

Price, Quality and Reputation:
Evidence from An Online Field Experiment ¹

Ginger Zhe Jin
University of Maryland and NBER

Andrew Kato
University of Maryland

November 13, 2005

¹Correspondence: Department of Economics, University of Maryland, College Park, MD 20742; phone (301) 405-3484; email jin@econ.umd.edu. We benefit from comments of Austan Goolsbee, Raphael Thomadsen, John Shea, Dan Vincent, David Reiley, Larry Ausubel, Peter Cramton, V. Joseph Hotz, Jeff Smith, Jimmy Chan, Vincent Crawford, Mark Duggan and attendants at numerous seminars and conferences. We are particularly grateful to Seth Sanders and John List for their constructive advice at the early stage of the research, to Timothy Bresnahan, Rachel Kranton and Thomas Hubbard for their detailed suggestions in reshaping earlier versions, and to Editor Ariel Pakes and two anonymous referees for their careful readings. Special thanks to 8 friends who acted as our agents in purchasing baseball cards in retail markets, and numerous sports card store owners who shared their insights regarding the sportscard industry. Excellent research assistance from Randy Alexander Moore and Krzysztof Fizyta is gratefully acknowledged. Any remaining errors are ours.

Abstract

This paper examines the link between price, quality, seller claims and seller reputation in Internet auctions. To obtain a precise measure of quality, we purchased actual baseball cards and have them professionally graded. These cards were systematically purchased so half of the sample came from sellers making high quality claims and the other half from sellers making modest or no claim. We compare the quality data to the prices paid by online buyers for goods with similar claims.

We find that some buyers in the online ungraded market are misled by non-credible claims of quality. They pay higher prices but do not receive better quality and in fact are defrauded more often. Online seller reputation is found to be effective for identifying good-faith sellers. But conditional on completed auctions, reputable sellers do not provide better quality. Evidence also suggests that high-claim sellers target less experienced buyers. We attribute these data patterns to misleading signals in the online ungraded market and two loopholes in the eBay rating system, namely universal ratings and costless switching of anonymous identities.

1 Introduction

Theorists argue that, if consumers do not observe product quality before purchase, they may use price, advertising, or seller reputation to infer quality. The inference is validated by monotone relationships between these “signals” and true quality. For example, in Nelson (1970), only high quality sellers can afford costly advertising because high quality products are more likely to attract repeat purchases.¹ In a slightly different setting, Klein and Leffler (1981) argued that high price in the company of high quality generates a continual stream of rental income, which motivates sellers to provide high quality in the long run rather than cheat and disappear in the short run. Similarly, Shapiro (1983) suggested that sellers who sold high quality products in the past have an incentive to maintain their reputation, because good reputations are rewarded with high prices and high profits in the future.

An ideal test of the theory is to link true quality with price, advertising and reputation. However, since quality is rarely observable, researchers often document positive correlations between price, advertising and reputation, and conclude that the positive correlation(s) support the theory.² Such exercises are particularly popular in the context of Internet auctions, because most auction sites (such as eBay) offer the complete record of transactions plus feedbacks from previous transactions.³

Unlike previous studies, we design a field experiment in Internet auctions, obtain a key variable – true quality⁴ – and conduct a direct comparison between quality, price, seller claims and seller reputation. We focus on baseball cards whose physical qualities have not been professionally certified at time of sale. As detailed in Section 2, quality is one of the most important determinants of card value but the quality of an ungraded baseball card is very difficult to detect online. In fact, the only information available to buyers in these auctions is quality claims and descriptions made by sellers in the auction listings.

¹A crucial assumption in Nelson (1970) is that consumers observe the true quality after purchase. Milgrom and Roberts (1986) formalized the idea, concluding that sellers of new products may signal high quality via both price and advertising.

²Such inference is less problematic if the true quality is perfectly observed after transaction and the observation helps update the reputation. For example, Gorton (1996) studied the early Bank Note market and found firms issuing debt for the first time paid higher interest rates than otherwise identical firms. Over time, the premium was decreased as the market observed the credit history.

³Resnick and Zeckhauser (2002) and Bajari and Hortacsu (2003) provide excellent surveys for these studies.

⁴Most studies on eBay seller ratings do not observe true quality. To the best of our knowledge, the only exception is Resnick *et al* (2003) who conducted a controlled experiment allowing the same (honest) seller to sell similar items under different eBay identities.

The evidence was gathered in two steps. First, a seven-month eBay watch revealed a stunning fact: 33-51% higher prices were paid to sellers claiming to be selling high quality cards (as opposed to modest or no claims). This would be reasonable if sellers making grand claims actually delivered better cards or provided more reliable service. We checked this in the second step. Immediately following the market watch, we purchased ungraded cards with the same types of claims and had them professionally evaluated. These cards were systematically purchased so half of the sample came from high claims sellers and the other half from modest or no-claim sellers. In a reverse of the theory, sellers making the best claims were more likely to commit fraud by defaulting (no delivery) or sending counterfeits. Even conditional on actual deliveries of authentic cards, the average quality from high claims sellers was indistinguishable from the average quality from sellers with more modest claims.

Given the freedom to make non-verifiable claims (aside from the reputation concerns which we discuss below), the incredibility of seller claims should not be a surprise to theorists. What is surprising is the amount of price premium that buyers are willing to pay for empty claims. After examining a number of alternative explanations, we conclude that the most likely answer is some buyers being uninformed. These buyers, unlike the uninformed buyers assumed in most of the literature⁵, are not only uninformed of the quality of a specific seller (or the quality of goods they sell), but also possess incorrect beliefs on the distribution of quality in the market. More importantly, the prices we observed would not correct such misbeliefs. On the contrary, if an uninformed buyer watches the eBay market as we did and he is naive enough to believe the application of the classical theory to the eBay setting, he is likely to fall for the empty claims.

Does the eBay rating system help lower the risks of cheating and signal better product quality? Our data suggests that reputable sellers are less likely to make bold claims and less likely to default or send counterfeits. However, conditional on authentic delivery, the actual card quality is uncorrelated with seller ratings.

Our eBay watch data also confirm a puzzle documented in other studies of online reputation: buyers are more willing to place bids on items offered by reputable sellers, but conditional on completed auctions, seller ratings has no significant impact on the winning price. Furthermore, we find that the boldest claims not only yield price premiums but also increase the

⁵Most literature involving uninformed buyers limits the buyer's uncertainty to information regarding specific sellers' characteristics. For example, though buyers may not know the quality of goods produced by a specific seller, they are assumed to be so sophisticated that they know the distribution of quality of all goods for sale in the market. Such assumption is explicit in Akerlof (1970), Shapiro (1983), and many follow-up papers summarized in Tirole (1988). To be precise, buyers in Shapiro (1983) are even over-sophisticated: they anticipate infinite supply of hit-and-run sellers, and therefore refuse to pay non-zero price for non-reputable sellers, although some non-reputable sellers are new honest ones.

probability of sale. This suggests that, if a seller is only concerned about the current sale, the lower likelihood of sale associated with lower seller reputation can be easily overcome by making an incredible claim of superior quality. This explains why bold claims are often made by sellers with relatively low ratings. If naive buyers understood the negative correlation between seller claims and seller reputation, even seller reputation alone could be very effective in helping them avoid fraud. This suggests that the key problem is not the lack of signals, rather it is the confusion that results from mixing good signals like the eBay seller rating with misinterpretations of bad signals like seller claims.

The confusion may be attributable to two reputation loopholes. First, eBay ratings are universal across categories and roles. It is easy to build up an eBay rating by buying cheap items and use those reputation points to sell expensive items.⁶ A random check of feedback profiles also suggests that most feedbacks focus on the common conduct behind various transactions: the ability to complete the transaction in a timely manner. Feedback on the actual quality of the delivered item is rare, as it is hard to compare quality across categories and the absolute level of quality is subject to interpretation. This explains why the information that seller ratings successfully convey in our data set is the seller's inclination to cheat (by making bold claims, default, or send counterfeits) rather than the actual physical quality conditional on authentic delivery.

To make matters worse, one may switch between anonymous identities with little cost in eBay. In the experiment we encountered two fraudulent sellers who intentionally built up positive ratings, committed a series of defaults, received over 20 complaints, and abandoned the accounts soon afterwards. Because the cheater can switch to a clean slate, warnings from today's buyers cannot reach new buyers tomorrow. Like universal ratings, this loophole reduces the precision and accessibility of information on past seller conduct, casting doubt on how applicable the classical theory is to eBay.

The rest of the paper is organized as follows. Section 2 explains why baseball cards are especially suitable for this study. Section 3 presents empirical findings from our market watch, followed by experimental design in Section 4 and experimental results in Section 5. Section 6 discusses the implications from our empirical findings and Section 7 concludes.

⁶Although eBay saves feedback comments indefinitely, users cannot view the auction being commented on after 90 days have passed. Random inspection reveals that very few comments mention the item in the comment itself.

2 Baseball Cards and eBay

Three features of baseball cards make them an excellent example. First, each type of baseball card is a homogeneous good and card quality is the most important determinant of card value. Second, industry accepted professional grading services exist to verify card quality, identity, and authenticity. This allowed us to directly observe quality and seller performance. Finally, online auction is one of the most complained about markets for consumer frauds, and quality concerns pertaining to online ungraded baseball cards are likely to be present for a multitude of goods. We now present institutions in support of these arguments.

Each year, card companies design and print sets of cards depicting players and events from the previous season. Once the print run of a particular set ends, the supply of each distinct card in the set is fixed.⁷ The value of a particular card depends on its scarcity, the player depicted, and the condition of the card. By card condition, we mean the physical condition of the edges, corners, surface and centering of the printing. People often use a 10-point scale to track card condition. For example, flawless characteristics (even under microscopic inspection) will rate a perfect 10 while obvious defects to the naked eye like minor wear on the corners of the card might lower a corners grade to 7. The card's overall grade is computed off all four characteristics.

Card flaws are often hard to detect by the naked eye. In retail markets, buyers often use a magnifying glass to carefully examine card condition before purchase. One may also refer to independent, professional grading for a more accurate measure of quality. Because professional grading is voluntary and costs 6 to 20 dollars per card⁸, this practice divides the market into two groups: graded cards and ungraded cards. Graded cards are encased in plastic and sealed with a sonic procedure that makes it virtually impossible to open and reseal the case without evidence of tampering. The casing indicates the grading service, grade received and a bar code with serial number that identifies the particular copy of the card. Anyone with Internet access can go to the grader's web site and verify the card's grade by looking up the serial number. As of 2005, Professional Sports Authenticator (PSA) and Beckett Grading Service (BGS) remain the largest and most respected of the existing 10-15 grading services.

Single baseball cards can be traded online or offline. Offline outlets include retail stores and local card shows, both involving substantial transaction costs in finding trading partners. These transaction costs are significantly reduced in online auctions.⁹ eBay offered the first

⁷The exact number of copies printed for a specific card is regarded as an industry secret.

⁸Depending on package size and turnaround time, but independent of the actual grades received.

⁹For a typical card with a winning bid of \$100, the total seller fees amount to about \$4, lower than the average 6.7% sales tax that would apply to a typical transaction in the retail markets (Goolsbee 2000).

standardized, user-friendly auction site where individual sellers and buyers could meet and transact conveniently. Since its public stock offering in 1998, eBay has always been the most popular online place to trade single baseball cards.¹⁰

Despite transaction cost savings, online trading intensifies the information asymmetry between buyers and sellers, especially for ungraded cards. An online seller can describe the card condition in a brief title and web page. He can also supplement the text description by one or multiple scanned pictures. If it is a graded card, the picture will reveal the name of the grader and the serial number which identifies the exact card copy. However, if it is an ungraded card, it is impossible to prove that the card depicted is the one for sale. Even if it were the card for sale, poor scan quality reduces the informational value of the picture. While centering on the front may still be observable, a blurry scan makes it impossible to judge the condition of the card corners, edges and surfaces. Also, sellers sometimes do not include a picture of the back of the card, which could be in a different condition from the front of the card. The other auction settings are standard, as described in Lucking-Reiley (2000).

Whether buying or selling, all trading parties refer to a standard price guide - Beckett Baseball Cards Monthly. For each single type of ungraded card, Beckett collects pricing information from about 110 card dealers throughout the country and publishes a high and low price reflecting current selling ranges for Near Mint-Mint (8) copies. It is widely agreed that the retail transaction prices are closer to the low end than to the high end. For graded cards, Beckett lists price ranges by grading company and grade. Card value is convex in grade: for example, the July-2001 low price for Griffey Jr. rookie card is \$60 for PSA8, \$150 for PSA9, and \$1200 for PSA10. The book price allows us to control for market differences across cards and across time.

Online auction is one of the most complained about markets for consumer frauds. According to the Federal Trade Commission (FTC), Internet auction fraud was ranked as the second most common category of consumer complaint, accounting for 10% of the 204,334 Internet related complaints in 2001. For three years in a row, auction fraud was the most frequent offense reported to the FBI's Internet fraud unit. More alarmingly, total reported losses involved in Internet auction frauds tripled from \$17 million in 2002 to \$54 million in 2003.

Quality concerns stated above for baseball cards apply to a multitude of goods in online auction. The most direct analogy is to the other collectibles such as stamps, coins, and antiques, all of which have similar industry features and professional grading services. It is also common to observe seller claims of quality in other top ten trading categories, such as used cars, consumer

¹⁰Based on our interviews with over 30 sports card stores and a phone survey of 1213 sports card stores listed in Yahoo! Yellowpages. See Jin and Kato (2005) for more details about the survey.

electronics, books, clothing, accessories, toys, and jewelry. In fact, the most common complaints about Internet auctions are that "after sending the payment, buyers may receive an item that is less valuable than promised, or worse yet, nothing at all."¹¹

eBay has adopted several measures to combat fraud. First of all, eBay requires either a verifiable email address or a real credit card number for anyone who wishes to buy or sell in eBay. However, because identity theft is the most common form of fraud, it is obvious that a fraudulent seller could open multiple dummy accounts using stolen credit card numbers. Second, eBay has an entire division devoted to investigating fraud complaints. The Fraud Protection Program in eBay reimburses up to \$200 per item (minus \$25 processing fees) for buyers who did not receive any delivery after 30 days of payment. However, for a specific buyer, the reimbursement is limited to 3 claims per six months. Also, unlike a physical auction house, eBay positions itself as a marketplace organizer and therefore does not guarantee the quality of goods actually delivered.¹² Because it is difficult to identify a unique copy of an ungraded baseball card, even retail stores do not allow any return or exchange after sale. Due to the same reason, an online ungraded card that turns out to be counterfeit or below the claimed quality is usually difficult to be returned, exchanged, or reimbursed.

The third and probably the most widely used device to combat fraud is the eBay feedback forum. Every eBay user has a *Feedback Profile* made up of comments from other eBay users - an official "reputation." This profile consists of a numerical score next to the user id in the listing page as well as a separate page of detailed comments left by the individual's previous trading partners. The numerical score is the total number of distinct trading partners that have left positive feedbacks minus the total number of distinct partners that have left negative feedbacks.¹³ By this definition, eBay ratings may include both buying and selling behaviors in the past. Throughout the paper, we use the term "seller ratings" to refer to the eBay ratings an eBay user had when he acted as a seller in our data, and "buyer ratings" for the ratings an eBay user had when he acted as a buyer in our data. Because an individual can open multiple accounts in eBay¹⁴ and on average 52.1% of transactions result in any feedback from one trading

¹¹See the FTC brochure *Top Ten Dot Cons*. Recent eBay fraud lawsuits have involved undelivered computers, counterfeit artwork and counterfeit sports memorabilia.

¹²According to The Wall Street Journal as of January 18, 2001, 6 buyers who were sold counterfeit autographed sports memorabilia in eBay auctions filed a class-action lawsuit against eBay Inc. in April 2000. A San Diego judge dismissed the \$100 million lawsuits, saying that the online auction company was not liable for the sale of phony sports memorabilia on its site.

¹³Note that it is possible for one individual to leave both positive and negative feedbacks on separate transactions with the same trading partner. eBay will count both and this will be canceled out in the net score for that trading partner.

¹⁴The multiple accounts created by the same individual are not linked to each other, unless the account owner

party on the other ¹⁵, the numerical rating is an imperfect measure of reputation.

Numerous papers use auction records alone, including price and auction features, to examine the effect of eBay’s user feedback rating system. Resnick and Zeckhauser (2002) surveyed the existing 12 studies on this topic and found that the literature has not reached a consensus on the price effect: some found zero price premiums for sold items (Eaton 2002), some found tiny price premiums pooling completed and non-completed auctions (Melnik and Alm 2002, Lucking-Reiley et al 2000), and some found more significant effect through lab or field experiments (Ba and Pavlou (2002), Resnick *et al.* (2003)). In contrast, most studies found some expected effects of seller reputation on the likelihood of sale, where sale is equivalent to auction completion.

3 Evidence from Market Watch

Jin and Kato (2005) presents a model where baseball card trading in equilibrium takes place in three market segments: the best quality cards are graded and traded online, mediocre ones stay ungraded in retail markets, and the worst lemons go online as ungraded cards.¹⁶ Because either professional grading or physical examination alleviates information asymmetry, we focus on the online ungraded sector and take retail ungraded and online graded sectors as two reference groups.

To better understand the online market, we observed real eBay auctions for 7 months from April to December 2001. At first, we tracked Ken Griffey, Jr.’s 1989 Upper Deck card because it was the most actively traded single card on eBay and had by far the largest graded population of any card. Based on store owner interviews, we added another four rookie cards¹⁷ that each had at least 10,000 graded copies as of April 2001 and attracted sufficient interest in both retail and online markets. They were the 1982 Topps Cal Ripken #21, 1985 Topps Mark McGwire #401, 1993 SP Derek Jeter #279 (foil), and 1994 SP Alex Rodriguez #15 (foil).

These five cards covered star players with a wide variety of characteristics and range in value from \$30 to \$120. While Ripken and McGwire have recently retired, Griffey is still in the middle of his career, and Jeter and Rodriguez are at the beginning of their careers. Nevertheless, all five are established names. Their card values are much more stable than other modern cards, instructs eBay to do so.

¹⁵According to Resnick and Zeckhauser(2002).

¹⁶The fourth segment– retail graded cards – cannot survive because it incurs higher transaction cost but has no information advantage over the online graded sector unless default risk is severe.

¹⁷The most valuable cards tend to be “rookie” cards of star players, the first card ever printed by a particular company of that player.

and they all have well defined book prices for us to control for any fluctuation in market value. From retail store interviews, we noticed that an overwhelming majority of customers are adults and only pay attention to cards of near-mint 7 or better quality. For this reason, we disqualified all online auctions for ungraded cards that claimed 6 or below or mentioned at least one defect consistent with a professional grade below 7.

Table 1 summarizes these market watch data. Sixty-seven percent of the 1124 auctions were for graded cards. Among graded cards, 88% were graded 8 or above. This is consistent with our expectation that only cards of the highest quality will be graded.¹⁸ Low minimum starting prices were common to auctions for both graded and ungraded cards, and secret reserve prices were relatively uncommon.¹⁹ For both graded and ungraded cards, the likelihood of completing the auction was about 81%, where we define an auction as completed if it received at least one bid above any minimum or reserve price. Among completed auctions, eBay prices closely track Beckett low book prices²⁰, and as we expect, the average winning price is much higher for graded cards (Table 2).²¹

For ungraded cards, there could be two potential signals of card quality: seller ratings and seller claims. Seller claims are abundant, although none of them are verifiable online. Of the 372 ungraded cards in our sample, 43 were self claimed Gem Mint (equivalent to 10), 105 claimed Mint (9 or 9.5), 29 NearMint-Mint (8 or 8.5), and 24 Near Mint (7 or 7.5). The other 171 did not make any quality claim. The high-end claims, especially the 10s, were hardly credible. In May 2001, the average eBay price for a graded 10 Griffey was \$1,450 while an ungraded Griffey with self claim 10 was only \$94.26. Since grading costs no more than \$20, it is obvious that sellers making such claims must be overstating the quality of the cards they are selling. Similar logic applies to Mint 9 or better.

Note that non-verifiable seller claim is a phenomenon unique to the online ungraded sector. During our visits to retail stores, card dealers were very reluctant to make any quality claim

¹⁸Book prices for ungraded cards are defined by Beckett as prices for ungraded cards that seem of quality 8. About 12% graded cards offered in the online markets are graded below 8. One explanation is that these sellers are not experienced enough to tell the difference between 7 and 8.

¹⁹Only 50 out of 372 ungraded and 109 out of 752 graded auctions employed secret reserve prices. Since eBay's listing fees are based on which price range the minimum starting bid falls in, sellers tend to cluster their public minimums at the top of a price range where listing fees change. One such common point is \$9.99. In our eBay watch data, 56.5% of ungraded and 65.3% of graded auctions started at \$9.99 or lower. More generally, 80.6% of all auctions observed in the eBay watch had starting prices at least 25% below the Beckett low price. We control for the presence of reserve prices in all regressions.

²⁰The winning price for graded cards was on average 101.5% of the Beckett low book price while for ungraded cards it was 105.9%.

²¹The price dispersion is also higher for graded cards, because card value is convex of the actual grade.

about their ungraded cards. When we insisted, they would say "7 or 8", which is not informative given that collectors rarely look at cards below 7 and cards that are significantly better than 8 should have been graded. The lack of seller claims offline makes seller claims online suspicious. In spite of this, online buyers are very willing to reward seller claims: the winning prices of cards claimed Gem Mint (10) was on average 28.6% higher than the prices of cards claimed Mint (9), and 75.7% higher than the prices of cards claimed below 9 or nothing. Interestingly, buyers seemed to prefer no claims to low claims: cards with no claims had a slightly higher probability of sale and received higher winning prices than the cards claimed 8.5 or lower. However, both premiums were much smaller than the premiums enjoyed by high claims (9-10).

A second potential signal of ungraded card quality is eBay seller ratings. Table 2 presents a sharp difference of eBay ratings between high-claim and low-or-no-claim cards. On the supply side, sellers that claimed 9 or better quality for ungraded cards on average had net ratings around 300, while sellers that made moderate or no claims had net ratings around 700-2000. This difference is statistically significant at 99% level. On the demand side, buyers of graded cards seemed more experienced than buyers of ungraded cards (221.70 vs. 140.03, significantly different with 99% confidence). Within ungraded cards, buyers of cards claimed Gem Mint on average had an eBay rating of 81.45, while that number for the other ungraded cards is around 140 to 240. This seemingly large difference is not statistically significant, because the range of buyer ratings is huge. However, the buyer rating for cards claimed 10 is significantly lower, if we include all the other cards in the comparison group.

Table 3 presents regression results on the same eBay watch data. Defining an auction as observation unit, we focus on two dependent variables: whether an auction was completed, and the log winning price of the auction if it was completed (normalized by corresponding Beckett low book price). The first regression is based on the full sample of 1124 auctions, and entails probit. The second regression is conditional on 778 auctions that were completed and had corresponding Beckett Book prices. To address the potential selection generated by this sampling criteria, we control for the estimated propensity of completing a sale. In particular, from the first regression, we predict the probability of completion for each auction. This predicted probability is used as propensity score in the second regression. Since the propensity score is insignificant in the second regression, this suggests that the selection issue is minimal. In an unreported table, we also use the Heckman two-step procedure, the results are very similar.

The key independent variables are whether the card is graded, the self claimed grade if it is ungraded, net seller ratings, and whether the seller has any negative ratings. Both regressions control for a full set of card identity dummies, payment methods, shipping options, length of auction, whether the auction ended on weekend, whether the auction ended in prime time, where

there was a public or secret reserve price, whether there was a Buy-it-Now option, whether there was at least one digital picture, whether the picture was clear or blurred, and whether the seller claimed to be a card dealer. In each regression, we adopt two specifications: the first treats seller claim of ungraded card quality as a continuous variable, while the second decomposes seller claim into four categories: 10, 9 or 9.5, below 9, or no claim.

Regression results confirm what we found in the raw data: buyers are very willing to reward non-verifiable seller claims in the ungraded market. In the first specification, all else equal, claiming one extra grade for the listed ungraded card will make the auction 4.63% more likely to be completed, and if it is completed, raises the winning price by 21.6%. In the second specification, compared to the omitted group (i.e. claimed below 9), claiming 10 made the auction 8.99% more likely to be completed. Claiming 9 also had a positive (but insignificant) effect on auction completion. Conditional on completed auctions, relative to cards claimed below 9, the winning prices were 33.4% higher for cards claimed 9 and 51.1% higher for cards claimed 10. For ungraded cards with no claims, the price premium was much more moderate (15.1%).

For seller ratings, we find two interesting facts: first, for both graded and ungraded cards, seller ratings are effective in improving the likelihood of auction completion, but ineffective in raising the winning prices conditional on completed auctions. Moreover, the effects of seller ratings on the completion likelihood differ significantly between graded and ungraded cards: the probit coefficient of log seller ratings for ungraded cards (0.178) is twice as big as that for graded cards (0.084). The existence of negative ratings had a big negative effect on ungraded cards, but zero effect on graded cards. These findings confirm the argument that the information problems are more severe for ungraded cards and seller ratings are more helpful in alleviating the problem for ungraded cards.

If a seller is only concerned about the current sale of an ungraded card, making bold claims may be a short cut to increase the probability of sale. In the first specification, doubling the net seller rating increases the likelihood of auction completion by 3.32%, but that is lower than the marginal effect of simply claiming one extra grade in quality (4.63%). In the second specification, the marginal effect of claiming Gem Mint (10) on completion rate (8.99%) is greater than that of doubling the net seller ratings. Given that buyers are willing to pay 33.4-51.1% more for a claim of 9+ but zero for seller ratings, a non-reputable seller making high quality claims is likely to receive more in a single sale than a reputable seller not making high claims. ²²

²²None of this should be taken to imply that a non-sale results in zero payoffs for the seller; it is a reduction in profits. Indeed, eBay allows discounted relisting of items for auctions that do not end in a sale and encourages sellers to relist unsold items. We assume that given a fair market price whenever their cards finally do sell, sellers ought to prefer to sell sooner than later and prefer to avoid any costs associated with relisting.

As a robustness check, we run a tobit regression of winning prices, pooling completed and incomplete auctions and assuming left censoring at minimum bids. While not reported in tables, we find that claiming 10 has a positive effect on winning prices, but seller ratings have zero impact. In an unreported specification, we add in the interactions of seller claims and seller ratings, all of which were insignificant. This contradicts the intuition that claims made by reputable sellers should be more credible. We also estimate the effects of seller claims and seller ratings on the number of bidders: both effects are positive, but only the effect of seller ratings is statistically significant.

To summarize, the regression results reinforce the description of this market that emerges from the summary statistics. At least some buyers are very enthusiastic about bidding in auctions run by sellers who claim to be offering high quality cards. The higher the claims made by a seller, the higher the average winning bid is in his auction. Though, *ceteris paribus*, buyers of ungraded cards prefer to bid in auctions run by high reputation sellers, they are very willing to bid in low reputation sellers' auctions if they claim to be selling high quality ungraded cards. Taken at face value, this would appear to indicate a well-functioning market where those who have high quality are able to somehow communicate that effectively to buyers, and high prices reflect high quality.

However, using eBay ratings as a measure of buyer experience and seller reputation, we find odd patterns that contradict this rosy picture. The majority of cards for sale are already graded, which is reasonable since shop owners and experienced market participants tell us that collectors are really only interested in cards that will grade 7 or higher. In fact, we find evidence that the average winner of a graded card auction is more experienced than a winner of an ungraded card auction. We also find evidence that sellers who decline to make fantastic claims of high quality are not less, but *more* experienced on average than those who do.

If high prices due to high claims in fact reflect high quality for ungraded cards, some empirical facts are hard to explain by the classical theory: the most experienced buyers are avoiding the best deals (ungraded cards with high claims that will actually grade out high); the most experienced sellers with the best reputations are either foregoing profits for no good reason when they run ungraded auctions (by not claiming high quality for their high quality cards) or are not selling high quality cards; and finally, sellers making high claims with presumably high quality cards are foregoing profits by not grading their cards before selling them (due to the convexity in graded card prices).

These puzzles are difficult to explain using only eBay watch data, because they are missing a key variable – true quality. In practice, true quality could have several components: the quality

of the seller in terms of prompt delivery, the authenticity of the delivered goods, the actual quality of the delivered goods, etc. Without reliable data on these quality measures, eBay watch data can only tell us what buyers *believe*. To discover whether buyer beliefs coincide with what really happened, we designed the following field experiment.

4 Experimental Design

We purchased ungraded cards from online internet auctions, and sent them for professional grading. This gives us a direct and objective measure of true card quality. Since seller claims play an important role in eBay auctions, we purchased ungraded cards online in two groups so that the sellers of one group claimed systematically higher quality than the sellers in the second group. The purchasing and grading procedures were designed to mimic actual practices in the market. We focused on the same types of cards as in the market watch, and to avoid any changes in card value due to players' performance, we restricted our purchases to the off season of baseball from December 8, 2001 to March 18, 2002. Since serious collectors only pay attention to cards of Near Mint 7 or better quality, we targeted cards that appeared to have "collectible quality" given the information available at the time of purchase.²³

Each week from December 8, 2001 to March 18, 2002, we searched for active auctions as of the Saturday of that week and ranked the auctions by seller-described card condition. For those cards that carried vague claims like "nice" or "perfect" without any specific grade suggestion or meaningful details for a collector, we treated them as NearMint-Mint (8) claims based on the fact that Beckett guides always quote the same prices for an ungraded NearMint-Mint card and a card graded 8.²⁴ The ranking process suggested that the typical supply of a particular type of ungraded single card consisted of three groups: One group of sellers indicated specific defects about the cards they were selling; a second group claimed their cards were of moderate quality or made no claims; the third group claimed their cards were of extremely high quality (Mint 9 or Gem Mint 10 depending on card type). We focus on the latter two groups (the first group is disqualified).²⁵

²³Specifically, we ruled out all online auctions that mention specific defects that would result in a grade below 7 by PSA or BGS standard.

²⁴This practice was further confirmed by our market watch data, which suggested that the bidding frequency on these cards were between those claimed 7 and those claimed 9+. Descriptions of this sort are the online equivalent of uninformative descriptions given to us by store owners in their shops.

²⁵Given the fact that some price premium (though insignificant) were paid to cards with no claims, one would like to detect the quality difference between cards with moderate claim and cards with no claims. However, the insignificant price premium implies that we need a much larger sample size to achieve any statistical difference of

Specifically, we bid on the best-ranked and median-ranked cards after ruling out the disqualified auctions. In case we lost either one, we bid on the card ranked immediately below it. Given the clustering of seller claims, the perceived difference between the primary and backup cards were negligible. Though we cannot guarantee that every copy we purchased would have been sold regardless of our presence, we restricted bidding to the last five minutes of the auction and did not bid on a card unless it attracted at least one bid by that time. Here we are assuming that the existence of a bid indicates a bona fide buyer's intent to purchase the card, and that transaction would really be carried out. As stated in eBay policies, a completed auction constitutes a legally binding contract. It is possible that sellers and auction winners may communicate after the auction ends and mutually agree to terminate the contract. There is no data for or against this conjecture. Even if it exists, our findings will be unbiased as long as claims are not systematically correlated with contract termination.²⁶

There is also the possibility of a bid not meeting a secret reserve price.²⁷ Consistent with the market watch, secret reserves were rare among the auctions we considered. Only two had secret reserves, of which we won and met the reserve in one and won but did not meet the reserve in the other. The auction where the reserve was met resulted in a bona fide delivery of authentic goods, and ruling these observations out does not change the results.

To ensure winning our targeted auctions, we deliberately overbid. However, we could not bid outrageous amounts because this might distort the market over time if someone realized it and used a shill to bid us up. Therefore, we varied our bidding strategy by adding a relatively large arbitrary amount (10-15 dollars above the leading bid) when we placed our bid very near to the end of the auction. Since eBay buyers tend to "snipe" towards the end of the auctions (Roth and Ockenfels(2002), Bajari and Hortacsu (2003a)), we were sometimes outbid by a very small amount in the last 5-10 seconds and did not have time to "snipe" back. Due to this technical difficulty, we lost 14 cards, 6 of which were best ranked. Since all the lost bids were

²⁶An extreme case would be to assume that the only other bidders we observe in auctions with high claims are actually shills. If we were truly the only non-shill bidder, then some cards we purchased would not have otherwise been sold. But, since listing an item is costly, sellers using shill identities must expect that someone will eventually spring their trap. Therefore, it is reasonable to assume that real bidders exist and these items will be sold sooner or later. However, as a referee points out, if shill bidding is more likely to occur in high-claim auctions, our market watch data may overestimate the probability of sales and the winning prices for cards with high claims. Unfortunately, we do not have sufficient data to detect shill bidding as suggested in Engelberg and Williams (2005).

²⁷To elaborate, a seller may use secret reserve to smoke out high value bidders and their valuations. Even if the secret reserve is not met in the listed auction, the seller can offer the item through E-mail at high prices near those revealed valuations. This story is not entirely impossible, but again there is no data for or against this conjecture.

made up with backup auctions and the loss was quite balanced between the two rank groups, we believe sample selection is not a big issue. As shown below, professional grading suggests that the highest reported professional grade is similar across the two groups. So even if we were subject to the winners' curse, it should not affect our conclusion regarding the *relative* quality differences between the two groups.²⁸ Moreover, our overbid strategy to ensure winning prevents us from drawing strong conclusion from the prices we paid for our sample cards. To avoid potential contamination in the price data, most analysis we present below focus on the quality data from the experimental sample, and compare them to the prices observed in our eBay watch.

We stress that the separation of best- and median-ranked cards serve two purposes: first, it allows us to observe variation in seller claims. Second, by bidding on the best- and median-ranked copies of the same card in the same week, we employ a pairwise comparison that controls for any card-specific or time-specific variations in consumer tastes or seller claims. For example, over 40,000 copies of Griffey have been graded 10, so the best claim for an ungraded Griffey is always 10 or 9.5. Subsequently, the median claim is often as high as 9. On the other hand, no 10 exists for Jeter, so the best claim for an ungraded Jeter is about 9 and the moderate claim lingers around 8. This forces us to design the sample by *relative* rather than absolute differences in seller claims.

In total, we paid for 107 cards online and received 96 of them before April 2, 2002, the date when we sent our purchases to BGS for professional grading.²⁹ Of the remaining 11 cards, 7 were received after April 2, 2002 and the other 4 were never received. These 7 late cards were graded in the second round. Because including or excluding these 7 late cards does not affect our conclusion in any significant way, all results reported in this paper are conditional on the cards graded in the first package.³⁰

²⁸The relative difference could be affected if the distribution of signals were different in the two groups. We thank Editor Ariel Pakes for pointing out this possibility. Since an equal variance test between the two groups cannot reject the null hypothesis, this concern is unlikely to affect our results.

²⁹Overall, we obtained 126 cards from retail markets for a related study, and paid for 107 cards online. Because the Jeter and the Rodriguez were hard to find in retail markets, we stopped purchasing them online after February 3, 2002 to maintain sample comparability in the other study. For all weeks afterwards, we doubled our online purchases for the other three cards.

³⁰Of the 7 late cards, 3 were best-ranked and 4 were median-ranked. Including these 7 cards will reduce the likelihood of counterfeit or default in our online purchases from 0.11 to 0.1028. Some store owners are concerned that grading companies may give better grades to cards sent in large packages. In order to rule out any potential grading discrepancy between the two rounds, we choose to exclude the 7 late cards from the reported tables. Should there be any grading bias, it should be the same for cards sent in the same package and therefore does not affect any comparison reported in this paper. We chose BGS over the other grading companies because BGS uses the most detailed grading scale (1 to 10 with half grade increments) and is the only major service that provides

5 Experimental Results

We now present two sets of experimental results: one on true quality and its correlation with seller claims, and one on reputation and quality.

5.1 Experimental Results On Quality

Table 4.1 summarizes the card quality data from the experiment. The first two columns report summary statistics for the best- and median-ranked cards within online purchases. Because the best- and median-ranked cards are paired up by card type and purchase week, column 3 reports their difference based on the *pairwise* comparison. In practice, the number of best-rank cards is slightly larger because several best-rank cards were sold in a package of multiple copies (of same card same claim). The pairwise comparison is conducted by controlling for card-week fixed effects. The full distribution of grades for each of the two samples are shown in Table 4.2.

Despite being forced to pay more for each best-ranked card, we find the average quality of cards from the best-rank group and median-rank group virtually indistinguishable. Assigning zero quality to defaults and counterfeits, the best-rank group’s average quality (6.06) was about a full grade lower than that of the median-rank group (7.05). If we limit the comparison to only authentic cards successfully delivered, the card quality from the two rank groups is extremely close (7.30 versus 7.37). Recall from the eBay watch data that average winning prices were 33.4% more for ungraded cards claiming 9 and 51.1% more for cards claiming 10 (as compared to those with moderate claims (7-8)). It is clear that higher winning prices in these eBay auctions for sportscards do not necessarily indicate higher true physical quality.³¹

The real difference between the two groups is the higher default and counterfeit rates that accompanied high seller claims: of the 11 fraudulent transactions, 9 came from the best-rank group and 3 of the 4 defaults were best-rank cards whose sellers claimed gem mint (10). A more detailed statistic is the absolute difference between seller claims and true card quality. If we denote undelivered or counterfeit cards as zero grade, the best-rank group overclaimed more than 3 grades, doubling that of the median-rank group. Conditional on authentic deliveries, overclaims from the best-rank group are still 0.84 grade higher (significant with 95% confidence).

sub-grades by centering, corners, edges and surface. In a separate study, Jin, Kato and List (2005) compare the grading practice of PSA, BGS and SGC. We find that BGS provides the most consistent grading among the three.

³¹Our overbid strategy to ensure winning prevents us from drawing strong conclusions from the prices we paid for our sample cards. That is why we focus on the quality difference from our experimental data, and compare that to the price difference observed in our eBay watch.

This is a dramatic difference given that a card of grade 7 is valued 50% less than a 8 and 70% less than a 9. Even if we include delivery service as part of the total bundle of characteristics of the product being purchased (Spence 1977), it appears that price premiums for high claim cards are not securing any quality benefit for the purchasers, who in fact are more frequently defrauded.

5.2 Experimental Results On Seller Reputation

Seller reputation is one of the market mechanisms that potentially alleviate the information asymmetry between online buyers and sellers. In the field experiment we are able to relate seller reputation to true seller quality thus examining the roots of reputation theories: if reputation holds in any meaningful way, we should observe a positive correlation between seller reputation and seller performance.

Table 5.1 summarizes log of seller net rating by a number of subsamples. First of all, reputable sellers are less likely to default or deliver counterfeit copies, suggesting that reputation is a meaningful signal for default and counterfeit risks. Based on this result, it is not surprising online buyers in our eBay watch data were more willing to bid on items offered by reputable sellers. More interestingly, seller reputation also differs by the magnitude of self claims. Although the average seller ratings appear quite similar in best- and median-ranked groups, sellers claiming mint or better were on average less reputable than the other sellers. Recall that the highest claiming sellers were responsible for most of the outright cheating. This explains why seller ratings and seller claims were negatively correlated in the eBay watch data. While seller rating is a good signal of seller performance (in terms of default and counterfeit risks), claiming 9 or 10 is in fact the strongest negative signal of dishonesty. Unfortunately, not every online buyers understood the true signaling effect of high claims. That is probably why in our eBay watch data experienced buyers tend to avoid ungraded cards, especially those with extremely high claims. Such buyer selection, if it exists, is consistent with the suspicion that claiming sellers may target inexperienced buyers.

Table 5.2 reports regression results on seller ratings and seller claims. There are three dependent variables: whether the online purchase is a default or counterfeit, the true card quality (defaults and counterfeits are coded zero), and log of the prices we paid as percentage of Beckett low book. The key independent variables are seller self grade (as a continuous variable), whether the seller claimed Mint or better (as a binary variable), log of seller net ratings, and whether the seller has any negative feedback. All control for card type and purchase time.

Regression results are consistent with the data summaries: reputable sellers are less likely to commit default or deliver counterfeits, and card quality for authentic deliveries is not correlated to seller ratings. This may explain a puzzle in the eBay watch data (and the existing literature): buyers are more willing to buy items from reputable sellers because they want to reduce the risk of default and counterfeit. But conditional on completed auctions, buyers are not willing to pay more to reputable sellers because their authentic deliveries are not of better quality.³² A second finding is that seller claims mean nothing once we control for seller reputation. Nevertheless, we were forced to pay more for seller claims, just like the winners in our eBay watch data. This points to buyer naiveness with regard to seller claims.

Note that negative feedbacks do not imply any difference in default/counterfeit risks or card quality. We believe this is due to a unique feature of Internet auctions: any eBay seller can costlessly switch his anonymous identity. This implies that some sellers *choose* to continue with negative feedbacks because the negatives do not harm them too much. As a result, we can only observe the impact of negative feedbacks on this selected group.

Costless seller identity switching also makes it possible for fraudulent sellers to game the eBay feedback system. One of the four sellers who did not deliver in our real purchases provides a good example. That seller – *eddie27g* – completed delivery for seven auctions to receive positive feedback, establishing a good seller reputation. He then waited more than three weeks to allow the eBay database to purge those completed transactions, preventing any future buyers from observing the value of the items he sold. After the eBay database erased the transaction records, 25 of the next 27 auctions he put on eBay resulted in fraud, leading to an explosion of complaints from buyers. All of these auctions occurred within one week, after which the seller abandoned the seller id that today according to eBay is “not a registered user.” Another seller who sent us a counterfeit card followed a similar strategy. That seller set up a web site to make himself look like a dealer and claimed he had been in the sports card business for 10 years. In his feedback file, he built up a net rating of 128 positives before a run of 31 negatives at the end resulting from fraud cases. We could only verify the last 10 feedbacks, which were all negatives, because he made his feedback file private and abandoned the seller id.

These fraudulent sellers may only account for a small proportion of trading activity, but their behavior introduces a great deal of noise into the seller reputation system. Given the fact that a seller receiving many complaints can easily switch to a new eBay identity, uninformed

³²It is also possible that buyers choose their bidding target and bidding amount after integrating the risk of default and counterfeit with the expected quality conditional on authentic delivery. In that case, reputable sellers should enjoy a higher probability of sale and higher winning prices, which is inconsistent with the pattern observed in the real data.

newcomers may never have access to the complaints to update their beliefs about that seller. This probably explains why there are less than 1% negative feedbacks in the whole eBay system (Resnick and Zeckhauser 2002), especially if we only count registered users.

6 Discussion

This section attempts to integrate the experimental results with the findings from the eBay watch. In doing so, we discuss a number of explanations and assess their relative strength in explaining the non-classical data patterns in eBay.

The most striking finding is that buyers in our eBay watch paid 33-51% more to cards with outrageous quality claims, but our field experiment suggests that such seller claims mean nothing but an elevated risk of default or counterfeit. This clearly violates the monotone relationships that the classical theory predicts between price, advertising and quality. However, as we have detailed in Section 5, this finding could explain some paradoxes that arise in our eBay watch: (1) experienced buyers avoid cards with highest claims probably because they better understand the true meaning of the claims; (2) sellers with better reputations are less likely to make bold claims probably because these claims entail cheating and tarnish reputations; and (3) sellers making high claims are not foregoing profits by not grading their cards before selling them. They probably do not actually believe they have high quality cards and are simply making statements that they know will maximize profits from their auctions.

However, two puzzles remain: why some buyers are willing to pay for the empty claims, and why not all sellers make empty claims. Underlying both puzzles are two more general questions: what accounts for the mismatch between the classical theory and the Internet auction settings? And why is eBay by and large successful? We now turn to discuss each in more details.

Why are some buyers willing to pay more for empty claims?

One obvious explanation is that some buyers do not know the true meaning of seller claims. If so, they are not only uninformed about the quality of a specific seller (or the quality of the goods sold by a specific seller), but also hold an incorrect belief about how true quality varies by seller claims. Before concluding this as the most likely explanation, we examine a couple of alternatives.

Buyers could still be acting rationally if their preferences are convex enough, the true quality of cards with high claims is more dispersed, and not all owners of truly high quality

cards realize their cards are worth grading. In fact, we did obtain a Mint 9 card from our best-ranked group, which suggests that either the seller was not experienced enough to judge the quality of the card or even experienced sellers cannot perfectly predict the grade a professional grader would assign to the card. Our experiment confirms the possibility that buyers may have a non-zero chance of getting a high quality ungraded card online. This non-zero probability plus the convex price schedule may motivate the buyer behaviors that we observed in reality.

To consider this alternative explanation, we need to compute expected values instead of expected quality. Taking the Beckett price schedule as a proxy for card values, the bottom of Table 4.1 translates card grade into Beckett low price and calculates the average card value in each rank group.³³ Regardless of whether we include defaults and counterfeits and whether we account for possible eBay reimbursement for defaults, the average card value is not significantly higher in the best-ranked group. This conflicts with the enormous price premiums that buyers were willing to pay for high claims in the eBay watch data.

An example is more illustrative. In May 2001, a typical PSA 10 Griffey was sold online at price \$1,450 and the average final price of ungraded Griffey's claiming gem mint 10 was \$94.26. Consider two options for a risk neutral buyer: he may pay \$1,450 and win a card that has already been graded as gem mint; or he may spend all \$1,450 on ungraded Griffey's claiming gem mint, send all delivered copies for grading at a constant grading cost of \$8 per copy and claim insurance for all undelivered copies from eBay.³⁴ In equilibrium, the two options must have equivalent expected values for a risk neutral buyer.

If buyers expected the same quality distribution from all cards claimed gem mint 10 cards as we discovered in our experiment, they must have expected a combined default and counterfeit risk less than 0% (It turns out to be -42.7%). If buyers correctly believe the combined risk for Griffey is 38.9% as discovered in the experiment, they must have expected the ungraded cards for sale online had a quality distribution similar to all Griffey's PSA graded between January 1998 and December 2001. That distribution had 1.83% gem mint 10s, 30.78% mint 9s, and

³³Beckett does not report prices for graded cards below 7 because "lower grade cards from the modern era draw little, if any, premium above non-graded cards in the same condition and are rarely seen in circulation." Instead, Beckett quotes full and low retail prices of ungraded cards for the equivalents of NM-MT (8) condition. For ungraded cards below the NearMint-Mint quality, Beckett provides guidelines for determining their prices from the book quotes. For example, for any card printed during 1981 to 1989 in Excellent (5) condition, Beckett recommends using 15-25% of book prices. Beckett provides a slightly different price schedule for cards printed after 1990, recognizing that older cards are less likely to be in good condition and were made out of inferior materials. We use the middle point of each range as the true discount value for each condition.

³⁴With a \$25 deductible, the Fraud Protection Program in eBay reimburses up to \$200 per item for buyers who did not receive any delivery after 30 days of payment. Therefore, for every default, the buyer can get $\$94.26 - \$25 = \$69.26$ back from eBay.

49.69% near mint-mint 8s. Though we may have over sampled from the high fraud best-ranked group, the true quality distribution of neither of our experimental samples comes anywhere close to producing 30% Mint 9 grades.

Another possibility is that buyers' utility function is more convex than the price schedule. Such risk loving buyers may derive psychological utility from the gambling aspect of buying ungraded cards and place high value on the small chance of receiving gem mint cards. For this explanation to hold, these buyers cannot just be eBay lovers, as they could have derived the same thrill from winning a bid in either the median-ranked or the best-ranked group. Rather, they must believe that the probability of getting high quality cards, however small it is, is higher in the best-ranked group. If they are fully informed of the risk, they must also understand that the presence of even one dishonest seller online could theoretically create many dishonest seller identities. The ease of creating anonymous identities on eBay would drive down the probability of receiving a Gem Mint 10 down to zero. Therefore, any reward for non-verifiable high claims requires extreme convexity in preferences.

These alternative explanations are not entirely impossible, but given the arguments above and the institutional reasons we elaborate later on, we believe the most likely explanation is some naive buyers underestimate the risk associated with high seller claims. Note that this conclusion does not imply the entire market is irrational or naive.³⁵ In fact many institutions in the online baseball card market appear to be effective: buyers interpret seller reputation as a signal of reliability, experienced buyers tend to move to the safer graded market, and prices reflect a large quality difference between graded and ungraded cards. What is puzzling is why the ungraded market has not broken down to a paradise of cheating, given the potential profits a seller could earn by making bold claims.

Why don't all sellers make empty claims?

One obvious explanation is that not all sellers realize the potential benefits from making

³⁵To measure the size of the problem, note that in our eBay watch data, 35% of transactions were on ungraded cards, and 40% of these ungraded cards had seller claims of mint or better. In our online experimental sample, 66% claimed mint or plus and 11% were fraudulent. Of the 11 defaults/counterfeits, 9 had claimed mint plus. Assuming no fraud on graded cards, these numbers imply roughly 3% default/counterfeit risk and 11% empty-claims risk in the whole online market of baseball cards. Since the retail market is not free of counterfeits, the online fraud rate seems to imply a trivial problem. One may argue that uninformed buyers underestimating risk is an inevitable phenomenon in a new market, eventually buyers would learn their lesson, and the market would converge to a desirable long-run equilibrium. But on the other hand, prices may be misleading and the eBay institutions (as elaborated later on) prevent some important information from flowing to new buyers. It is unclear what long run consequences would arise from the current level of buyer naivete. The topic itself warrants future study.

empty claims and starting fresh afterwards. We believe this is unlikely. The transaction record is publicly available to everyone with access to the Internet. In our eBay watch data, the price premium for bold claims stand out from the simplest summary statistics. Even an inexperienced seller could search all completed auctions in the last three calendar months for the same item they are selling, and mimic the most successful auction listings without fully understanding the strategies involved. However, it is not the inexperienced seller who avoids making claims; the data indicates it is typically an experienced seller with a high rating that refrains from making claims. If relative newcomers like us can figure it out, it is hard to imagine that those experienced sellers are not aware of the price pattern.

Nevertheless, experienced sellers are less likely to make bold claims probably because they are more concerned with future sales. If a seller expects to use the same seller identity to sell many more items in the future, he may be reluctant to make bold claims, as transactions completed under bold claims may hurt future ratings. Although the positive effect associated with seller ratings is not very strong in our data, it may add up to a significant amount if the trading volume is high.

Moreover, our analysis may not capture all the returns to a high eBay ratings. In an extreme example, if an eBay seller has a retail store and gives out his eBay id to retail customers, a positive eBay rating may generate positive spillover to his retail business.³⁶ In a more realistic example, online auction sellers in our eBay watch data might be involved in selling many other items on eBay. It is possible that reliable and on-time delivery is more important for some classes of goods, and sellers must have some minimum level of reputation before buyers will even consider bidding for those items.³⁷ If a seller is engaged in listing auctions for both sportscards and those “minimum credibility required” goods, then making empty claims may be suboptimal because it will affect their business in other markets. Using data from a single market, we cannot quantify the differential return of eBay ratings across multiple markets.

The third possibility is sellers differing in the costs of starting fresh. This may involve the conscience cost in abandoning a bad history and pretending to be new and fresh, or aversion to the legal and moral risk of being caught. Our data does not allow us to distinguish among these possibilities.

Why do buyer naivet and seller cheating exist on eBay?

³⁶In our eBay watch data, about 28% sellers seem to be card dealers (by self claims or the amount of sportscards listed simultaneously on eBay). However, the claim of dealership is difficult to verify.

³⁷For instance, buyers may be drawn by empty claims of quality for a used DVD box set worth fifty dollars, but they may be a lot more careful and concerned about being defrauded when purchasing a top of the line plasma television worth five thousand dollars.

We speculate that buyer naivet and seller cheating exist on eBay because eBay institutions differ from a typical reputation system in two ways.

First, eBay ratings are universal. Every feedback is counted equally, regardless of the amount of the transaction and the individual's buying or selling position in previous trades. As a result, it is easy to build up an eBay rating by buying cheap items and use that reputation to sell expensive items. This reduces the precision of eBay ratings as a measure of seller type and facilitates cheating.

Another consequence of universal ratings is that eBay ratings can only convey the information universal across all previous transactions, i.e. whether the individual has honored completed auctions by paying the full amount as a buyer or delivering the good on sale as a seller. In practice, eBay saves feedback comments indefinitely but users cannot view the auction being commented on after 90 days have passed. Random inspection reveals that very few comments mention the item in the comment itself. Since the actual quality of delivered items is not always comparable across categories, a previous buyer's complaint of mediocre item quality is at most a noisy measure of what quality the seller would deliver in the future. By this design, it is not surprising that in our data seller ratings are a good indicator of authentic delivery, but do not correlate with the actual quality conditional on authentic delivery. The wide use of professional grading online also suggests that eBay reputation is less effective than professional grading in handling information pertaining to the actual quality of delivered items.

The second unique feature is that eBay allows costless switching between anonymous identities. A cheater can abandon the identity that receives the negative feedback and switch to a clean slate. A cheater can also open multiple accounts and create positive feedbacks between them. Though it may be difficult to generate 1000 positive feedbacks in this way, the possibility of self-producing at least low levels of positive feedbacks plus the escape from negative ratings mean cheaters could easily mingle with new honest sellers in a lower range of positive ratings. As a result, positive ratings contain much noisier information about seller types than they could in the absence of costless switching and multiple identities.

A number of theories have examined situations where past behavior does not stick when a player changes its anonymous identity (Friedman and Resnick 2001) or starts a new relationship (Ghosh and Ray 1996, Kranton 1996, and Watson 1999). They all find that some initial costs are needed to encourage players to behave under current identities. These costs can be a slow start in developing a long run relationship, newcomers receiving poor treatment from established players, or, like in our context, lower likelihood of sale for less reputable sellers. These theories imply that loopholes in eBay rating system do not invalidate seller reputation; they just make

it more difficult for the reputation system to function.

Specifically, the noise in eBay's reputation system translates into information costs for buyers. Because it is costless to shed bad histories, lessons learned by one naive buyer today cannot reach the other naive buyers in the same market, nor can they reach new buyers entering tomorrow. Consequently, both current and future buyers have difficulty updating their beliefs, not only about whether a specific seller is honest, but also about the actual distribution of dishonesty. It is the latter that potentially makes the naiveté persist over time, even if most naive buyers are able to actively process whatever information is available to them.

Why is eBay by and large successful?

Despite the problems discovered in our field experiment, eBay is still the most successful site for Internet auctions. As of today, the eBay community includes more than a hundred million registered members, making it the most popular shopping destination on the Internet.³⁸

The eBay success is not necessarily inconsistent with our findings. In fact, we can think of two reasons for the success: first, eBay buyers learn over time. As shown in our data, experienced buyers tend to avoid the ungraded market and rely on professional grading to minimize the risk associated with incredible seller claims. The use of professional certificates is particularly popular in the types of merchandises that are likely to suffer from the information problems, for example, collectible coins, stamps, jewelries (especially diamonds), antiques, art works, and motor vehicles.

More importantly, eBay has been constantly fighting with the problems mentioned in this paper.³⁹ To address the fundamental information asymmetry between buyers and sellers, eBay engages in a number of activities to facilitate the usage of professional certificates. For example, eBay allows logos of professional certifiers to be posted in the auction description. More aggressively, eBay is implementing a program that explicitly indicates eBay preference of prestigious certifiers over non-reputable certifiers.

Furthermore, after our data collection period, eBay has adopted numerous methods to amend the loopholes in their rating system. To go beyond universal ratings, eBay started to allow buyers (in some countries other than the U.S.) to sort seller ratings. In doing so, buyers can differentiate parts of the seller history and combine all the detailed feedbacks associated with

³⁸Source: eBay company overview accessed at <http://pages.ebay.com/aboutebay/thecompany/companyoverview.html> on Nov. 13, 2005.

³⁹The following examples are based on eBay Senior Counsel Joe Sullivan's presentation at the "Roundtable on the Economics of Internet Auctions" organized by the Federal Trade Commission, Oct. 2005.

negative ratings. To address the problems arising from costless identity switch, eBay introduced “ID Verify” which allows a seller to get his identity verified by a third-party company. eBay also implements an internal program to identify high-risk sellers and remove them from the eBay system. Paypal, now a subsidiary of eBay, has recently raised its insurance limit from \$500 to \$1000. All these actions intend to bring the eBay institution closer to the classical reputation system.

One may wonder why eBay did not conduct a study like ours and published the results to all the (potential) eBay users. Given the steps that eBay has undertaken to ensure trading safety, most likely eBay has been aware of the problems for a long time. However, in any official document, eBay only reports the percentage of fraudulent transactions that got reported to its Buyer Protection Program (which is usually less than 0.1%). We suspect this reluctance is due to two reasons: first, any scientific study like ours will focus on a specific category and it is difficult to generalize category-specific results to the whole range of eBay merchandises. Second, although consumers are keen to know the actual fraud rate in eBay auctions, reporting this number may significantly change the buyers’ prior about eBay safety thus reducing the number of buyers (especially new ones) who are willing to use eBay. From a profit point of view, this strategy is not as appealing as keeping the uncertainty in the air and fighting frauds inside eBay.

7 Conclusion

This paper examines the link between price, quality, seller advertising and seller reputation in Internet auctions. Unlike most empirical studies that infer the effectiveness of online seller ratings from price data only, we purchased ungraded baseball cards and had them professional evaluated. The field experiment, plus the eBay watch data gathered before the experiment, gives us a rare opportunity to link seller ratings, prices, and seller claims with true quality.

We find that some relationships are inconsistent with the classical theory. In particular, buyers in our eBay watch paid 33-51% more for cards that claimed quality of mint or better, but high claims result in more fraud (i.e. default or sending counterfeits) and no better card quality in our field experiment. Moreover, reputable sellers are less likely to make bold claims and less likely to commit fraud. But conditional on authentic delivery, cards from reputable sellers are no better in quality. Correspondingly, buyers are more willing to bid on items listed by reputable sellers and more likely to complete such auctions. However, conditional on completion, seller ratings have no significant impact on the winning prices.

These findings suggest inconsistency between the classical theory and the institution of Internet auctions. Specifically, universal ratings and costless switching of anonymous identities on eBay clearly violate the most basic assumptions of classical reputation theories. We suspect such violations contribute to the non-classical data pattern in eBay. While the causal link between data and institutions warrants more research, our findings suggest that assuming price is a monotone function of quality is misleading, especially in a market where quality is important but no verifiable information about quality is available to consumers. To what extent this conclusion applies to markets other than Internet auction is still an open question, but our findings call for a closer look at the non-parametric implications of auction theory before applying it to explain auction outcomes.

8 References

- Akerlof, George (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 222: 488-500.
- Ba, S. and Pavlou, P.A.(2002): "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior." *MIS Quarterly* 26(2002): 243-268.
- Bagwell, Kyle and Michael H. Riordan (1991): "High and Declining Prices Signal Product Quality," *The American Economic Review* 81(1): 224-239.
- Bajari, Patrick and Ali Hortacsu (2003a): "The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions," *Rand Journal of Economics* Vol. 3, No. 2, pp. 329-355.
- Bajari, Patrick and Ali Hortacsu (2003b): "Online Auctions: A Review of Empirical Findings," *forthcoming Journal of Economic Literature*.
- Brown, Jeffrey R. and Austan Goolsbee (2002): "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *Journal of Political Economy* 110(3): 481-507.
- Dana, James D. (1994): "Learning in an Equilibrium Search Model," *International Economic Review* 35(3): 745-771.
- Darby, Michael R. and Edi Karni (1973): "Free Competition and the Optimal Amount of Fraud," *Journal of Law and Economics* April 1973, pp. 67-88.
- Diamond, P.A. (1971): "A Model of Price Adjustment," *Journal of Economic Theory*, 3: 156-168.
- Eaton, D.H. (2002): "Valuing Information: Evidence from Guitar Auctions on eBay," working paper.
- Engelberg, Joseph and Jared William (2005): "A License to Shill: Incremental Bidding and eBay's Proxy System," Kellogg School of Management working paper.
- Epple, Dennis and Artur Raviv (1978): "Product Safety: Liability Rules, Market Structure and Imperfect Information," *American Economic Review* 68: 80-95.
- Friedman, E. and P. Resnick (2001). "The Social Cost of Cheap Pseudonyms." *Journal of Economics and Management Strategy* 10(2): 173-199.

- Genesove, David (1993): "Adverse Selection in the Wholesale Used Car Market," *The Journal of Political Economy* Volume 101, Issue 4 (Aug. 1993) 664-665.
- Ghosh, P. and D. Ray (1996): "Cooperation in Community Interactions without Information Flows," *Review of Economic Studies* 63: 491-519.
- Goolsbee, Austan (2000): "In a World Without Borders: The Impact of Taxes on Internet Commerce," *The Quarterly Journal of Economics* May 2000.
- Gorton, Gary (1996): "Reputation Formation in Early Bank Note Markets," *The Journal of Political Economy* 104(2): 346-397.
- Hubbard, Thomas N. (2002): "How Do Consumers Motivate Experts? Reputational Incentives in an Auto Repair Market," *Journal of Law and Economics*, Oct. 2002, 437-468.
- Jin, Ginger Z. and Andrew Kato (2005): "Dividing Online and Offline: A Case Study," *University of Maryland Working Paper*.
- Johnson, Carrie A. (2003a) "Q1 2003 Online Sales: Bucking the Retail Trend," *Forrester Research* (April 29, 2003), accessed at <http://www.forrester.com/ER/Research/Brief/0,1317,16779,FF.html>.
- Johnson, Carrie A. (2003b) "The Future of Amazon.com," *Forrester Research* (September 2003), accessed at <http://www.forrester.com/ER/Research/Report/Summary/0,1338,16873,00.html>.
- Klein, Benjamin and K. B. Leffler (1981): "The Role of Market Forces in Assuring Contractual Performance." *The Journal of Political Economy* 89(4): 615-641.
- Kranton, Rachel E. (2003): "Competition and the Incentive to Produce High Quality," *Economica* 70: 385-404. es
- Linnemer, Laurent (2002): "Price and Advertising as Signals of Quality When Some Consumers Are Informed," *International Journal of Industrial Organization* 20(7): 931-947.
- Lucking-Reiley, David (2000): "Auctions on the Internet: What's being auctioned and how?" *Journal of Industrial Economics* 48(2): 227-252.
- Lucking-Reiley, David; D.Bryan, N.Prasad and D. Reeves (2001): "Pennies from eBay: the Determinants of Price in Online Auctions" *University of Arizona Working Paper*.
- Melnik, Mikhail I. and James Alm (2002), "Does a Seller's eCommerce Reputation Matter?" *Journal of Industrial Economics* 50(3): 337-349.
- Nelson, Phillip (1970): "Information and Consumer Behavior," *The Journal of Political Economy* 78(2): 311-329.
- Resnick, Paul and Richard Zeckhauser (2002): "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System," *The Economics of the Internet and E-Commerce* M. R. Bayes, eds. Amsterdam, Elsevier Science, 11: 127-157.
- Resnick, Paul; Richard Zeckhauser; John Swanson; and Kate Lockwood (2003): "The Value of Reputation in eBay: A Controlled Experiment" working paper, accessed at <http://www.si.umich.edu/presnick/papers/postcards/>.
- Rob, R. (1985): "Equilibrium Price Distributions," *Review of Economic Studies* 52: 457-504.
- Robert, Jacques and Dale O. Stahl (1993): "Informative Price Advertising in a Sequential Search Model," *Econometrica* 61(3): 657-686.

- Roth, A. and Ockenfels, A. (2002) "Last Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon on the Internet" *American Economic Review* 92(4).
- Salop, Steven and Joseph Stiglitz (1977): "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion," *Review of Economic Studies* 44: 493-510.
- Shapiro, Carl (1983): "Premiums for High Quality Products as Returns to Reputation." *Quarterly Journal of Economics* 98(4): 659-680.
- Spence (1977): "Consumer Misconceptions, Product Failure and Producer Liability." *The Review of Economic Studies* 44(3): 561-572.
- Stahl, Dale O. (1989): "Oligopolistic Pricing with Sequential Consumer Search," *The American Economic Review* September 1989, 700-712.
- Watson, J. (1999): "Starting Small and Renegotiation," *Journal of Economic Theory* 85(1): 52-90.
- Wolinsky, Asher (1983): "Prices as Signals of Quality," *The Review of Economic Studies* 50(4): 647-658.
- Wolinsky, Asher (1986): "True Monopolistic Competition as a Result of Imperfect Information," *The Quarterly Journal of Economics* 101(3): 493-512.

Table 1: Summary of eBay watch data

Variable	Full Sample					Completed Transactions	
	OBS	Mean	Std Dev	Min	Max	OBS	Mean
Final Price	1124	179	510	0	13500	776	165.5
Full Book Price	950	379.6	677.6	50	2500	776	370.6
Low Book Price	950	226.3	412.6	25	1500	776	221.3
82 Topps Cal Ripken	1124	0.03	0.18	0	1	776	0.03
85 Topps Mark McGwire	1124	0.12	0.33	0	1	776	0.1
89 Upper Deck Ken Griffey	1124	0.69	0.46	0	1	776	0.72
93 SP Derek Jeter	1124	0.09	0.29	0	1	776	0.09
94 SP Alex Rodriguez	1124	0.06	0.24	0	1	776	0.06
Dummy=1 if the transaction is incomplete	1124	0.19	0.39	0	1	776	0
Number of Bids	1124	9.65	7.3	0	37	776	11.1
Shipping cost	929	4.06	2.68	0	25	659	3.95
Dummy=1 if no shipping information	1124	0.17	0.38	0	1	776	0.15
Length of auction (days)	1124	5.66	2.25	0	10	776	5.61
Dummy=1 if auction ends in weekend	1124	0.42	0.49	0	1	776	0.41
Dummy=1 if auction ends in prime time	1124	0.26	0.44	0	1	776	0.25
Dummy=1 if accept Master or Visa Card	973	0.24	0.43	0	1	674	0.23
Dummy=1 if accept Amer. Ex or Discover	973	0.1	0.3	0	1	674	0.1
Dummy=1 if accept money order	973	0.98	0.13	0	1	674	0.99
Dummy=1 if accept personal check	973	0.65	0.48	0	1	674	0.68
Dummy=1 if accept Paypal alike	973	0.31	0.46	0	1	674	0.34
Dummy=1 if accept cash	973	0.01	0.09	0	1	674	0.01
Dummy=1 if no payment method information	1124	0.13	0.34	0	1	776	0.13
Dummy=1 if dealer	1124	0.28	0.43	0	1	776	0.29
Dummy=1 if has a buy_it_now price	1124	0.11	0.32	0	1	776	0.03
Dummy=1 if has a reserve price	1124	0.14	0.35	0	1	776	0.09
Dummy=1 if post a legible scanned picture	1124	0.72	0.45	0	1	776	0.72
Dummy=1 if post an illegible scanned picture	1124	0.11	0.31	0	1	776	0.12
Dummy=1 if no scanned picture	1124	0.18	0.38	0	1	776	0.16
Dummy=1 if graded	1124	0.67	0.47	0	1	776	0.65
Dummy=1 if graded 7	1124	0.07	0.26	0	1	776	0.06
Dummy=1 if graded 8	1124	0.26	0.44	0	1	776	0.29
Dummy=1 if graded 9	1124	0.14	0.35	0	1	776	0.16
Dummy=1 if graded between 9 and 10	1124	0.07	0.25	0	1	776	0.02
Dummy=1 if graded 10	1124	0.12	0.33	0	1	776	0.12
Seller estimated grade for ungraded cards only	201	8.86	0.89	7	10	143	8.91
Dummy=1 if does not have self-reported grade (ungraded cards only)	372	0.46	0.5	0	1	269	0.47
Dummy=1 if report serial number (graded cards only)	752	0.68	0.47	0	1	507	0.68
Seller reputation: net ratings	1064	612.6	999.4	-1	8440	730	638
Natural Logarithm of Net Rating +1	1064	5.26	1.8	0	9.04	730	5.34
Seller reputation: dummy=1 if there has been any negative feedback	1070	0.37	0.48	0	1	730	0.38
Seller reputation: the number of negative feedback	1070	1.51	4	0	46	730	1.51
Missing negative ratings	1124	0.05	0.22	0	1	776	0.06

Note: Some variables have missing values for some observations, either because the variables are not applicable (e.g. seller claims only apply to ungraded cards), or because they are missing in the listing page (e.g. not all ungraded card sellers made quality claims).

Table 2: eBay watch data - summary of auction completeness and winning prices

		Full Sample			Completed Transactions		
		OBS	Auction completed?	Net seller rating	OBS	Winning price (\$)	Winner's net buyer rating
Graded	Near Mint or below (<8)	88	72.73%	382.63	64	51.02	207.19
	Near Mint-Mint (8 or 8.5)	290	83.79%	706.89	243	75.47	266.45
	Mint (9 or 9.5)	237	82.70%	739.24	196	166.03	200.20
	Gem Mint (10)	137	78.10%	874.79	107	690.23	166.64
	Total	752	81.25%	709.73	610	209.84	221.70
				(1088.63)	(362.16)	(366.77)	
Ungraded	No Claim	171	83.63%	757.05	143	54.16	135.69
	Claimed Near Mint (7 or 7.5)	24	75.00%	1271.67	18	46.47	152.44
	Claimed Near Mint - Mint (8 or 8.5)	29	82.76%	2444.07	24	49.14	243.43
	Claimed Mint (9 or 9.5)	105	78.10%	290.07	82	70.01	137.58
	Claimed Gem Mint (10)	43	83.72%	320.81	36	90.60	81.45
	Total	372	81.45%	739.53	303	61.93	140.03
				(1980.40)	(26.53)	(272.70)	

Note: standard deviation in parentheses.

Table 3: Regression results of eBay watch data

Dep. Var.	=1 if complete the auction Probit (coefficients)		ln (winning price / low book price) Controlling for propensity scores (coefficients)					
=1 if Graded	2.286	*			1.813	***		
	(1.261)				(0.291)			
=1 if Graded below 9			0.202				0.308	***
			(0.337)				(0.107)	
=1 if Graded 9			0.268				0.159	
			(0.357)				(0.112)	
=1 if Graded 10			0.223				-0.398	***
			(0.376)				(0.133)	
=1 if Ungraded with no self grade	2.417	**	0.372		1.740	***	0.151	**
	(1.205)		(0.254)		(0.277)		(0.059)	
Seller self grade * ungraded	0.248	*			0.212	***		
	(0.137)				(0.031)			
=1 if self grade 9 * ungraded			0.033				0.334	***
			(0.283)				(0.068)	
=1 if self grade 10 * ungraded			0.769	**			0.511	***
			(0.378)				(0.081)	
ln(seller ratings+1) * graded	0.084	*	0.097	**	-0.026		-0.005	
	(0.049)		(0.049)		(0.017)		(0.015)	
=1 if any negative ratings * graded	0.099		0.097		-0.015		-0.038	
	(0.158)		(0.159)		(0.050)		(0.043)	
ln(seller ratings+1) * ungraded	0.178	***	0.169	***	-0.018		-0.014	
	(0.058)		(0.052)		(0.015)		(0.015)	
=1 if any negative ratings * ungraded	-0.649	***	-0.683	***	-0.008		0.002	
	(0.226)		(0.222)		(0.057)		(0.055)	
=1 if scanned picture is clear	0.307	**	0.357	**	0.031		0.074	*
	(0.140)		(0.150)		(0.046)		(0.040)	
=1 if scanned picture is unclear	0.757	***	0.714	***	-0.089		-0.053	
	(0.235)		(0.233)		(0.073)		(0.063)	
=1 if the seller claims to be a card dealer	-0.146		-0.202		0.015		0.011	
	(0.170)		(0.173)		(0.049)		(0.045)	
Predicted probability of sale					0.339		0.318	
					(0.411)		(0.426)	
Obs	1124		1124		778		778	
R-Square	0.461		0.463		0.167		0.392	

Notes: All regressions control for a full set of card identity dummies, payment methods, shipping options, length of auction, whether the auction ended on a weekend, whether the auction ended in primetime, where there was a public or secret reserve price and whether there was a Buy It Now option. The number of observations in the price regression is less than the total number of completed auctions, because Beckett Baseball Monthly does not report book prices for some cards graded below 8 (due to low trading volume). Fluctuation of card value is controlled by Beckett low book price (by card-month). Standard errors in parentheses. ***p<0.01, ** p<0.05, *p<0.1, two-tail.

Table 4.1: Experimental Results Summary I - Price and Quality

	Best Ranked	Median Ranked	Best vs. Median Ranked (pairwise comparison)		Online Overall
N	53	47			100
N of defaults	3	1			4
N of counterfeits	6	1			7
Total of defaults and counterfeits	9	2			11
Full sample					
=1 if default or counterfeit	0.17 (0.05)	0.04 (0.03)	0.12 (0.06)	**	0.11 (0.03)
Card Quality ¹	6.06 (0.40)	7.05 (0.24)	-0.90 (0.49)	**	6.53 (0.24)
Seller claim of card quality ²	9.21 (0.09)	8.39 (0.13)	0.99 (0.16)	***	8.87 (0.09)
Seller claim -True card quality ²	3.17 (0.44)	1.42 (0.32)	1.71 (0.62)	***	
Final winning price we paid (\$)	73.49 (7.16)	47.65 (2.27)	24.50 (8.35)	***	61.34 (4.13)
Card Value ³	34.62 (6.59)	36.65 (3.07)	0.09 (8.26)		35.57 (3.99)
Card Value accounting for eBay reimbursement for defaults ⁴	37.79 (6.52)	37.14 (2.98)	3.465 (8.111)		37.48 (3.71)
Authentic Deliveries Only					
Card Quality	7.30 (0.14)	7.37 (0.11)	0.03 (0.21)		7.33 (0.09)
Seller claim of card quality ²	9.12 (0.10)	8.41 (0.13)	0.88 (0.18)	***	
Seller claim -True card quality ²	1.81 (0.18)	1.04 (0.18)	0.84 (0.32)	**	
Final winning price we paid (\$)	73.18 (8.49)	47.96 (2.35)	25.17 (10.49)	**	60.43 (4.54)
Card Value ³	41.70 (7.51)	38.28 (2.98)	7.62 (9.70)		39.97 (3.99)

Note: Standard errors in parentheses. ***p<0.01, ** p<0.05, * p<0.1, two-tail. The number of best-ranked cards is slightly bigger than the number of median-ranked cards because several best-ranked cards are sold in a package of multiple copies (of same card same claim). Pairwise comparison between best- and median-ranked groups is conducted within card-week. (1) Defaults and counterfeits are coded as zero quality. (2) The online comparison of seller claims is conditional on cards that have any seller claims. It does not change any conclusion if we impute no claim as claiming 8. Retail dealers rarely make any quality claims for ungraded cards. (3) Card value is calculated based on Beckett low book prices. (4) Assuming all the undelivered cards are reimbursed by eBay and the reimbursement equals to the winning price minus \$25 processing fee.

Table 4.2 Detailed distribution of self claim and true quality of the experimental sample

Grade	Best Rank		Median Rank	
	self claim	true quality	self claim	true quality
0		9		2
0.5				
1				
1.5				
2				
2.5				
3				
3.5		1		
4				
4.5				
5				
5.5				
6		5		4
6.5		2		7
7	1	8	6	8
7.5		15		9
8	5	10	22	12
8.5		2		5
9	30	1	17	
9.5			1	
10	17		1	
Total Count	53	53	47	47

Note: Empty cells mean no observations.

Table 5.1: Experimental Results -- Summary of Online Seller Reputation

Sample: Online purchases only

		N	ln (seller net ratings +1)	Significant Difference?
Best ranked?	Yes	53	5.42 (0.23)	
	No	47	4.92 (0.30)	
Have any self grade?	Yes	89	5.05 (0.22)	
	No	11	5.99 (0.62)	
Self grade 9 or 10?	Yes	66	4.72 (0.23)	**
	No	34	6.00 (0.30)	
Self grade 10?	Yes	18	3.63 (0.43)	***
	No	82	5.49 (0.20)	
Default or counterfeit?	Yes	11	3.83 (0.66)	**
	No	89	5.32 (0.19)	

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1, two-tail.

Table 5.2: Experimental Results -- Regression on Online Purchases

Dep. Var.	Default or Counterfeit? Probit (coefficients)		Card Quality OLS (coefficients)				ln (winning prices/book low) OLS (coefficients)	
Sample	All Online Purchases		All Online Purchases		Authentic Deliveries Only		All Online Purchases	
Seller self grade	0.303 (0.308)		-0.542 (0.484)		-0.020 (0.137)		0.276 *** (0.069)	
=1 if seller self grade 9 or 10		-0.080 (0.468)		0.345 (0.727)		0.366 (0.248)		0.355 *** (0.126)
=1 if miss self grade	dropped	dropped	-2.647 (3.893)	2.022 * (1.094)	0.113 (1.215)	0.415 (0.462)	2.400 *** (0.614)	0.239 * (0.123)
ln(seller net rating +1)	-0.224 (0.142)	-0.272 ** (0.134)	0.415 ** (0.199)	0.505 ** (0.212)	-0.020 (0.055)	0.000 (0.061)	0.020 (0.029)	-0.001 (0.031)
=1 if has any negative feedback	0.589 (0.443)	0.553 (0.444)	-0.130 (0.781)	-0.377 (0.711)	0.014 (0.294)	-0.095 (0.336)	0.103 (0.111)	0.143 (0.130)
OBS	100	100	100	100	89	89	100	100
R2	0.1182	0.0965	0.728	0.7152	0.6444	0.66	0.7476	0.6827

Note: Standard errors in parentheses. *** p<0.01, **p<0.05, *p<0.1, two-tail. In addition to a full set of card type dummies, the Probit specification includes cluster of card-week. Quality and price regressions include a full set of card-week dummies. All defaults and counterfeits have seller self grades, so the missing dummy for seller self grade drops out of the Probit equation.