where $U_H > U_L$. Since we know that all sellers will default in the last period of their lives, even the most optimistic buyer must account for that possibility. Therefore, we let $V_H = \rho U_L + (1 - \rho)U_H$ denote the expected valuation of the most optimistic buyer. $V_L = U_L$ is the expected valuation of the most pessimistic buyer.

The seller’s costs for High and Low qualities are $C_H$ and $C_L$, respectively, where $C_H > C_L$. We assume that the difference in valuations between high and low valuations exceeds the difference in the respective costs. That is, $(V_H - V_L) > (C_H - C_L)$.

Consider the following strategy for buyers: Buyers who observe a default on the part of a seller “punish” the defaulting seller by refusing to pay above $V_L$ to that seller forever. In the absence of observed default in the seller’s past, buyers are willing to pay $V_H$. This “punishment” for observed default is rational since the seller has revealed himself as a potential risk and since there are many other identical sellers to transact with. Luckily for defaulting sellers, not all buyers are as perceptive as others. Some buyers do not observe a default immediately. Delays of information are common in all information environments and the Internet is no exception. We say that buyers who perceive a default with a lag are not vigilant and are “fooled” for the duration of the lag. However, over time all buyers will eventually come to observe a seller’s folly and resort to the punishment strategy. Once a seller defaults at time $t$ a proportion of buyers $0 < \beta(t) < 1$ are “fooled.” That is, they do not attend to seller ratings and do not punish the seller. We let $\beta(t) = \beta$. Specifically, $\beta(t)$ declines over time as more buyers gradually become aware of the negative ratings. We call $\beta$ the vigilance parameter since it indicates the level of vigilance among buyers.

Consider the terminal seller—a seller who is about to exit in the next period. He will choose low quality in the current period if the following holds:

$$(V_H - C_L) + (1 - \beta)(V_L - C_L) + \beta(V_H - C_L) > (V_H - C_H) + (V_H - C_L).$$

That is, for a terminal seller, low quality can be expected if

$$(C_H - C_L) > (1 - \beta)(V_H - V_L).$$

Similarly, the non-terminal seller—the seller who is not about to default would choose low quality in the current period if

$$(V_H - C_L) + \sum_{i=1}^{\infty} (1 - \rho)^i[(1 - \beta^i)(V_L - C_L)] + \beta^i(V_H - C_L) > \sum_{i=0}^{\infty} (1 - \rho)^i[(V_H - C_H)],$$

which we can simplify to obtain:

$$(C_H - C_L) > \sum_{i=1}^{\infty} (1 - \rho)^i[(1 - \beta^i)(V_H - V_L) - (C_H - C_L)].$$

This finding means that a separating equilibrium exists. That is, for moderate levels of vigilance, only terminal sellers will provide low quality and all non-terminal sellers will provide high quality. This assertion can be tested with the hypothesis that seller ratings are indicative of default probabilities.

**H2:** Toward the end of the seller’s activity, seller ratings will drop.

We choose 4 months as the appropriate terminal period based on the observation that four months tend to encompass most of the default activity. The mathematical support for hypothesis 2 (in the Appendix) is based on the simple observation that when all buyers are vigilant, all sellers provide high quality, except for the last period of business. When no buyers are vigilant, all sellers provide low quality. For levels of vigilance in between the two extremes, a separating equilibrium exists that provides different optimal strategies for terminal and nonterminal sellers. Nonterminal sellers can credibly signal their types, and no incentive compatible mimicry is possible that will enable a terminal seller to pass himself off as a nonterminal seller.
Empirical Evidence

We begin with some anecdotal evidence. In Figure 1, we show negative ratings for four eBay sellers who have terminated their eBay membership. A linear trend implies that the seller has a constant default rate. A convex trend implies an increasing default rate as the seller approaches the end of its life.

All four sellers have accumulated negative ratings. The \( y \)-axis represents the cumulative number of negative feedbacks a seller has accumulated by a particular month. The \( x \)-axis represents the month, from the start month to the last month of membership. The seller named Socktimer has accumulated over 30 negative feedbacks by the time he terminated his membership. However, the rate of accumulation appears constant. Socktimer did not increase his rate of default prior to termination. Not so with the next three sellers, banrighaceo, bellboy221, and northshoregiftsandmore. All three had a steep increase in their rate of default near the end of their membership. Seller northshoregiftsandmore accumulated 100 negative feedbacks in 3 months, more than tripling his total negative feedbacks of the previous 5 months. Seller banrighaceo similarly went from 10 to nearly 30 negatives in 1 month after 10 months of relatively flat increase.

We next examine the predictive usefulness of past default frequency for predicting future default likelihood. The data consist of 49 no-longer-registered eBay sellers who had made sales in the “collectibles” category (though they may be cross-category sellers). We refer to these as “terminal” sellers. That is, for various reasons these sellers or eBay deemed it proper to terminate their presence on eBay. As a control group, we add 34 randomly selected nonterminal sellers to the sample. For each seller, data were expressed in monthly terms to reflect the expected delay in feedback (negative feedback is generally issued between 14 and 30 days following the transaction).

For each seller we collected the following variables:
1. proportion of negative ratings in each month
2. number of negative ratings in each month
3. volume of sales in each month
4. track record in terms of total sales by a given month
5. negative track record in terms of total negative sales by a given month
6. seniority in terms of months registered on eBay.

Summary statistics are provided in Tables 3 and 4. The average monthly proportion of negative ratings was 0.034, which is substantial. Overall, the proportion of negative ratings to overall transactions was 0.026, which is still high. This means that 2.6 out of every 100 transactions end in negative outcome, assuming all negative outcomes are reported.

Three regressions were run. The first regression had the monthly number of negative ratings as the dependent variable. The second regression had the monthly proportion of negative ratings as the dependent variable. The third regression had the difference in the proportion of negative ratings from the previous month to the current month. Explanatory variables were the lagged dependent variable, lagged monthly volume, the seniority of the seller (not lagged since this variable can be observed in real time), a “terminal” flag to account for whether the seller is terminal or not, and an “end” flag to indicate whether the observation is in the last 4 months of the data available on that seller. In the case of terminal seller, the “end” flag also indicates whether the seller is in the last 4 months of his eBay activity. All regressions allowed for individual seller effects by differencing. That is, the regressions control for seller’s propensity to default and, given this propensity, seek to capture temporal dependencies.

Differencing the regression equation removes the individual effects and produces an equation that is estimable by instrumental variables. From Table 5, we see that last month’s negative rating is indicative of this month’s negative rating. That is, negative ratings are not independent draws each month but are rather serially correlated. We also see that, for terminal sellers, the number of negative ratings rises substantially and significantly towards the end of their lives (End*Terminal). In contrast, the last 4 months’ observations on nonterminal sellers (End) display an insignificant drop in negative rating. This finding is consistent with the separating equilibrium of in the previous subsection, A Model of Seller Default and Termination. Seniority and track record are both insignificant.

We next examine whether the above result is invariant to choosing the proportion, rather than the absolute number, of negative ratings per seller.
Different sellers sell different volumes and so this dependent variable may be deemed more appropriate. As Table 6 demonstrates, the results are consistent with the first regression (and therefore with the separating equilibrium of the previous subsection on the model of seller default and termination). Terminal sellers are significantly more likely to default toward the end of their lives on eBay, but nonterminal sellers are not. In addition, lagged proportions of negative ratings are significant indicators of future negative ratings, indicating strong serial relationship. It appears that despite the obvious shortcomings of the feedback system, past negative ratings can be used as a signal of future default rates.

The Arellano-Bond test for first- and second-order autocorrelation in residuals rejected the null hypothesis of no first-order correlation, but it is not possible to reject the null of no second-order autocorrelation. However, the presence of first-order autocorrelation does not imply that the estimates are inconsistent. (See Arellano and Bond, 1991, pp. 281–282).

Finally, it is important to look at the dependency of the change in the default propensity over time on the “end” and “terminal” flags. That is, do sellers—terminals and nonterminal—accelerate their default propensities over time, and does the rate of change vary closer to the termination date? Though our separating equilibrium makes no prediction on the rate of change over time, it is intuitive that as the terminal seller nears the date of termination, the rate of default would increase. With the change in the monthly proportion of negative ratings as the dependent variable, we are able to investigate this question.

Table 7 indicates that nonterminal sellers lower their propensity to default over time. This makes sense if default is due to inexperience. As sellers accumulate experience, their default rate falls. However, terminal sellers accelerate their default rate towards the end of their lives, increasing the proportion of negative ratings from one month to the next.

CONCLUSIONS

The Internet has been lauded for its great potential to reduce or even eliminate competitive barriers. This is because small retailers, or even individuals wishing to sell an item, can avoid the prohibitive costs of setting up a brick and mortar shop, finding a good location, and advertising. However, these same benefits come with the Internet’s greatest disadvantage—anonymity. That is, the potential for fraud, default, or misrepresentation is far greater on the Internet. Hence one of the most important roles of the infomediary—endorsing or at least providing some credible information on the seller. Many infomediaries—exemplified here by the auction houses—have resorted to buyer-based ratings. Such rating systems face a major hurdle: insurance—implicit or explicit—by the infomediary may remove any incentive on the part of the buyer to take these ratings into consideration.

The present study addresses this issue both theoretically and with empirical investigation. We constructed a theoretical model that shows that, even in the
presence of insurance, buyers will pay attention to seller rating as long as it is correlated with default probability. This is a result of the assumption that a dispute process is not costless. This same model yields the interesting and perhaps surprising finding that buyers will pay greater attention to seller rating when buying less expensive items, since they are less likely to enter a dispute on such items. A dispute involves a nonnegligible cost and the insurance deductible may make a dispute a losing proposition. As such, seller ratings are critical for items such as CDs and books but not for larger items such as new PCs. Empirical evidence lends validity to the findings of the model. Specifically, both statistical significance and rating elasticity were higher for Harry Potter books and Shrek DVDs than they were for IBM laptops and Dell PCs. This finding has far-reaching implications for online retailers whose quality and delivery are insured or guaranteed by an intermediary. These include retailers on eBay, third-party sellers on Amazon, and retailers selling through a variety of online intermediaries. If these online retailers are mostly dealing in inexpensive items, maintaining a high seller rating is of utmost importance. The retailers focused on inexpensive items should quickly build a high rating, possibly by offering steep discounts initially. Later on, they can recoup any losses by charging a premium for their reputation. Retailers who primarily sell high-price items, on the other hand, should not focus their efforts on reputation since buyers would not be very sensitive to ratings. Rather, they should focus on communicating the items’ features.

The next issue we examined was whether past default rates were indicative of future default probability or even outright exit by the seller. We first presented a model that showed that asymmetric information, where the seller knows of his own imminent demise but the buyer is left in the dark, would result in terminal sellers defaulting prior to the period of termination. Anecdotal evidence from eBay sellers prior to termination demonstrated this pattern. Regressions showed that prior default is indicative of future default and that terminal sellers are more likely to default.

The above findings have both optimistic and pessimistic implications for the future of Internet retailing through infomediaries. The optimistic implication is that feedback ratings are effective because they are credible. They also have predictive power, and they impact both bids and final prices in auctions. The apparent lack of incentives for buyers to provide negative feedback appears to be irrelevant. Perhaps buyers’ altruistic motives, the need for reciprocity, strong feelings of community, or self-gratification drive them to contribute positive and negative ratings without any need for pecuniary compensation. The infomediary (and its buyers), however, derive real value from these contributions. To the extent that other business models can be built that extract surplus from the Web community, this should be a promising direction for research.

The negative implication lies in the findings that sellers who intend to defraud buyers will build up reputation to extract the maximum surplus and then default on all agreements one period prior to exiting. This idea is not new and was first raised by Sobel (1985), who referred to this problem as the principal-double-agent problem, but the voracity with which some Internet auction sellers have grasped this concept is unparalleled. A well-known fraud case (Freedman, 2002) is that of Stewart Richardson, who sold collectible figurines. In over 6,000 auctions, he earned only 43 negative feedbacks. Then in a span of a few days he sold over 100 figurines, left his shop, employees, and wife, withdrew at least $261,000 from the business bank account, and disappeared.

Fraud is a challenge for eBay and other infomediaries. The reliance on buyer feedback requires little interference on the part of the infomediary. Any interference could risk crowding out of buyers’ sense of obligation (for example, Andreoni, 1993, shows that taxation may crowd out voluntary contributions to public goods) and could serve as legal admission of responsibility for fraud (Warner, 2003). Future research should offer remedies for this infomediary dilemma.

Online auctions are not the only infomediaries to offer feedback ratings. Shopping sites such as BizRate.com, Epinions.com, and Imandi.com have adopted successful customer-rating features. The success of these rating systems implies that they are deemed as credible by customers and relatively free of conflicts of interest that would be present with retailer ratings. However, the same moral hazard faced by sellers in online auction sites applies to other retailing formats. Specifically, fraudulent sellers could build reputations and defect at a later stage, costing both customers and the online intermediary. As feedback systems become
more and more prevalent in online environments, the infomediary dilemma should be recognized by management of infomediary sites as well the sellers that conduct business through such sites.

REFERENCES
Proof to Claim 1
If the buyer were to bid his valuation, then
\[ P = \int_{(1-2m)/2Q}^{\infty} \delta((P-d)-c)f(Q)dQ \]
\[ + \int_{\infty}^{\infty} Qf(Q)dQ \]  
(A.1)

Earlier we defined quality uncertainty with respect to \( \epsilon \). As we are taking expectations in equation (A.1) with respect to quality, \( Q \), we must transform pdf \( g(\epsilon) \) into a function of \( Q \). That is, \( f(Q) = g(\epsilon)\delta(\epsilon/\epsilon Q) = \frac{S}{2mQ} \). Denote \( \theta = \delta(P-d) - c \). Then, as long as the buyer disputes with some positive probability, the bid price is
\[ P = \frac{S(\theta^2 + \frac{e^2}{Q}) - 2\theta(S - 2m)}{4mQ} \]  
(A.2)

From equation (A.2) we can investigate the impact of seller ratings on the winning bid in second price auctions. We can do this by rearranging equation (A.2) (i.e., moving \( P \) to the RHS) to form the function \( f(P, S) = 0 \) and then using the implicit function theorem to obtain the derivative of bid price with respect to seller ratings. That is,
\[ \frac{df}{dP} \frac{dP}{dS} + \frac{df}{dS} = 0. \]  
(A.3)

Thus,
\[ \frac{dP}{dS} = \frac{(\theta - \frac{e}{Q})^2}{2(2mQ(1-\delta) + S\delta(Q - \theta))}. \]  
(A.4)

When \( \theta = Q, dP/dS = 0 \), otherwise, \( \theta < Q \) and \( dP/dS > 0 \). Since the bid price \( P \) is bounded above by \( Q \), so is \( \theta \). As long as \( \theta < Q \), bid prices increase with seller ratings. Similarly, from equation (A.4), we can see that the variance parameter \( m \), adversely affects sensitivity to seller ratings.

Proof to Accompany Hypothesis 1
Hypothesis 1 is based on finding that when \( dP/dS > 0 \), the impact of seller ratings on bid prices declines as price increases. To see this, notice that as long as, \( dP/dS > 0 \),
\[ \frac{d^2P}{dPdS} = -\frac{\delta(Q - \theta)(4mQ(1-\delta) + S\delta(Q - \theta))}{2(2mQ(1-\delta) + S\delta(Q - \theta))^2} < 0. \]  
(A.5)

Proof to Accompany Hypothesis 2
When \( \beta = 0 \), both terminal and nonterminal sellers always choose high quality, except for the last period of their lives. When \( \beta = 1 \), both terminal and nonterminal sellers always choose low quality. When \( 0 < \beta < 1 \) and \( \rho \) is sufficiently close to zero, a terminal seller will choose to sell low quality when \( (C_H - C_L) > (1-\beta)(V_H - V_L) \) and a nonterminal seller (until he becomes terminal) will choose to sell a high quality as long as \( (V_H - V_L) > (C_H - C_L) \). This is because, when \( \rho \rightarrow 0 \) and \( \beta < 1 \), the RHS of equation (13) approaches infinity, ensuring that the nonterminal seller chooses high quality. Note that \( (V_H - V_L) > (C_H - C_L) \) is true by assumption. That is, for sufficiently small \( \rho \) and sufficiently high \( \beta \), there exists a separating equilibrium between terminal and nonterminal seller.