The Value of a Reputation: Evidence from Amazon.com

Alexander M. Cannon
Department of Economics, Pomona College
Claremont, California 91711
alexander.cannon@pomona.edu

Abiel Reinhart
Department of Economics, Pomona College
Claremont, California 91711
abiel.reinhart@pomona.edu

Abstract

In this paper we attempt to quantify the value of a reputation for third-party sellers in Amazon.com's market for used books. Specifically, we investigate whether a strong reputation allows a seller to command higher prices than rivals possessing a weaker reputation. We find that a seller's reputation, which consists of an average numerical feedback score and the total number of pieces of customer feedback, has a positive, small but statistically significant effect on price. Using 16,281 transactions collected during April 2005, we find that an increase in a seller's reputation will cause a 0.013 increase in the used to retail book price ratio.

I. Introduction

In this paper we examine the value of seller reputation and corresponding buyer behavior in electronic exchanges using data collected from Amazon.com. While a number of earlier papers have examined the value of reputation using data from auction giant eBay, no previous study has focused on Amazon.com. The segment of Amazon we study does not use auctions to sell goods, but instead forces sellers to quote their prices, and buyers then purchase goods based on that set price.

Asymmetric information has long been recognized as an influence on economic transactions. For instance, Akerlof (1970) studied the market for lemons and theorized that in a market where one could not determine the true quality of any item, such as the market for used cars, market failure could occur. In addition, Spence (1973) discussed education and job market signaling, and Stiglitz and Weiss (1981) modeled credit rationing as an equilibrium condition.

How then can a market characterized by asymmetric information operate? One solution involves reputation, since reputation can signal quality to a potential buyer. In Spence (1973), education communicates ability to employers. Shapiro (1983) notes that the quality of past products will determine a seller's future reputation. Those who have sold high quality products will be able to sell their future products for a premium. If Seller A has a slightly higher price than Seller B, but also has a better reputation for selling higher quality goods, then Seller A will be able to sell her product for a premium.

Electronic exchanges provide a good testing ground for empirically studying the effects of reputation, due to the existence of asymmetric information between buyers and sellers. The electronic environment does not allow customers to directly observe merchandise before purchasing, and sellers can remain relatively anonymous. Further, due to the low cost of entry, sellers can easily change identities if their ratings decline. Lastly, electronic exchanges can

constantly tabulate and distribute customer feedback, making sure numerical measures of reputation are publicly available. The last point is particularly important, because it implies that reputation is well-defined and that researchers and buyers have identical information.

II. Quantitative Research on eBay

Existing research on reputation value has focused on eBay, the Internet's largest auction site. eBay has been a popular research target because of its size, the broad array of goods it offers, and its vast quantity of information on bids, seller reputation, and actual transactions. Moreover, because observed bidder competition sets the final price, one can use this price as a proxy for a buyer's willingness to pay for a good.

Past research has either used existing transaction data on a cross section of sellers, or has taken the form of a controlled experiment. For instance, in Resnick, Zeckhausen, Swanson and Lockwood (2004), the authors team up with an eBay seller who had established a positive reputation through a long history of transactions. This seller then sold matched pairs of vintage postcards using both his regular eBay identity and that of a new seller. The results showed that buyers were willing to pay 8.1% more for items offered for sale by the established seller.

Cabral and Hortaçsu (2004) carry out multiple tests to examine the connection between reputation, and price and sale frequency. They also examine the impact of negative reviews on seller effort. In particular, they analyze whether a negative review decreases subsequent seller effort, leading to a higher rate of negative reviews in the future. With respect to sales frequency, they found that a seller's first negative review decreased sales growth by 14%. Since sales cannot be directly observed, the paper uses comments as a proxy, under the assumption that a constant fraction of buyers will leave comments. They also found that a 1% decrease in the fraction of positive comments for a seller is correlated with a 9% decrease in the winning bid.

Houser and Wooders (2001) and Melnik and Alm (2002) also find a correlation between reputation and price on eBay. Using a study of Intel Pentium 3 processors, Houser and Wooders show that a 10% increase in the number of positive reviews will lead to a 0.17% increase in price, and a 10% increase in negative reviews will lead to a 0.24% drop in price. Additionally, increasing the number of positive ratings from zero to 15 was correlated with a 5% increase in price. Melnik and Alm study the sale of a Five Dollar gold coin and find that when one's rating doubles from 452 to 904, one will receive a price premium of $.18 on average.

III. The Amazon.com Marketplace

While Amazon is itself one the largest booksellers in the world, it also allows third parties to sell many types of goods (not just books) through a service called Marketplace. To post an item for sale, sellers navigate to the item's Amazon listing and click a link marked "Sell yours here." They then provide information on quality and state how many items they have available. Once they finish this process, their item becomes listed alongside all other sellers' offerings.

Listing an item is free. When a sale is actually made, Amazon charges the seller a fee of $0.99 plus a percentage of the sale price. For books, this percentage is 15%. The fees are collected by automatically deducting them from the buyer's payment. "Amazon Payments" handles buyer payments for goods. From the buyer's perspective, this is the same as purchasing from Amazon directly: they simply use a credit card number to purchase an item. The money is then routed to Amazon, and within a couple weeks is deposited into the seller's bank account. Items expire after sixty days, though they can be reactivated. However, sellers can keep their item listed continuously for an indefinite period by purchasing a "Pro Merchant Subscriber" account, which costs $39.99 per month.

All purchases are covered by Amazon's A-to-Z guarantee, which provides $2,500 of insurance against seller fraud. This insurance covers cases in which the seller did not ship an

item, or the item was materially different than the seller's description. The buyer can return the item, and Amazon covers the shipping costs. To submit a claim, buyers must first wait 30 days, after which they have a 60 day window. This policy differs from eBay, which does not guarantee the stated condition of goods sold, as mentioned by Ba and Pavlov (2002).

Amazon's system for rating sellers is straightforward and similar to the systems used by other online retailers. For each transaction, buyers can leave a rating of 1 to 5 ("awful" to "excellent"), in addition to a text comment which can be no longer than 200 characters. Multiple ratings are averaged and rounded to one decimal point.

When a buyer looks at the list of used books, s/he is presented with a screen showing all third-party copies listed in order of ascending price. Each listing also includes the seller's belief of the quality of the item, the name of the seller's account, the seller's average rating over the last 12 months, the number of reviews from the past 12 months, the lifetime number of reviews, the location from where the item will ship, and lastly, the seller's description of the product. These descriptions generally range from one to three lines of text and cannot include any images of the actual item.

Each seller has a Marketplace profile page. This page displays a variety of summary statistics for customer feedback, such as the percentage of positive reviews (ratings of 4 or 5) in the last 30, 60, and 365 days. One can also view every piece of feedback, including the date of the review, the numerical score, and buyer comments. The profile screen also provides a link to a seller's zShops profile. The zShops allow the seller to organize all of the items they are selling via Amazon.com into custom categories and also gives the seller a place to list their address, website, and logo.

We consider Amazon to be a good research target because Amazon lists detailed reputation statistics, which are prominently displayed for each potential buyer who looks at

the item's page (i.e. each buyer has the same amount of information regarding the seller's reputation). In addition, because there is almost no unobserved information passing between buyers and sellers, a researcher can make more exact predictions regarding buyer and seller behavior. (Cabral and Hortaçsu, 2004)

IV. Data

Amazon Web Services (AWS), which provides access to almost all product data on Amazon.com, facilitated the data collection process. In order to collect a random sample of book ISBN numbers, we used the AWS search engine to group books by the first three letters of the author's last name and identified the ten bestselling books in each group. We also specified that each book have at least ten used copies available, so that we have enough data associated with each book. We then used a random number generator to select 500 books from this large set of ISBN numbers.

Two times a day for an eleven day period (April 16 to 26, 2005), we retrieved the complete list of used copies that were currently selling for each of the almost 500 books in our sample. For each used book listing, we collect information on the book's quality, the price, and seller, the seller's average feedback rating, the number of pieces of feedback the seller has received, and the short block of text the seller has used to describe his/her item.

The most significant problem in the data collection process is that transactions are not directly observable—that is, no seller log exists that shows a complete receipt of sales. Transaction data is essential to determine the connection between reputation and price. For instance, consider the case of two sellers both selling the same used book with the same quality at the same price. Now suppose that one seller has no feedback, but the other seller has a long history of positive feedback. With prices at the same level, reputation still has value,

for the seller with positive feedback could be selling copies with greater frequency than his competitor. Therefore, transaction data is required but is unfortunately, difficult to ascertain.

Fortunately, a partial solution to this problem does exist. While we cannot directly observe transactions, we can observe changes in a seller's inventory. Therefore, we have taken a snapshot of seller inventory every twelve hours. For each sample, we compare the current quantity in stock to the quantity of the previous snapshot. If the amount has declined, we will assume this indicates a sale. We are able to check, however, if a decrease in quantity reflects a cancelled listing. Therefore, once proper eliminations of cancelled (and not sold) books have been made, we have a working data set. In all cases we use the price from the period just before quantity decreased as the transaction price.

We have used the critical assumption that sellers do not constantly restock their inventory immediately after a sale. If they do restock their inventory, presumably they would restock in bulk at one moment in time. In other words, we do not believe that sellers are prone to constantly restocking their inventory on a daily basis. If a seller does restock their inventory by any amount, we should capture this increase when the quantity increases. One problem presents itself on a restocking day, however, when new inventory could offset a quantity decrease caused by sold items. We do not believe that this will be a considerable problem.

We have also assumed that if a seller has a premium account, the seller will not remove the book from the used book listing. In addition, if a seller does not have a premium account, his/her book will automatically drop out of circulation after 60 days, forcing the seller to manually re-open the listing. Therefore, we will compare the list of used offerings from the current snapshot with the previous day. If a listing has disappeared, we need to check

if its quantity has gone to zero, or if it has simply expired. AWS provides access to the expiration date, so the expiration of the book can be viewed.

V. Methodology

1. Testing for a Price Premium

The first regression attempts to explain the price of a used book by looking at the book's new retail price, quality, and its seller's reputation.

$$(P/RP) = \alpha_1 + \beta_1 R + \beta_2 TF + \beta_3 LN + \beta_4 VG + \beta_5 G + \beta_6 B + \beta_7 \ln(1+(R*TF)) + \varepsilon,$$

P is the used book price, RP is the retail price charged by Amazon, R is the seller's average numerical (1 to 5) rating, and TF is the total quantity of customer feedback for the seller. LN (like new), VG (very good), and G (good) are three dummy variables that represent the seller defined condition of the book (the "Acceptable" condition has been omitted), and B is a dummy variable for the binding of the book (0 for paperback and 1 for hardback). In addition, we include an interaction term, (1+(R*TF)), which accounts for the possibility that the effect of the rating term depends on the level of customer feedback. Test One only considers actual transactions; items that are listed but never sold are not analyzed here.

We expect to find the coefficients on book quality are all positive and that the coefficients on variables indicating better seller reputation are also all positive. However, we do not necessarily expect that the coefficient for total feedback is constant. Rather, it likely decreases but remains positive as total feedback rises, reflecting some form of diminishing marginal value of feedback.

2. Testing for a Price Premium under Different Market Conditions

We expect that competition will reduce a seller's ability to set prices, thus the second regression extends the first to take into account the presence of competing sellers.

$$(P/RP) = \alpha_1 + \delta_1 R + \delta_2 TF + \delta_3 LN + \delta_4 VG + \delta_5 G + \delta_6 B + \delta_7 \ln(1+(R*TF)) + \delta_8 \ln(NL) + \delta_9 (PVAR) + \varepsilon,$$

Here, NL is the number of other listings of the same book not controlled by the current seller. Note that sellers may have multiple listings for one book (for instance if they have multiple copies of differing quality), so the total number of listings is larger than or equal to the total number of competing sellers. The expected sign of the coefficient is negative.

PVAR is the variance of listed prices for the given book. It is possible that if prices are more dispersed (perhaps due to greater quality differences) a seller will have more freedom in setting their price. Another issue is that when prices are all close together, raising a listing's price could push it very far down the price-ordered list of offerings, decreasing its probability of being sold.

It may be possible to enhance this regression in the future. For example, it might be worthwhile to examine the proportion of sellers of different reputations in the marketplace and the number of other books listed that have the same quality as the one we are looking at.

3. Price Ratios of Sold and Unsold Items

A limitation of the tests above is that they consider only items that were sold. Thus they do not account for the situation in which an item sells but there was an item with equal or greater quality listed at a lower price. The following test focuses on this issue.

$$(P/(PLUS*RP)) = \alpha_1 + \gamma_1 TF + \gamma_2 LN + \gamma_3 VG + \gamma_4 G + \gamma_5 B + \gamma_6 \ln(1+(R*TF)) + \gamma_7 \ln(NL) + \gamma_8 (PVAR) + \varepsilon,$$

Here P is the price at which a given item sold, just as before. PLUS is the price of a copy of the same used book with the following properties: 1) PLUS < P; 2) quality of item PLUS is equal to or greater than item P and it has the same binding; 3) item PLUS did not sell in the period in which item P was sold; 4) item P and item PLUS are not being offered by the

same person; and 5) PLUS is the minimum price of all items matching the first four properties.

The independent variables are as before and are for the item that sold. The idea here is to measure how a seller's reputation affects the ratio of their price to the prices of lower-priced goods that did not sell. We expect that a better reputation leads to a larger ratio.

Note that the ratio P/PLUS is a lower bound on a price premium, not the price premium itself. For instance, if seller A sells at $15, and seller B's offering of equal or greater quality which did not sell was priced at $10, it is incorrect to say A's price premium over B was $15/$10 = 1.5. Rather, we know the premium was at least 1.5; it could be that had A raised his price to $16, he still would have sold ahead of B.

VI. Results and Analysis

The results from Test One, with 16,283 data points, are summarized in Table I. A book that has either a "Like New" or "Very Good" condition label will cause the used book price to retail price ratio to increase by .36 and .09 respectively. This result is as expected, with better quality pushing the used price closer to the retail price. The "Good" condition label does not show any significant effect on the price ratio. The binding term's coefficient is small and statistically significant; however, it is likely meaningless. This is because a single ISBN number only represents a single type of binding. This variable is only included because sometimes sellers list an edition with a given binding under the wrong ISBN. Therefore in most cases the binding variable does not vary at all for a particular book.

The average feedback term is not statistically significant. This is likely because there is little variation in the average rating for sellers, with most averages falling between 4.0 and

5.0 on a scale of 1 to 5. The total feedback coefficient is very slightly negative, the opposite sign of that expected, and significant.

The most valuable reputation term is the interaction term between the average feedback score and total feedback quantity. Interacting these two terms is important because both a high score and a large volume of feedback convey reputation and permanence. This variable is statistically significant, with a one percent increase yielding a 0.013 increase in the price to retail price ratio.

The first test resulted in an adjusted R^2 of 0.3393. We then moved to Test Two (please see results in Table II), and included two further variables, NL and PVAR. The inclusion of these two variables has strengthened our results in terms of significance, while also increasing the adjusted R^2 value to 0.3458.

All else equal, a one dollar increase in the price variance will yield a very slight increase in the price ratio ($1.56*10^{-9}$). However, if the quantity of books increases, the price ratio will decrease. This is consistent with the idea that increased competition drives down transaction prices. When the quantity increases by one percent, then the price ratio will drop by 0.0083.

The results of Test Three, shown in Table III, become largely insignificant when transactions are tested that involve the sale of a book which has a price greater than the lowest price of that same book. With an adjusted R^2 value of 0.022, this test explains little with regard to price premiums.

VIII. Conclusion

This study has shown a small positive correlation between reputation and price premiums on Amazon.com. Future directions for research include an examination of certain

niche markets, such as the market for used textbooks. This market may have a different group of buyers and sellers than the market for used mystery novels, for example. The group of textbook sellers may be much smaller, which may allow for a wider (or perhaps smaller) range of prices for specific textbooks. As textbooks are in high demand only at certain times of the year (particularly at the beginning of semesters and quarters), prices and quantities of books may fluctuate seasonally. Also, because textbook buyers usually require that their items arrive quickly, they may place a higher value on reputation.

Further research should address the issue of collusion between sellers in the online marketplace. Sellers have the potential to collude together to keep prices at high levels, and therefore make higher profits. In our casual exploration of Amazon.com's used book listings, we found that some books only included sellers with large operations. This could be related to certain niche markets and could therefore be explored with the previous suggestion.

References

Akerlof, G. A., 1970, 'The Market for 'Lemons': Quality Uncertainty and the Market Mechanism', *Quarterly Journal of Economics*, 84, pp. 488-500.

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

Cabral, L. and Hortaçsu, A., 2004, 'The Dynamics of Seller Reputation: Theory and Evidence from eBay', NBER Working Paper 10363.

Diamond, D., 1989, 'Reputation Acquisition in Debt Markets', *Journal of Political Economy*, 97, pp. 828-862.

Houser, D. and Wooders, J., 2001, 'Reputation in Auctions: Theory and Evidence from eBay', University of Arizona Working Paper.

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

Resnick, P., Zeckhauser, R., Swanson, J. and Lockwood, K., 2004, 'The Value of Reputation on eBay: A Controlled Experiment.' University of Michigan Working Paper.

Shapiro, C., 1983, 'Premiums for High Quality Products as Returns to Reputations', *Quarterly Journal of Economics*, 98, pp. 659-680.

Spence, M., 1973, 'Job Market Signaling', *Quarterly Journal of Economics*, 87, pp. 355-374.

Stiglitz, J. E. and Weiss, A., 1981, 'Credit Rationing in Markets with Imperfect Information', *American Economic Review*, 71, pp. 393-410.

Table I

Results of Test One (16,281 Observations)

Variable	Coefficient	Standard Error	t-Statistic
Constant	**0.267259**	**0.016689**	**16.01452**
Average Rating	0.001296	0.003395	0.381783
Total Feedback	**-2.37E-06**	**8.39E-08**	**-28.30035**
New Condition	**0.367616**	**0.010262**	**35.82338**
Very Good Condition	**0.090443**	**0.011621**	**7.783027**
Good Condition	0.008505	0.01209	0.703432
Binding	**-0.02319**	**0.002252**	**-10.29527**
Interaction Term	**0.012585**	**0.001128**	**11.15362**
Adjusted R-squared	0.33933		

Table II

Results of Test Two (16,281 Observations)

Variable	Coefficient	Standard Error	t-Statistic
Constant	**0.297231**	**0.017609**	**16.87962**
Average Rating	0.002198	0.003379	0.650502
Total Feedback	**-2.43E-06**	**8.44E-08**	**-28.81889**
New Condition	**0.36221**	**0.010248**	**35.34556**
Very Good Condition	**0.087034**	**0.011576**	**7.518572**
Good Condition	0.006413	0.012037	0.532794
Binding	**-0.024393**	**0.002264**	**-10.77401**
Interaction Term	**0.01261**	**0.001123**	**11.23027**
Total Used	**-0.008339**	**0.001359**	**-6.136868**
Variance	**1.56E-09**	**1.44E-10**	**10.81575**
Adjusted R-squared	0.345826		

Table III

Results of Test Three (15,150 Observations)

Variable	Coefficient	Standard Error	t-Statistic
Constant	**-0.00332**	**0.000799**	**-4.154311**
Total Feedback	**-9.29E-09**	**4.15E-09**	**-2.239396**
New Condition	0.000148	0.000559	0.264799
Very Good Condition	**0.001507**	**0.000644**	**2.338397**
Good Condition	**0.001805**	**0.000667**	**2.705233**
Binding	-5.27E-05	0.00011	-0.480769
Interaction Term	0.000108	5.58E-05	1.929566
Total Used	**0.001062**	**6.56E-05**	**16.17546**
Variance	**2.05E-11**	**6.68E-12**	**3.077047**
Adjusted R-squared	0.022976		