Buying Reputation as a Signal of Quality: Evidence from an Online Marketplace*

Lingfang (Ivy) Li† Steven Tadelis‡ Xiaolan Zhou§

June 20, 2018

Abstract

Seller reputation, generated by buyers leaving feedback, is critical to foster trust in online marketplaces. We argue that signaling theory predicts that only high quality sellers would reward buyers for truthful feedback, and we explore this scope for signaling using the "reward-for-feedback" mechanism on Alibaba Group's Taobao marketplace. We find that items with rewards generate sales that are nearly 30% higher and are sold by higher quality sellers, consistent with a signaling equilibrium. The market design implication is that marketplaces can benefit from allowing sellers to use rewards to build reputations and signal their high quality in the process. *JEL* Classifications: D47, D82, L15, L86.

Keywords: reputation, feedback, ratings, signaling, rebate mechanism, Taobao.

*We thank Alibaba Group for providing us access to the data. We are grateful to to Luis Cabral, Mo Xiao, Dimitriy Masterov, Hanghui Zhang, and Tom Blake for helpful discussions, and to participants at several seminars for valuable feedback. Lingfang Li is grateful for support from the National Natural Science Foundation of China (grants #71402031, #71490721 and #71522007), Shanghai Education Development Foundation and Shanghai Municipal Education Commission (Shuguang Program, No. 17SG05), and Alibaba Huoshui Research Plan. Xiaolan Zhou is grateful for support from the National Natural Science Foundation of China (grant #71603159) and the Fundamental Research Funds for the Central Universities (grant #2017ECNU-HLYT005). This paper was previously presented under the title "Does Money Really Talk? Evidences from Reward-for-Feedback Mechanism in Alibaba's C2C Platform." The authors are listed alphabetically and have contributed equally to the paper.

†School of Management, Fudan University, e-mail: lingfangivy.li@gmail.com

‡UC Berkeley, NBER and CEPR, stadelis@berkeley.edu

§Faculty of Economics and Management, East China Normal University, e-mail: xiaolan.z.zhou@gmail.com

1 Introduction

The growth of trade in online marketplaces such as eBay, Taobao, Etsy, and others in the past two decades is remarkable because buyers are purchasing items they cannot inspect from anonymous and far away sellers. It has been argued time and again that reputation and feedback systems foster the trust needed to make buyers feel comfortable transacting in these large anonymous marketplaces (Dellarocas, 2003). Many studies provide evidence that buyers indeed respond to a seller's reputation in predictable ways.[1]

However, feedback is a public good that may be under-provided (Bolton et al., 2004; Chen et al., 2010; Lafky, 2014). Furthermore, recent studies have shown that user-generated feedback is often biased, with extreme levels of "grade inflation" across several platforms (Nosko and Tadelis, 2015; Horton and Golden, 2015; Zervas et al., 2015) Several market design questions naturally arise: First, can feedback be generated to be more informative? Second, should marketplaces offer rewards for informative feedback or should they encourage sellers to do this? And last, if sellers offer rewards for feedback, how will this affect market outcomes?

In this paper we argue that established theory can shed light on these questions, and turn to a unique data set to test whether the theoretical implications are borne in the data. We build on Nelson (1974) who argued that sellers with high quality goods will pay to advertise their goods as a signal of quality, an idea that was formalized by Kihlstrom and Riordan (1984) and Milgrom and Roberts (1996). Intuitively, high quality sellers will be willing to spend on advertising because they will benefit from repeat purchases by happy buyers, which will not be the case for low quality sellers. We argue that, in a similar way, only high quality sellers who expect to receive positive feedback will pay buyers to leave feedback.[2] For feedback to be truthful so that high quality sellers get positive feedback and low quality sellers get negative feedback, payment for feedback must be unconditional on sentiment and mediated by the marketplace. In equilibrium, the market for feedback separates the wheat from the chaff, allowing good sellers to buy reputation through rewarding buyers for feedback so as to distinguish themselves early on and increase their sales.

[1] For recent surveys of this literature see Cabral (2012) and Tadelis (2016).

[2] In the rest of the paper, "high quality sellers" refer to the sellers who provide high quality transactions. High quality of transaction means the good is received by the buyer, the quality of the good is the same as the seller promised, and the good is shipped on time. A transaction is of low quality if it fails any of these criteria.

Using proprietary data obtained from Alibaba Group's Taobao marketplace,[3] we shed light on three questions that naturally follow from the theoretical framework we follow. First, do high quality sellers choose to signal their quality by rewarding buyers for leaving informative feedback? Second, do buyers seem to infer sellers' quality from their decision to offer rewards for feedback? And if so, what are the returns to sellers who reward buyers for feedback in terms of sales and feedback?

We exploit the fact that on March 1, 2012, Taobao introduced a feedback reward mechanism called "Rebate-for-Feedback" (RFF). Sellers could set a rebate amount for any given items they sold (in the form of cash back or a store coupon) as a reward for a buyer's feedback after purchasing that item. If a seller chooses the RFF feature then Taobao guarantees that the rebate is transferred from the seller's account to a buyer who leaves high-quality feedback. Feedback quality *does not* depend on whether it is favorable but instead depends only on how informative it is, which is determined by a machine learning algorithm that examines the content and length of the buyer's detailed feedback and whether key features of the item are mentioned.

The RFF feature plays a dual role: first, it obviously can induce more feedback, and second, it can offer sellers a way to signal their high quality. Therefore, before a high quality seller accumulates sufficient positive feedback, the RFF feature allows him to send a signal about his quality and intentions to satisfy buyers. In addition, it solves the "cold start" problem where new sellers, even if being of high quality, will face a barrier to entry due to the fact that they have no past feedback. Furthermore, because RFF encourages buyers to leave feedback, high quality entrants will gain in the long run from attracting future buyers without having to reward them for feedback.

Our panel dataset consists of all transactions purchased from 13,018 randomly selected sellers who sold at least one unit between September 2012 and February 2013 on Taobao.com in four distinct categories: cell phones, memory cards, cosmetic masks, and jeans. A key feature of the data is that products are given a unique identifier for each an every seller. This allows us to use a fixed-effects panel model that controls for seller and product attributes simultaneously, and identifies variation within seller-product pairs to establish how sellers behave across time holding their quality

[3]Taobao is world's largest online consumer-to-consumer (C2C) e-commerce platform, owned by the world's largest retailer Alibaba Group (NYSE: BABA), Taobao sold RMB 1.173 trillion (US$190.6 billion) in gross merchandise volume (GMV) from Q1 2013 to Q1 2014, which is more than 2.3 times eBay's 2013 GMV. See http://www.techinasia.com/alibaba- updates-ipo-filing-reveals-taobao-tmall-sales-figures-2014/ and http://www.marketwatch.com/story/ebay-inc- reports-fourth-quarter-and-full-year2013-results-2014-01-22 (accessed on August 21, 2014).

fixed. We use feedback given during the period in which sellers choose to use RFF as a measure of their underlying quality while controlling for past performance using past feedback information.

Our main findings confirm the implications of RFF acting as a potent signaling mechanism. First, the RFF feature is chosen by higher quality sellers, suggesting that sellers self-select and use it as a signal of quality. Second, more experienced and well established sellers are less likely to use the RFF feature suggesting that it is indeed a substitute to established reputation. Third, buyers respond strongly to the signal content of using a RFF – sales of an item are about 30% higher when the seller chooses the rebate option, implying that buyers understand the signaling intent of the RFF feature and act on it. This effect is large in relative terms, but due to the low number of sales per seller, it results in very reasonable effects. The median item sells about 3 items per month, implying that when a seller uses RFF then he sells one more item. Last, RFF induces buyers to write more detailed feedback but does not bias buyers towards positive feedback.

Our paper offers three contributions to the literature. First, we believe we are the first to empirically analyze a novel feedback enhancing mechanism in a large online marketplace, and show that it provides both signaling and public good provision benefits. Second, we provide compelling empirical evidence on the role of signaling in online markets, showing that sellers send credible signals to which buyers respond rationally, allowing them to select better sellers or with higher quality goods.[4] This also shows the broader signaling implications of Nelson (1974), which was directly tested in a recent paper by Sahni and Nair (2016). Last but not least, we contribute to the growing market design literature with respect to managing asymmetric information in online markets. Unlike Hui et al. (2016), Nosko and Tadelis (2015) and Masterov et al. (2015), who emphasize how marketplaces can manage the asymmetric information problem by regulating seller quality, we show that marketplaces can leverage the signaling incentives of high quality sellers and allow them to self-select using RFF mechanisms to improve market outcomes.

Several papers have focused on the public goods nature of feedback and proposed two ways to incentivize buyers to leave more feedback: either online marketplaces can provide incentives to buyers to leave feedback (Miller et al., 2005; Fradkin et al., 2015), or sellers can provide these incentives (Li, 2010). Using a mixed model of adverse selection and moral-hazard, Li (2010) shows that sellers will choose to rebate buyers for feedback if sellers decide to exert effort and provide a

[4] Two recent studies that use online marketplaces to provide evidence of signing equilibria are Backus et al. (2018) and Kawai et al. (2013).

high-quality transaction. As a consequence, buyers will avoid sellers who do not choose the rebate option. Li and Xiao (2014) extend this idea to listed-price online markets and test the rebate mechanism in lab experiments, finding evidence consistent with the theoretical predictions in Li (2010). Cabral and Li (2015) run a series of controlled field experiments on eBay where sellers propose monetary rewards for providing (any) feedback, and find that buyers grant these sellers more frequent feedback and and more favorable feedback when transaction quality is high, but when transaction quality is low, offering a rebate significantly decreases the likelihood of negative feedback. The mixed results of Li and Xiao (2014) and Cabral and Li (2015) most likely emerge because the implementation of the mechanism is different across the two studies. In Li and Xiao (2014), once a seller choose RFF the experimenter guarantees the reward will will be transferred, but in Cabral and Li (2015), the sellers promise a reward, but buyers may not believe negative feedback will be rewarded. Importantly, when Taobao implemented RFF, the platform guaranteed the reward is paid by the seller to the buyer not based on whether it is positive or negative, but on the informative content that is monitored using a machine-learned Natural Language Processing algorithm. Our findings are consistent with those in Li and Xiao (2014), showing the importance of ensuring that the sellers' commitment to the signal is creditable and implying the platform must play an active mediating role in designing the market for feedback effectively.

Our paper also contributes to a growing literature that empirically studies the workings of reputation systems in online markets with asymmetric information. A series of recent studies have shown that, because feedback is user generated, reputation scores are often biased (Dellarocas and Wood, 2008; Nosko and Tadelis, 2015), inflated (Horton and Golden, 2015; Zervas et al., 2015), and possibly manipulated by market players (Mayzlin et al., 2014; Luca and Zervas, 2016; Xu et al., 2015).[5] We argue that RFF mechanisms can help promote honest and informative feedback while offering the added benefit of a signaling mechanism that promotes high quality sellers. This reduces the inefficiencies of adverse selection in markets with asymmetric information.

The paper proceeds as the follows. In Section 2 we describe Taobao's RFF mechanism. In Section 3 we lay out our hypotheses. In Section 4 we describe the data, while Section 5 presents our analyses and results. Section 6 concludes and discuss some limitations of our analysis.

[5] Some sellers start selling cheap items to climb up the reputation ladder, as noted in Brown and Morgan (2006) and Fan et al. (2016). Proserpio and Zervas (2017) show that when hotel management respond to online reviews, they then receive fewer but longer negative reviews. Online reviews also help consumers learn, which is explored in the context of restaurant reviews by Wu et al. (2015).

2 The Taobao Reward-for-Feedback Mechanism

Launched in 2003, Taobao Marketplace (www.taobao.com) has become the most popular C2C online marketplace in China with close to half a billion registered users. On an average day, more than 60 million visitors have access to more than 800 million product listings, and an average of 50,000 items of merchandise are traded every minute.[6] Like eBay, Taobao intermediates between buyers and sellers. But unlike eBay, Taobao earns revenues from advertising and other services, not from charging listing or commission fees from sellers. The majority of the products are new merchandise sold at fixed prices. On Taobao, customers can pay using Alipay which is comparable to eBay's users using PayPal. But unlike Paypal, Alipay takes money up front and puts it in an escrow account and verifies that the customer is satisfied with the product before payment is released.

An "item" on Taobao refers to a product-seller combination and any product sold by a seller is assigned a unique item ID. If another seller sells the same product, it will be assigned a different item ID. This is unlike the definition of an item in other marketplaces such as eBay, in which an item refers to a product and not a product-seller pair.

Taobao's feedback system is similar to eBay's. After a transaction is completed, buyers and sellers can leave each other positive, neutral, or negative feedback, as well as detailed comments about the transaction. However, Taobao's and eBay's feedback systems differ in three ways. First, Taobao separately reports a user's rating score as a seller and as a buyer, whereas on eBay a user's total rating score is aggregated for sales and purchases.[7] Second, Taobao reports buyers' feedback of an item on both the item's rating page and also on the seller's rating page. An item's rating page refers to the "detailed feedback comment" page as shown in figure 1. Buyers can see all ratings for the item when browsing an item's page. A seller's rating page displays ratings for all items sold by a seller on the seller's profile page. By contrast, eBay only has seller's rating profile page like Taobao, but does not provide item's rating page. Third, if a buyer does not leave any feedback 15 days after a seller leaves feedback for the buyer, the Taobao system will leave an automatic positive feedback for the seller. In the "comments" area for the transaction it displays "Feedback provider didn't leave feedback on time, the system offered automatic positive rating!" which is a message containing

[6]Source: BBC India. See https://www.facebook.com/bbcindia/posts/741334802577552 (accessed on February 16, 2017).

[7]We use "rating" to represent both positive/neutral/negative feedback and detailed comments for the rest of this paper. A rating for an item reflects a buyer's opinion about the merchandise and the seller.

18 Chinese characters.[8] By contrast, on eBay, if no feedback is left after a transaction, future buyers will not even know the transaction has happened. On one hand, this automatic feedback feature on Taobao makes it possible to track all sales of an item, but on the other hand, the automatic positive ratings may bias the ratio of positive ratings to be higher than it should, thus affecting the informativeness of the feedback system.[9]

On March 1, 2012, Taobao launched a "Rebate-for-Feedback" (RFF) feature for sellers.[10] A seller has an option to set a rebate value for any item he sells (in the form of cash-back or a store coupon) as a reward for a buyer's feedback. A seller can choose which items to opt into RFF and the form and amount of rebates, and can specify the period of RFF. If a seller chooses this option then Taobao guarantees that the rebate will be transferred from the seller's account to a buyer who leaves high-quality feedback. Feedback quality only depends on how informative it is, rather than whether the feedback is positive or negative. Taobao measures the quality of feedback with a Natural Language Processing (machine learning) algorithm that examines the comment's content and length and finds out whether key features of the item are mentioned. The rewarded rating will be identified on the item's rating page as one for which RFF was granted, so that future buyers can know this rating was rewarded.

In the Appendix we provide a translation of Taobao's announcement of the new online service. One of the announced goals is to "increase the ratio of non-automatic ratings for sellers." Another is "increasing the quality of buyers' comments," using Taobao's machine learning algorithm to judge feedback quality. Finally, with respect to Taobao's role in helping the seller offer a rebate, the seller deposits a certain amount for a chosen period and Taobao freezes the deposit until the end of the rebate period so that funds are guaranteed for buyers who meet the rebate criterion.

According to a Taobao survey (published in March 2012), 64.8% of buyers believe that they will be more willing to buy items that have the RFF feature, and 84.2% of buyers believe that the RFF option will make them more likely to write detailed comments.[11] Figure 1 shows a Taobao.com page

[8]Sellers always leave feedback for a buyer in order to get an automatic positive feedback in case the buyer leaves none. Fan et al. (2016) also provide an introduction to Taobao's feedback system.

[9]Dellarocas and Wood (2008) suggest that feedback not left on eBay is more likely to be negative because of fear of retaliation. On Taobao, the retaliation concern is even larger because sellers know buyers' cell phone and address, so buyers may get unpleasant calls from some sellers if they leave negative feedback.

[10]The RFF feature implemented by Taobao is similar to the mechanism proposed in Li (2010) and Li and Xiao (2014). In fact, Li suggested the RFF mechanism to Alibaba Research towards the end of 2011, and several months later, Taobao launched the RFF mechanism. See http://www.aliresearch.com/blog/article/detail/id/20486.html (accessed on June 15, 2015).

[11]http://bbs.taobao.com/catalog/thread/513886-256229600.htm, accessed June 24, 2012.

with the RFF feature. The box just below the 4.9 score includes a feedback reminder, that reads "Dear customer, you will have a chance to get 0.50 RMB reward if you leave feedback conscientiously on the product from April 19–27, 2012." The box on the lower right corner, in turn, includes a notice "this comment is informative, so it is rewarded with 0.50 RMB."

3 Theory and Hypotheses

Rather than lay out a formal model of RFF signaling, we adapt the seminal theory of advertising (Nelson, 1974; Kihlstrom and Riordan, 1984; Milgrom and Roberts, 1996), which translates naturally to our setting. As a precursor to this literature, Nelson (1970) introduced the concepts of *search goods* and *experience goods*: an experience good is a product or service where product characteristics are difficult to observe in advance, but these characteristics can be ascertained upon consumption; a search good is a product or service with features and characteristics easily evaluated before purchase. Hence, adverse selection problems are severe with experience goods, and some mechanism is needed to help buyers infer which products are high quality and which are not.

The advertising-as-a-signal literature established that only high quality sellers will spend money on ads to promote their experience goods because only high quality sellers can be confident that they will receive positive returns from their expenditures. According to the theory, advertising—which is a form of "burning money"—acts as a signal that attracts buyers who correctly believe that only high quality sellers will choose to advertise. Incentive compatibility is achieved through repeat purchases: buyers who are attracted to sellers who advertise will buy the good, experience it, and will return in the future only if the goods sold are of high enough quality. Advertising has to be costly enough to deter low quality sellers from being willing to spend the money and sell only once, because they will not attract repeat customers. Hence, ads act as signals that separate high quality sellers, and in turn attract buyers to their products.

It is easy to see that a RFF mechanism can play a similar signaling role as ads do. Assuming that consumers express their experiences truthfully in written feedback, any consumer who buys a product and is given incentives to leave feedback, will leave positive feedback only if the buying experience was satisfactory. Hence, a seller will offer RFF incentives to buyers only if the seller expects to receive positive feedback, and this will happen only if the seller will provide high quality transactions. If a seller knows that their goods and services are unsatisfactory, then paying for feedback will generate negative feedback that will harm the low quality seller.[12] Equilibrium behavior then implies that RFF, as a signal of high quality, will attract more buyers and result in more sales. This generates our main signaling hypotheses,

S1: *When a seller chooses RFF, he is more likely to provide high quality transactions. (Seller Signaling Hypothesis)*

B1: *Choosing RFF signals high quality and increases an item's sales. (Buyer Belief Hypothesis)*

Note that if in equilibrium buyers infer that an item with the RFF feature has higher quality (S1), then its sales will increase (B1), but for some buyers the cost of leaving feedback may be higher than the compensation from the RFF itself. For example, imagine that buyers have some idiosyncratic costs of leaving feedback that is distributed over an interval $[0, \bar{c}]$, and assume that the RFF offered by high quality sellers is set at r, where $0 < r < \bar{c}$. Then, even though all buyers should infer that the RFF is a signal of high quality, only those with a cost of leaving feedback that is below r will leave feedback, while those with costs above r will not. Therefore, some buyers will be attracted by the signal of quality and *not* leave any ratings. As described in Section 2, transactions with no buyer feedback will automatically receive a positive rating via Taobao's system. Buyers who leave non-effective ratings (i.e., automatic ratings or zero-word ratings as described in Appendix-2) should not expect to be rewarded. Hence, the signaling value of the RFF has another testable implication,

B2: *Choosing RFF increases an item's number of non-effective ratings through sales with no detailed feedback. (Increased Ineffective Rating Hypothesis)*

Interestingly, because of the ongoing dynamic nature of reputation in online marketplaces, this will create a virtuous cycle: the increased number of sales and detailed positive feedback will attract more future buyers in the long run. This ties in with the theoretical literature on seller reputation, which shows that building a reputation is more valuable in earlier stages of a seller's career (Bar-Isaac and Tadelis, 2008). Similarly, the advertising literature suggests that a firm will

[12] A moral hazard model as in Li (2010) and Li and Xiao (2014) can also demonstrate a role for RFF as a commitment to exert effort and provide high quality. When a seller chooses RFF, it operates as a bond. If he exerts effort he will get a positive review that encourages future buyers, while if he does not exert effort then a negative review will be left, which in turn negatively impacts future sales.

"burn money" to promote brand awareness in its early stages (Milgrom and Roberts, 1996; Bagwell, 2007), after which its reputation will be established. Therefore, a seller has a higher incentive to choose RFF as a signal to attract buyers during earlier stages of his selling career, which will also be an investment in reputation. This implies,

S2: *A seller is more likely to choose RFF before establishing a good reputation on Taobao. (Reputation Building Hypothesis)*

As a last implication, the exact nature of Taobao's RFF mechanism has further implications that we can test in our data. The automatic algorithm rewards buyers for what is deemed to be informative feedback, which is highly correlated with length. It is therefore obvious that any buyer who is leaving feedback to get a rebate must leave longer feedback. Since RFF is designed to offer rebates based on the informativeness of feedback regardless whether the feedback is positive or negative, the feedback induced by RFF should be unbiased as suggested in the models of Li (2010) and Li and Xiao (2014). Hence,

B3: *Choosing RFF increases the average length of an item's ratings but will not create a positive bias. (Long Unbiased Ratings Hypothesis).*

4 Data Description

Our data consist of all transactions sold by 13,018 randomly selected sellers who sold at least one unit in the four chosen categories between September 2012 and February 2013. All products are new merchandise offered at fixed prices. The categories we chose represent both search goods and experience goods in various price ranges. Cell phones and memory cards (or "TF cards") are search goods because quality is generally known prior to purchase (as long as the products are authentic). Only the service quality, such as the shipping speed and return policy, may vary across sellers. Cosmetic masks and jeans are experience goods because their true quality can only be determined by using them.[13] In the data set, there are $114,090$ items in the four categories with positive sales

[13]Nelson (1970) used clothes as an example of search goods, because a buyer can try it before buying it in a store, and a TV is an experience good as a buyer doesn't know how well it works. However, with e-commerce there is less information asymmetry about TVs than clothes. Most cell phones and TF cards are well-known branded products for which buyers can read online reviews, whereas most jeans or facial masks sold on Taobao cannot be found or tried in offline stores, so a buyer cannot try them before paying for them.

in the sample period. Among the $13,018$ sellers, 60.82% used a rebate at least once during the six months for some product (not necessarily in the four chosen categories).

Because an *item* on Taobao refers to a product-seller combination, the same product sold by different sellers will be assigned different item IDs. A *transaction* is defined as a sale of an item (one or multiple units) from a seller to a buyer. For each transaction, our data contain transaction ID, item ID, category ID that the item belongs to, buyer ID, seller ID, quantity sold, total transaction price (including shipping fees), time stamp (e.g., 2013-01-08 23:15:49), and the corresponding rating information if it has been rated. The *rating information* includes whether it is positive, neutral, or negative, its length (in Chinese characters), time stamp, and whether it is a first- or second-time rating.[14] Table 1 provides summary statistics of item attributes at the item-month level.

We define an *"effective rating"* is one that we determine to be deliberately left by a buyer. These exclude automatic ratings (coded as 18 Chinese characters positive ratings in Taobao database) as well as zero-word ratings. By excluding ratings with 18 characters we are discarding some ratings that are from genuine buyers but these are a tiny fraction of ratings. From Figure 2, we see that the ratings with 17 and 19 characters are less than one percent, so the authentic consumer comments that contain 18 characters are likely to be less than one percent as well. We define a *"long rating"* to be a rating with at least 24 Chinese characters, which is the 75^{th} percentile of all the ratings excluding the 18-character ratings that predominantly are automated feedback.

For each seller our data set contains information about his location (province), his service promises (e.g., whether he accepts any returns within seven days, etc.) and daily reputation data, including his seller rating score, his corresponding seller rating grade, and the ratio of positive ratings. As described in the Appendix, there are a total of 21 seller grades, where a higher grade is identified with a higher interval of rating scores, where scores are the number of positive minus negative ratings. This is very similar to the "reputation score" on eBay.

We define a *period* as a month because information about an item on Taobao is reported in terms of the last 30 days. For example, when a potential buyer searches for a product and sorts the search outcomes by items' sales, the results are ranked based on items' sales in the past 30 days. A seller's ratio of positive ratings is reported for the last month and last six months on Taobao. For each item, we summarize its product characteristics as monthly sales, monthly average price, and

[14]For each transaction, a buyer is allowed to leave feedback more than once because for some products quality cannot fully be determined until it is used for a while.

monthly number of positive, neutral, negative, and long ratings. For each seller, we use the seller rating grade at the beginning of each month, and the ratio of seller positive ratings at the monthly level as seller characteristics. As a robustness test we used other time frames as the period duration, and the main findings still hold.

There are two types of *rebates* a seller can choose: cash or coupon.[15] For all transactions with the RFF feature, we know whether the buyer received a cash or coupon rebate. About 26.33% of rebates are in the form of cash and the average value of cash rebates is RMB 1.41.[16] Since the conditions of using a coupon are not available in the dataset obtained from Taobao, we use only the value of cash rebates. Also, since the starting time of a rebate is recorded in the data but not its ending time, we define a rebate-month dummy that takes the value of 1 if either a seller initiated at least one rebate for a given item in a given month, or some buyer gets a rebate from a transaction from the given item in the given month. Since some sellers changed an item's rebate form (cash or coupon) within a month, we define the rebate form as a cash rebate if a cash rebate has been chosen at least 50% of the time in the month, otherwise it is defined as a coupon rebate.

5 Empirical analysis

Before turning to the regression analyses, it is illustrative to look at buyer behavior with respect to items for which RFF were sometimes available and sometimes not during our sample period. First, with rebates there are more ratings of any character length. Second, Figure 2 shows the distribution of the length of ratings in Chinese characters with and without a rebate. Without rebates, there are relatively more shorter ratings (under 14 characters) while with rebates there are relatively more longer ratings. This suggests that rebates play a role in motivating people on the extensive margin to write more ratings and on the intensive margin to write longer comments.

Figure 3 shows the relationship between rebates and sales, as well as the relationship between rebates and ratings. The average monthly sales of an item with a rebate is much higher than without a rebate. The same applies to the average number of monthly effective and non-effective ratings,

[15]The seller can also choose whether only the first rater will receive the rebate or whether any rater who writes something informative gets a rebate. In our dataset, only 1.4% of rebates are chosen to reward only the first rater. Therefore, for simplicity, we classify rebates into two types: coupon and cash.

[16]In the dataset, some values of cash rebate are missing. The average value of cash rebate for cell phones is RMB 2.81, for TF cards is RMB 2.01, for masks is RMB 0.85, and for jeans is RMB 1.24.

and long ratings. Interestingly, positive ratings of an item with a rebate are almost the same as without a rebate, suggesting that rebates do not create a bias towards positive ratings.

The raw data suggest that the sale of items, the number of non-effective ratings, the ratio of long ratings, and ratio of effective ratings are positively correlated with offering a rebate, all suggestive of our hypotheses. We continue with a more nuanced empirical analysis to tease out each of the hypotheses more carefully.

5.1 Rebate Adoption

In this subsection, we empirically examine the hypotheses related to seller strategic behavior, namely that higher quality sellers use RFF as a signal (S1) and that they use it more when they have less feedback (S2). To identify the factors that may affect a seller's adoption decision, as well as whether the RFF acts as a signal of quality, we estimate the following panel regression:

$$y_{i,s,t} = \alpha + \pi \cdot Product_{i,s,t-1} + \gamma \cdot Seller\_Grade_{s,t-1} + \beta Seller\_Pos\_Rating_{s,t} + \delta_i + \mu_t + \varepsilon_{i,s,t}, \quad (1)$$

where $y_{i,s,t} \in \{0, 1\}$ is an indicator equal to 1 if item i of seller s in period t offers a rebate; $Product_{i,s,t-1}$ is a vector of product characteristics of item i in period $t-1$, including the logarithm of the following variables: sales, number of negative and neutral ratings, number of positive long ratings, and number of negative and neutral long ratings.

To measure a seller's experience selling on Taobao before selecting whether or not to use RFF in periot t, we use $Seller\_Grade_{s,t-1}$, which is one of 21 seller rating grades in period $t-1$. Because there are relatively few negative ratings, $Seller\_Grade_{s,t-1}$ is highly correlated with the number of transactions a seller has made.

To measure a seller's quality during the period t in which he chooses RFF we $Seller\_Pos\_Rating_{s,t}$, which is a seller's positive ratings in period t, which is the number of positive ratings divided by the number of all ratings in month t (alternatively we also use the ratio of positive effective ratings, which is the number of positive effective ratings divided by the number of all ratings in month t.) As mentioned in section 2, a seller's rating page displays ratings for *all* items (including items with RFF and without RFF) sold by him, so the ratio of positive (effective) ratings in month t reflects the seller's overall transaction quality during that period. Unlike the analysis of the impact of a rebate in the next subsection, we use the ratio of seller positive (effective) ratings in period t instead of period $t-1$ to test whether a seller uses a rebate as a signal. Since a seller's quality, especially service quality, may change over time, only the ratio of seller positive ratings in period t, instead of period $t-1$, reflects a seller's quality in period t.[17] δ_i and μ_t are item and month fixed effects to control for unobserved attributes of items and months.

Table 2 shows the results of the regression in equation (1). Our preferred specification is column (1), which uses item fixed effects to identify the decision to adopt RFF *within* product-seller. Column 3 does not include item fixed effects and identifies the effects across items, and hence, across sellers. As all specifications show, a seller is more likely to choose RFF for an item during period t when he provides high quality transactions, which we proxy by his ratio of seller positive ratings in month t. To interpret the results, let $R = \frac{\Pr\{rebate\}}{\Pr\{no\ rebate\}}$ denote the ratio of the probability of choosing a rebate to the probability of not choosing a rebate. The estimated coefficient for "Ratio of seller positive ratings, t" in column (1) is 3.7086. In the data, the 25th, 50th, 75th percentiles of ratio of seller positive ratings are 0.988, 0.995, and 1 respectively. Hence, if ratio of seller positive ratings increases from the 25th to the 50th, then R changes by a factor of $\exp\{3.7086 * (0.995 - 0.988)\} = 1.026$, while if ratio of seller positive ratings increases from the 50th to the 75th percentile, then R changes by a factor of $\exp\{3.7086 * (1 - 0.995)\} = 1.046$. We also run the same model with ratios of seller positive effective ratings as shown in column (2) (column (4) without item fixed effects), and find estimated coefficients that are significantly positive as well. These results show that a seller generally provides higher quality transactions for all items in a period during which he chooses RFF for any item he sells than when he does not choose RFF for an item.[18] These findings support the *Seller Signaling Hypothesis* (**S1**) when we compare the same seller's behavior across his choice to use RFF or not.

Turning to when a seller uses rebates, a seller is more likely to offer RFF when his rating grade is low, confirming the *Reputation Building Hypothesis* (**S2**). The coefficient for "seller rating grade, $t-1$" in column (1) is -0.1767, so if the seller's rating grade last month is one level higher, then R changes by a factor of $\exp\{-0.1767\} = 0.84$. Hence, the likelihood of choosing a rebate declines if the seller gets a higher rating grade (identified by within-seller changes in ratings.) Note also that a

[17] According to Table 1, the average number of days before leaving ratings is 10. Hence, the feedback for transactions occurring towards the end of the month may be displayed in a seller's rating page in month $t+1$. We also use period $t+1$ ratio as a robustness test and find similar results. We do not include the item's (as opposed to seller's) ratio of positive ratings in product characteristics for two reasons. First, it is highly correlated with other product characteristics, such as sales, and number of negative and neutral ratings. Second, it severely fluctuate if there are few ratings. One single negative rating can significantly swing the item's ratio of positive ratings.

[18] Usually, a high ratio of seller positive ratings are associated with high ratio of item positive ratings, as seller positive ratings are calculated by the sum of all positive ratings of items sold by him.

seller chooses to use RFF for an item if it has fewer positive long ratings in the previous month. The coefficient on the number of an item's positive long ratings in the previous month is significantly negative. This suggests that when comparing the same item sold by the same seller in different months, if an item has less long reviews in the last month, then it is more likely to be chosen to offer rebates in this month. This is further supports *Reputation Building Hypothesis* (**S2**).[19]

Column (3) identifies effects across items, and hence, across sellers. The estimated coefficient of ratio of seller positive ratings in month t is also significantly positive, implying that sellers who choose RFF for an item are more likely to provide higher quality transactions than sellers who don't choose RFF in the same month and further confirming the *Seller Signaling Hypothesis* (**S1**). Comparing product characteristics to column (1), the coefficient on the number of negative/neutral ratings in the past month becomes significantly negative and the coefficient on the number of positive long ratings in the last month for an item becomes significantly positive in column (3). These results suggest that an item with better reviews is more likely to be chosen to offer rebates than an item with worse reviews, which adds an interesting and intuitive insight related to the *Seller Signaling Hypothesis* (**S1**). Namely, sellers seem to choose their *best* items to further guarantee a higher likelihood of making their customer happy so that she leaves a positive review.

Unlike column (1), column (3) shows that the coefficient on a seller's past rating grade reverses sign and becomes significantly positive. As explained above, the within product-seller regression of column (1) suggests that sellers use the RFF feature to improve their reputation during early stages of their selling careers, while the result in column (3) suggests that sellers with a higher seller rating grade may be more savvy about using this feature when comparing across different sellers and items. Intuitively, sellers with higher grades invest more time on Taobao, are more experienced, and are more familiar with new tools on Taobao than sellers with low seller rating grades.

Turning to item categories in column (3), the excluded category is jeans and all item coefficients are negative and significant. All else equal, the likelihood of RFF adoption from largest to smallest is jeans, cosmetic mask, cell phone and memory cards when comparing across items and thus sellers. This suggests that a seller is more likely to choose the RFF feature for *experience goods* and

[19]With only a few positive long ratings for an item, consumers may be uncertain about the quality of transactions even if the ratio of item positive ratings is 100%. Recall that all default ratings are positive on Taobao, making it easy to get 100% ratio of item positive ratings when there are only a few sales. Therefore, a seller has an incentive to adopt RFF for the items with less positive long ratings in order to induce more detailed ratings and let future buyers know more about the item's quality.

expensive goods, further supporting the signaling hypothesis in the spirit of the original insights in Nelson (1974), as the effects are largest for experience goods and smallest for search goods.

In summary, when comparing within the same seller-item in different months, we find that a high ratio of seller positive ratings for the month is associated with RFF adoption; when comparing across different item and thus sellers, we find that the ratio of seller positive ratings is positively associated with choosing RFF for an item in that month. We also find that a seller is more likely to choose RFF when he is in the earlier stage of career in Taobao.

5.2 The Impact of Rebates on Sales and Ratings

In this subsection, we empirically examine hypotheses B1, B2, and B3. To investigate the response of buyers to items that offer RFF, we estimate the impact of a rebate on an item's sales and on the ratings it receives using the following panel regression model:

$$y_{i,s,t} = \alpha + \beta \cdot Rebate_{i,s,t} + \pi \cdot Product_{i,s,t-1} + \gamma \cdot Seller_{s,t-1} + \delta_i + \mu_t + \varepsilon_{i,s,t}, \tag{2}$$

where $y_{i,s,t}$ is the dependent variables of interest as shown in each regression model of Tables 3 - 7. $Rebate_{i,s,t}$ is a vector that indicates the rebate status of item i sold by seller s in period t, which includes an indicator if item i of seller s offers any rebate in period t, and an indicator if the rebate is in the form of cash. As before, $Product_{i,s,t-1}$ includes product characteristics of item i, $Seller_{s,t-1}$ includes seller characteristics, and δ_i and μ_t are item and month fixed effects. All the variables (except price) take the value in the previous month because only the lagged value is observed by buyers.

We regress the variables of interest on rebate variables including item fixed effects, which control for product and seller characteristics. During the six-month period in our data, a seller may vary his adoption of a rebate for an item. Since the rebate has two types, coupon and cash, controlling for item fixed effects allows us to use the variation in rebate type to examine not only the effect of adopting the RFF feature, but also to measure the potentially differential effect of the type of RFF.

Note that the seller's choice of adopting the RFF feature is our endogenous (strategic) variable of interest. The previous subsection confirmed the hypothesis that it is adopted as a signal of high quality, and here we test whether in equilibrium buyers respond to it as an informative signal of quality. As mentioned above, we run the regressions of buyer behavior with item fixed effects. This means that changes in behavior as a response to the RFF is identified from variation *within* seller and *within* product by the definition of the item ID. This is a more refined variant of the "matched listing" approach first used by Elfenbein et al. (2012) who study how sellers on eBay use charity as a substitute for reputation.[20] Hence, if product or seller characteristics drive some results, then our use of item fixed effects within our panel structure should alleviate any such concerns and absorb any product or seller heterogeneity.

5.2.1 Buyer Responses to Rebates

Table 3 uses $\ln(\text{sales}_t)$ as the dependent variable in equation (2). The estimated coefficient on the rebate dummy is large and significant, showing that a rebate increases the quantity sold of an item by 30.65% on average. This supports the *Buyer Belief Hypothesis* (**B1**), which is consistent with equilibrium behavior of signaling. To put this effect in context, Table 1 shows that the median seller-item sells about 3 items a month, so using RFF will result in about one more item sold. We also estimated the effects of different types of rebates. Column 3 shows that, on average, a coupon rebate increases the quantity sold of an item by 32.82% while a cash rebate increases the quantity sold by 24.59%. We conjecture that this is probably because a coupon usually has a higher value compared with the average value of a cash rebate.[21]

Most of the coefficients on product and seller characteristics are as we expect them to be and they do not vary much across specifications. Lower price, higher number of previous sales, fewer number of negative/neutral ratings, larger number of positive long ratings, and a higher ratio of seller positive ratings in the past all attract more sales.[22] The upshot is that the effect of a rebate on sales is large.

In Table 4 the dependent variable in equation (2) is number of non-effective ratings. By the nature of the RFF feature, buyers whose ratings contain no comment are not likely to obtain the rebate according to Taobao's conditions. Hence, as we explain in Section 3, buyers who understand the signaling value of rebates but are uninterested in obtaining the rebate may still flock to sellers

[20]On Taobao, an item sold by the same seller has a unique item ID. By contrast, on eBay, an item sold by the same seller several times will be recorded as several "listings". Elfenbein et al. (2012) identify a "matched listing" as a situation where a seller posted multiple items with the same title, subtitle, and starting price that differ in other listing attributes such as committing a fraction of the sale to charity. Einav et al. (2014) rely on variation within matched listings to investigate various sale strategies in eBay.

[21]Sellers are probably more generous with coupons because they can be used only against a future sale from the same seller, whereas a cash rebate is paid out regardless of a buyer's future purchases.

[22]It is curious that in all specifications we find that a higher seller rating grade is associated with fewer sales. Our conjecture is that sellers with a higher seller rating grade are too busy and impersonal, which may slightly deter Taobao buyers.

who offer rebates because of their signaling content. Hence, the number of non-effective ratings reflect the number of buyers who are not attracted by the discount effect of rebate while still appreciating its signaling value. Column (1) of Table 4 shows that rebates increase non-effective ratings by 12.32%. Column (2) shows that a coupon rebate generates more non-effective ratings than a cash rebate (13.25% versus 9.71%.) These results confirm the *Increased Ineffective Rating Hypothesis* (**B2**).

We further consider the heterogeneous effect of rebates across product categories. Column (3) in Tables 3 and 4 reports estimates from dummies for each product category. We find that a rebate has the largest effect on cell phones and jeans, and the lowest on TF cards. This is consistent with the signaling story, which suggests that signaling is more important when the buyer is more concerned about asymmetric information, which is the case when the product is either more expensive (e.g., cell phone), or when it is an experience good (e.g., jeans).

5.2.2 Effect of Rebates on an Item's Ratings

In Tables 5 through 7 the dependent variable in equation (2) are a variety of rating measures. We define the *ratio of effective ratings* as the number of effective ratings divided by the number of all ratings for an item. Columns (1) and (2) in Table 5 show that offering a rebate raises an item's ratio of effective ratings by 7.11%, and that coupon rebates have a larger effect than cash rebates (7.65% versus 5.56%). As column (3) shows, the effects are strongest for jeans and cell phones, weaker for masks, and nonexistent for TF cards, similar to previous patterns.

Columns (4) and (5) in Table 5 show that offering a rebate raises an item's *ratio of long ratings* (the number of long ratings divided by the number of all ratings) for an item by 6.81%, and that coupons have a larger effect than cash rebates (7.38% versus 5.16%). This confirms the first part of the obvious *Long Unbiased Ratings Hypothesis* (**B3**). Column (6) shows that the effects are strongest for jeans and cell phones, weaker for masks, and nonexistent for TF cards.

Table 6 reports the effect of rebates on the number of positive long ratings and number of negative and neutral long ratings for an item. Columns (1) and (4) of Table 6 show that choosing a rebate will increase the number of positive long ratings and the number of negative and neutral long ratings by 30.43% and 2.01%, respectively. The other columns confirm past patterns, that a coupon rebate has a larger effect than a cash rebate, and the impact on experience and expensive goods are larger. From the estimated coefficient we can get a flavor of how the long ratings affect an item's sales as shown in Table 3, together with the estimated effects of rebates on long ratings in Table 6, and we find that the reputation effect caused by RFF is relatively small compared with the total effect of rebates for increasing the next month's sales.[23]

Columns 1-6 in Table 7 exploe whether rebates change the likelihood of receiving positive feedback both using item fixed effects (column 1-3) and also when comparing across items (column 4-6). Because columns 1-3 use item fixed effects, the quality of sellers should be the same with and without rebates, and as such we don't expect to see an effect of rebates on the ratio of positive ratings for the same item.[24] As columns 1-3 show, offering a rebate is not associated with more positive feedback. In fact, the ratio of positive feedback slightly declines, but the magnitude, though significant, is very small. One possible explanation is that buyers have higher expectations from RFF items chosen for RFF as implied by the signaling story. Columns 4-6 of Table 7 drop item fixed effects, allowing the quality of sellers to vary with and without rebates, and we still find no bias towards positive feedback. This supports the later part of the *Long Ratings Hypothesis* (**B1**).

Interestingly, as columns 7-9 in Table 7 show, offering a rebate gives the seller earlier ratings, shortening the time to receiving a rating by 7.27%. A coupon shortens the period by 8.01% while a cash rebate shortens it by 5.16%. This benefits a seller in two ways: he will get money transferred from Alipay faster and it reduces the likelihood of a dispute. Again, the effects are strongest for jeans and cell phones.

In summary, using item fixed-effects to compare outcomes for the same item, when using RFF the monthly sales are nearly 30% higher, and the number of non-effective ratings is nearly 12% higher. We also find that both the number of positive long ratings and negative long ratings of an item will increase when the item is chosen for RFF, but the ratio of positive ratings for the item does not seem to increase.

[23] The reputation effect (separate from the signaling effect) of RFF for increasing next month's sales is calculated as 0.0136 ($= 0.0453 \times 0.3043 + (-0.0101) \times 0.0201$, where 0.0453 and -0.0101 are the estimated coefficients for ln(number of positive long feedback, $t-1$) and ln(number of negative long feedback, $t-1$) in column (1) of Table 3, and 0.3043 and 0.0201 are the estimated coefficients for "Dummy if a rebate, t" in column 1 and 4 of Table 6. The total effect of a rebate on sales is estimated as 0.3065 as shown in model 1 in Table 4. Since 0.0138 is much smaller than 0.3065, it suggests that the short run reputation effect is small, yet there are certainly cumulative effects that are larger. Similarly, comparing with the total effect on the number of non-effective ratings, the reputation effect is also small (0.022 vs. 0.1232) where $0.022 = 0.3043 \times 0.0711 + 0.0201 \times 0.0175$.

[24] A concern may be that offering rebates causes buyers to feel a need to reciprocate and offer more positive feedback than the seller deserves. Cabral and Li (2015) find some support for reciprocal behavior when a rebate is offered by the seller in a field experiment on eBay. Unlike Taobao, eBay offers no enforcement to pay the rebate so buyers may believe that they will not get the rebate if they leave negative feedback.

5.3 Robustness Checks

The data set we obtained from Taobao only covers the period after RFF was implemented, so we cannot conduct a natural experiment study of RFF by using before and after RFF implementation data. Fortunately, as described above, an item on Taobao is a product-seller pair and any product sold by a seller is assigned a unique item ID. So all changes in behavior as a response to using RFF are identified off of variation within seller and within product by the definition of the item ID. A limitation of our study is that when we run the analysis for ratio of seller positive ratings (t), some ratings of transactions in month t may be shown in month $t+1$ since the average number of days before leaving ratings is 10. To address this concern, we run a robustness check for rebate adoption by using ratio of seller positive ratings in month $t+1$ and find similar results.

We run several additional robustness checks to address some potential concerns. Tables are omitted for space considerations and are available upon request.

5.3.1 Endogeneity Concerns

There are two sources of endogeneity in a seller's decisions: price and RFF adoption. To address the potential problem of price endogeneity, we use an instrumental variables approach as a robustness check. Commonly adopted instrumental variables for prices include cost variables, observed exogenous product characteristics, and the number of products in the same market.[25] In our study, cost information is absent, and many product characteristics such as size and color of a cellphone are not observed in our data. Instead, we use the number of items in the same market as an instrument. To define the market we first classify items by prices into four quantiles, and then select items with non-zero sales in the same product category, same price quantile, and the same province. We run regressions using price instruments and find that the main results hold. We also run regressions without price as a control variable and find that the main results still hold.

Note that the seller's choice of adopting the RFF feature is the endogenous (strategic) variable of interest in our analysis. Our empirical analysis is geared towards testing whether it is adopted as a signal of high quality, and whether buyers respond to it as an informative signal of quality. As mentioned above, we run the regressions of buyer behavior with item fixed effects. This means that

[25] Berry et al. (1995) suggests using observed exogenous product characteristics such as the horsepower of a car, the sums of the values of the same characteristics of other products offered by that firm, and the sums of the values of the same characteristics of products offered by other firms.

changes in behavior as a response to the RFF are identified off of variation *within* seller and *within* product by the definition of the item ID. Hence, if product characteristics drive some results, then our use of item fixed effects within our panel structure should alleviate any such concerns.[26]

5.3.2 Product category

In sections 5.1 and 5.2, we report the estimated impact of rebates averaged across our four categories. Because detailed comments of other buyers are more important for goods with more risk involved in the purchase, like experience goods and more expensive goods, we estimate the average impact of a rebate for each category as a robustness check. We divide our data into four categories: cell phones, TF cards, masks, and jeans, and run the panel regressions within each category.

Using item fixed effects as before, we find that a seller tends to adopt a rebate when its rating grade is low, except for sellers of jeans, which supports the Reputation Building Hypothesis *S2*. Dropping item fixed effects, we find that sellers with a high ratio of seller positive ratings are more likely to adopt a rebate in each category, consistent with the Seller Signaling Hypothesis *S1*.

Similar to our findings with the full sample of goods, we find for each category that rebates increase item sales, the ratio of item long ratings, and the number of item long ratings, while they shorten the days between transaction and rating. A rebate has almost no effect on an item's ratio of positive ratings for each category.

5.3.3 Alternative Period Windows

Recall that we only observe when the RFF offering started and do not observe when it ended. In the analysis in sections 5.2 and 5.3, we use one month as a period window and use item-months as the observation measurement. One concern is that, if an item's rebate period is less than a month, our results may be biased. To address this concern we use two-week blocks instead of one month for the period window as a robustness check.

For each two-week period t, the explanatory variables for $t-1$ are created using data from the 30 days prior to t because when a buyer considers an item for purchase, she can see the information

[26] The only remaining endogeneity problem with the rebate strategy may be due to some unobserved time-varying characteristics associated with an item. For example, a seller may change an item's picture in different months together with the choice of rebates, which may cause some correlated responses. Because we have controlled for many item observed characteristics such as product and seller characteristics, we don't believe that unobserved time-varying item characteristics cause a serious endogeneity problem. This is basically the same approach taken by Elfenbein et al. (2012).

about the item for the past 30 days, such as sales, seller's ratings, etc. For example, $\ln(\text{sales}, t-1)$ means the log sales of the item in the 30 days prior to the two-week period t. We continue to obtain results very similar to those in Tables 4-6.

Another concern may be that the periods are adjacent, implying that the first days of period t are closer to the last days in period $t-1$ than they are to most days in period t. To address this concern we conduct another set of robustness tests in which a period includes the first 15 days of each month from the 6 months we have. This gap creates periods for which each day in the period is closer to other days in that same period than to days in any other period. The main results are robust to this specification.

6 Conclusion

The burgeoning growth of online marketplaces and the increased access they provide to data offer new and exciting opportunities to empirically test how markets work in practice. We use a unique data set from Taobao's online marketplace to examine the effects of an RFF mechanism. Our empirical evidence supports the notion that RFF help create a missing market that allows sellers to signal their high quality. Namely, higher quality sellers use RFF to send signals, and buyers respond to these signals rationally, which in turn alleviates some of the adverse selection and cold-start problems in anonymous online marketplaces. Our results shed light on the strategic interaction between buyers and sellers in online marketplaces, which in turn offers insights into the design of online markets.

Specifically, we find that a seller is more likely to choose the RFF option in earlier stages of selling, when other signals like accumulated feedback are not yet established. Upon choosing a RFF, a seller is more likely to provide a high quality transaction, suggesting the positive selection signaling content of RFF. We also show that buyers respond to the RFF signal in ways that are consistent with equilibrium behavior: sales of an item are about 30% higher when the seller chooses a RFF. It is as effective, on average, as quadrupling last month's sales, or increasing the number of positive long ratings in the last month by nearly seven times. Furthermore, the use of RFF does not create bias in feedback, and they afford the seller the opportunity to build up a good reputation faster, creating a "flywheel" of sorts. That is, the signaling content of the RFF encourages both more sales and more feedback, the latter rapidly increasing the seller's reputation, which in turn attracts more buyers and generates more sales. This removes a barrier to entry that new sellers without a reputation face, and alleviates the "cold start" problem that new sellers typically suffer from.

Turning to the market design implications of our study, our analysis suggests that marketplaces can help reduce the asymmetric information problem by letting sellers engage in RFF signaling practices. That is, online marketplaces can rely on the strategic sophistication of both sellers and buyers to alleviate some of the asymmetric information problem by leveraging the signaling incentives of high quality sellers. This last point offers insights into the question of whether marketplaces need to be regulated to improve quality. It is in the interest of marketplaces to reduce the asymmetric information problem, and we show that established market mechanisms such as signaling can be used to enhance marketplace quality, relaxing the need for external forms of regulation.

References

Backus, M., Blake, T., and Tadelis, S. (2018). On the empirical content of cheap-talk signaling: An application to bargaining. *Journal of Political Economy*, forthcoming.

Bagwell, K. (2007). The economic analysis of advertising. In *Handbook of Industrial Organization*, volume 3, pages 1701—1844.

Bar-Isaac, H. and Tadelis, S. (2008). Seller reputation. *Foundations and Trends® in Microeconomics*, 4(4):273–351.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Bolton, G. E., Katok, E., and Ockenfels, A. (2004). How effective are electronic reputation mechanisms? an experimental investigation. *Management Science*, 50(11):1587–1602.

Brown, J. and Morgan, J. (2006). Reputation in online auctions: The market for trust. *California Management Review*, 49(1):61–81.

Cabral, L. (2012). Reputation on the internet. In Peitz, M. and Waldfogel, J., editors, *The Oxford Handbook of the Digital Economy*. Oxford University Press.

Cabral, L. and Li, L. I. (2015). A dollar for your thoughts: Feedback-conditional rebates on ebay. *Management Science*, 61(9):2052–2063.

Chen, Y., Harper, M., Konstan, J., and Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review*, 100(4):1358–1398.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management science*, 49(10):1407–1424.

Dellarocas, C. and Wood, C. A. (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science*, 54(3):460–476.

Einav, L., Kuchler, T., Levin, J., and Sundaresan, N. (2014). Assessing sale strategies in online markets using matched listings. *American Economic Journal: Microeconomics*, 7(2):215–247.

Elfenbein, D., Fisman, R., and McManus, B. (2012). Charity as a substitute for reputation: Evidence from an online marketplace. *Review of Economic Studies*, 79(4):1441–1468.

Fan, Y., Ju, J., and Xiao, M. (2016). Reputation premium and reputation management: Evidence from china's largest e-commerce platform. *International Journal of Industrial Organization*, 46(1):63–76.

Fradkin, A., Grewal, E., Holtz, D., and M., P. (2015). Bias and reciprocity in online reviews: Evidence from experiments on airbnb. *16th ACM Conference on Electronic Commerce, June 2015, (EC 2015)*, page 641.

Horton, J. and Golden, J. (2015). Reputation inflation: Evidence from an online labor market. *Working paper, New York University*.

Hui, X., Saeedi, M., Shen, Z., and Sundaresan, N. (2016). Reputation and regulations: evidence from ebay. *Management Science*, 62(12):3604–3616.

Kawai, K., Onishi, K., and Uetake, K. (2013). Signaling in online credit markets. Working Paper.

Kihlstrom, R. E. and Riordan, M. H. (1984). Advertising as a signal. *Journal of Political Economy*, 3(92):427–450.

Lafky, J. (2014). Why do people rate? theory and evidence on online ratings. *Games and Economic Behavior*, 19(2):554–570.

Li, L. I. (2010). Reputation, trust, and rebates: How online auction markets can improve their feedback mechanisms. *Journal of Economics & Management Strategy*, 19(2):303–331.

Li, L. I. and Xiao, E. (2014). Money talks: Rebate mechanisms in reputation system design. *Management Science*, 60(8):2054–2072.

Luca, M. and Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427.

Masterov, D. V., Mayer, U. F., and Tadelis, S. (2015). Canary in the e-commerce coal mine: Detecting and predicting poor experiences using buyer-to-seller messages. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, pages 81–93, New York, NY, USA. ACM.

Mayzlin, D., Dover, Y., and Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104(8):2421–55.

Milgrom, P. and Roberts, J. (1996). Price and advertising signals of product quality. *Journal of Political Economy*, 4(94):796–821.

Miller, N., Resnick, P., and Zeckhauser, R. (2005). Eliciting honest feedback: The peer prediction method. *Management Science*, 51(9):1359–1373.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 2(78):729–754.

Nelson, P. (1974). Advertising as information. *Journal of Political Economy*, 4(82):729–754.

Nosko, C. and Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment. *NBER Working paper*, (No. w20830).

Proserpio, D. and Zervas, G. (2017). Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science*, 36(5):645–665.

Sahni, N. and Nair, H. (2016). Does advertising serve as a signal? evidence from field experiments in mobile search. Stanford University working paper.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8:321–340.

Wu, C., Che, H., Chan, T. Y., and Lu, X. (2015). The economic value of online reviews. *Marketing Science*, 34(5):739–754.

Xu, H., Liu, D., Wang, H., and Stavrou, A. (2015). E-commerce reputation manipulation: The emergence of reputation-escalation-as-a-service. *Proceedings of 24th World Wide Web Conference*, (WWW 2015):1296–1306.

Zervas, G., Proserpio, D., and John, B. (2015). A first look at online reputation on airbnb, where every stay is above average. *Boston University Working paper*.

Figure 1: Taobao.com page with feedback reward scheme

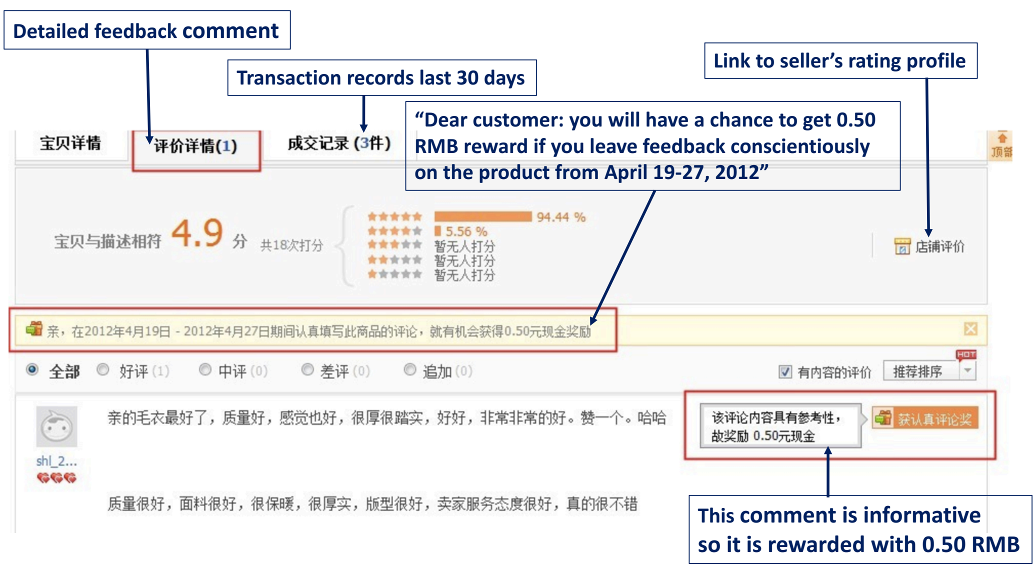

Figure 2: Ratings with and without a Rebate

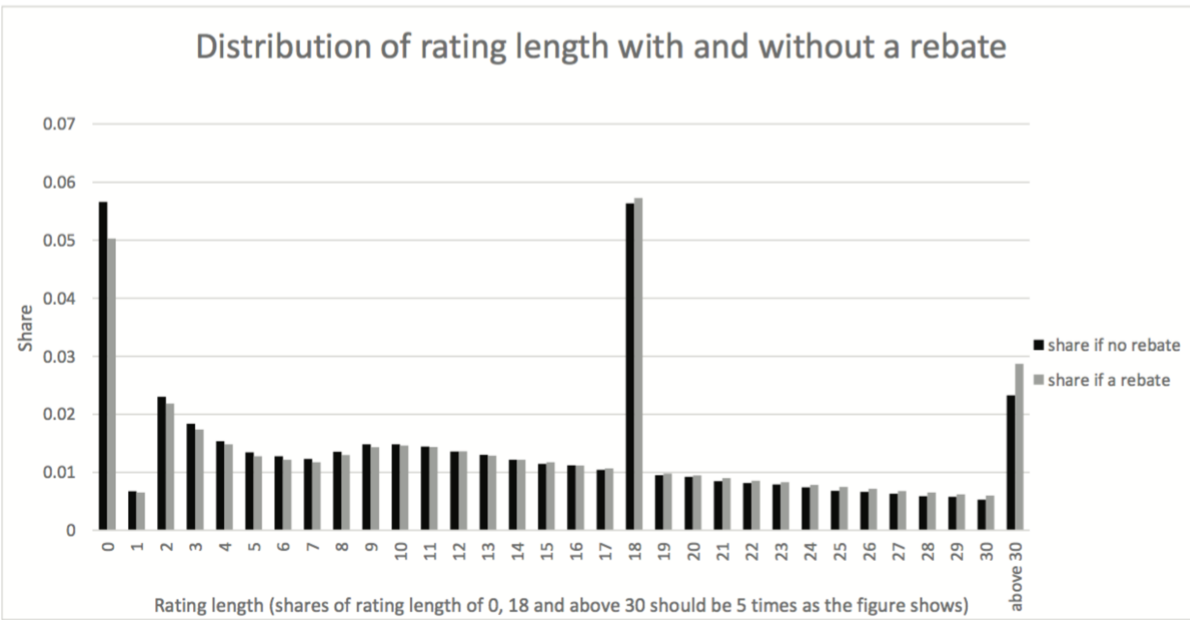

Figure 3: Sales and Ratings with and without a Rebate

Table 1: Summary Statistics (item-month)

	Obs.	Mean	Std.	25%	Median	75%.
Dummy if a rebate	284,263	0.181	0.385	0	0	0
Dummy if a coupon rebate	284,263	0.136	0.343	0	0	0
Dummy if a cash rebate	284,263	0.0454	0.208	0	0	0
Rebate Amount (RMB) conditional on a cash rebate	6,373	1.409	2.235	0.5	1	1
Monthly sales (quantity)	284,263	39.29	692.7	1	3	11
Average transaction price (RMB)	284,263	172.8	892.4	29	68.09	117.6
- Cell phone	30,068	956.4	2216	223.8	505.3	1138
- TF card	7,263	134.2	1842	28	40.77	80
- Mask	111,515	49.10	501.6	5.610	17.50	56.02
- Jeans	135,417	102.8	196.2	60.36	80	114.5
Number of positive ratings	284,263	21.67	533.9	1	2	6
Number of negative and neutral ratings	284,263	0.289	16.54	0	0	0
Number of positive long ratings	284,263	3.799	96.72	0	0	1
Number of neutral and negative long ratings	284,263	0.139	9.678	0	0	0
Ratio of effective ratings	240,308	0.477	0.374	0.111	0.474	0.889
Ratio of positive ratings	240,308	0.990	0.0705	1	1	1
Average number of days before leaving ratings	238,452	10.01	6.332	5	8.625	14
Ratio of long feedback	240,308	0.178	0.287	0	0	0.250
Number of long ratings	284,263	3.937	102.3	0	0	1
Seller rating grade*	273,790	9.878	3.481	8	10	12
Ratio of seller positive ratings*	283,565	0.990	0.0191	0.988	0.995	1
Ratio of seller positive effective ratings*	283,565	0.442	0.166	0.333	0.419	0.529

Notes: The observations are at the item-month level. An item in a month is excluded if the sales are zero. Statistics marked with " * " include sales made by the seller that are not used in our analyses.

Table 2: Adoption of a rebate for an item

Dependent variable: Logit indicator = 1 if adopting a rebate in t	(1)	(2)	(3)	(4)
Seller characteristics				
Seller rating grade, $t-1$	-0.1767***	-0.1334***	0.0557***	0.0771***
	(0.0123)	(0.0124)	(0.0015)	(0.0016)
Ratio of seller positive ratings, t	3.7086***		5.7810***	
	(0.5669)		(0.3363)	
Ratio of seller positive effective ratings, t		2.0246***		1.3953***
		(0.0710)		(0.0335)
Product characteristics				
ln(sales, $t-1$)	0.2853***	0.2905***	0.2328***	0.2467***
	(0.0100)	(0.0100)	(0.0051)	(0.0051)
ln(no. of item ngtv & ntrl ratings, $t-1$)	-0.0660	-0.0454	-0.2385***	-0.2251***
	(0.0560)	(0.0565)	(0.0361)	(0.0360)
ln(no. of item pdtv long ratings, $t-1$)	-0.0927***	-0.0872***	0.4855***	0.4522***
	(0.0176)	(0.0177)	(0.0100)	(0.0100)
ln(number of item ngtv & ntrl long ratings, $t-1$)	0.0918	0.0929	0.0573	0.0496
	(0.0740)	(0.0746)	(0.0482)	(0.0481)
Item category				
Dummy if cellphone			-0.4857***	-0.5340***
			(0.0186)	(0.0186)
Dummy if TF card			-1.2332***	-1.1416***
			(0.0500)	(0.0500)
Dummy if mask			-0.2798***	-0.1744***
			(0.0113)	(0.0111)
Item fixed effect	Yes	Yes	No	No
Month fixed effect	Yes	Yes	Yes	Yes
Observations (item-month)	145,7871	145,781	537,895	537,895

Note: Standard deviations are in parentheses.

Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).

t refers to a variable in month t, and $t-1$ refers to a variable in month $t-1$.

Table 3: Impact of rebate on sales of an item

Dependent variable:	ln(sales (quantity), t)		
	(1)	(2)	(3)
Rebate			
Dummy if a rebate, t	0.3065***		
	(0.0067)		
Dummy if a coupon rebate, t		0.3282***	
		(0.0077)	
Dummy if a cash rebate, t		0.2459***	
		(0.0124)	
Dummy if a rebate for cellphone, t			0.3557***
			(0.0229)
Dummy if a rebate for TF card, t			0.1199*
			(0.0655)
Dummy if a rebate for mask, t			0.1749***
			(0.0106)
Dummy if a rebate for jeans, t			0.3978***
			(0.0091)
Product characteristics			
ln(price, t)	-0.3817***	-0.3820***	-0.3832***
	(0.0097)	(0.0097)	(0.0097)
ln(sales, $t-1$)	0.0665***	0.0664***	0.0659***
	(0.0025)	(0.0025)	(0.0025)
ln(no. of ngtv & ntrl ratings, $t-1$)	-0.0608***	-0.0611***	-0.0618***
	(0.0142)	(0.0142)	(0.0142)
ln(no. of pstv long ratings, $t-1$)	0.0453***	0.0451***	0.0450***
	(0.0045)	(0.0045)	(0.0045)
ln(no. of ngtv & ntrl long ratings, $t-1$)	-0.0101	-0.0102	-0.0087
	(0.0184)	(0.0184)	(0.0183)
Seller characteristics			
Seller rating grade, $t-1$	-0.0311***	-0.0318***	-0.0309***
	(0.0038)	(0.0038)	(0.0038)
Ratio of seller positive ratings, $t-1$	0.9634***	0.9709***	0.9471***
	(0.1647)	(0.1647)	(0.1646)
Item fixed effect	Yes	Yes	Yes
Month fixed effect	Yes	Yes	Yes
Observations (item-month)	225,197	225,197	225,197

Note: Standard deviations are in parentheses.

Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).

We add 1 to all values where logs are taken to avoid $ln(0)$.

Table 4: **Impact of rebate on non-effective ratings of an item**

Dependent variable: ln(number of non-effective ratings, t)	(1)	(2)	(3)
Rebate			
Dummy if a rebate, t	0.1232***		
	(0.0052)		
Dummy if a coupon rebate, t		0.1325***	
		(0.0060)	
Dummy if a cash rebate, t		0.0971***	
		(0.0096)	
Dummy if a rebate for cellphone, t			0.1287***
			(0.0177)
Dummy if a rebate for TF card, t			0.0466
			(0.0508)
Dummy if a rebate for mask, t			0.0694***
			(0.0082)
Dummy if a rebate for jeans, t			0.1627***
			(0.0070)
Product characteristics			
ln(price, t)	-0.1239***	-0.1240***	-0.1244***
	(0.0075)	(0.0075)	(0.0075)
ln(sales, $t-1$)	0.3684***	0.3684***	0.3682***
	(0.0020)	(0.0020)	(0.0020)
ln(no. of ngtv & ntrl ratings, $t-1$)	0.0274**	0.0273**	0.0270**
	(0.0110)	(0.0110)	(0.0110)
ln(no. of pstv long ratings, $t-1$)	0.0711***	0.0709***	0.0709***
	(0.0035)	(0.0035)	(0.0035)
ln(no. of ngtv & ntrl long ratings, $t-1$)	0.0175	0.0175	0.0181
	(0.0142)	(0.0142)	(0.0142)
Seller characteristics			
Seller rating grade, $t-1$	-0.0101***	-0.0104***	-0.0099***
	(0.0030)	(0.0030)	(0.0030)
Ratio of seller positive ratings, $t-1$	0.4476***	0.4508***	0.4425***
	(0.1277)	(0.1277)	(0.1276)
Item fixed effect	Yes	Yes	Yes
Month fixed effect	Yes	Yes	Yes
Observations (item-month)	225,197	225,197	225,197

Note: Standard deviations are in parentheses.

Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).

Table 5: Impact of rebate on informativeness of item ratings

Dependent variable	Ratio of effective ratings, t			Ratio of long ratings, t		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebate						
Dummy if a rebate, t	0.0711***			0.0681***		
	(0.0030)			(0.0023)		
Dummy if a coupon rebate, t		0.0765***			0.0738***	
		(0.0035)			(0.0026)	
Dummy if a cash rebate, t		0.0556***			0.0516***	
		(0.0056)			(0.0042)	
Dummy if a rebate for cell phone, t			0.0657***			0.0694***
			(0.0104)			(0.0078)
Dummy if a rebate for TF card, t			0.0481			0.0345
			(0.0304)			(0.0228)
Dummy if a rebate for mask, t			0.0515***			0.0561***
			(0.0047)			(0.0036)
Dummy if a rebate for jeans, t			0.0869***			0.0773***
			(0.0041)			(0.0031)
Item fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations (item-month)	189,774	189,774	189,774	189,774	189,774	189,774

Note: Standard deviations are in parentheses.
Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).
We add 1 to all values where logs are taken to avoid $ln(0)$.
All models include product and seller characteristics (same as models in Table 3) as control variables.

Table 6: Impact of rebate on number of long item ratings

Dependent variable	ln(no. of positive long ratings, t)			ln(no. of ngtv & ntrl long ratings, t)		
	(1)	(2)	(3)	(4)	(5)	(6)
Rebate						
Dummy if a rebate, t	0.3043***			0.0201***		
	(0.0050)			(0.0016)		
Dummy if a coupon rebate, t		0.3274***			0.0217***	
		(0.0057)			(0.0019)	
Dummy if a cash rebate, t		0.2398***			0.0157***	
		(0.0092)			(0.0030)	
Dummy if a rebate for cell phone, t			0.3440***			0.0411***
			(0.0169)			(0.0056)
Dummy if a rebate for TF card, t			0.1408**			0.0041
			(0.0484)			(0.0160)
Dummy if a rebate for mask, t			0.1996***			0.0014
			(0.0078)			(0.0026)
Dummy if a rebate for jeans, t			0.3771***			0.0308***
			(0.0067)			(0.0022)
Item fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations (item-month)	225,197	225,197	225,197	225,197	225,197	225,197

Note: Standard deviations are in parentheses.
Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).
We add 1 to all values where logs are taken to avoid $ln(0)$.
All models include product and seller characteristics (same as models in Table 3) as control variables.

Table 7: Impact of rebate on bias and time of item ratings

Dependent variable	Ratio of positive ratings, t						ln(average number of days between transaction and rating, t)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rebate									
Dummy if a rebate, t	-0.0016**			-0.0002			-0.0727***		
	(0.0006)			(0.0004)			(0.0045)		
Dummy if a coupon rebate, t		-0.0013**			0.0001			-0.0801***	
		(0.0007)			(0.0005)			(0.0052)	
Dummy if a cash rebate, t		-0.0023**			-0.0010			-0.0516***	
		(0.0011)			(0.0008)			(0.0084)	
Dummy if a rebate for cell phone, t			-0.0033*			-0.0010			-0.0617***
			(0.0020)			(0.0013)			(0.0155)
Dummy if a rebate for TF card, t			-0.0004			0.0038			-0.0756*
			(0.0058)			(0.0044)			(0.0455)
Dummy if a rebate for mask, t			-0.0008			0.0025***			-0.0461***
			(0.0009)			(0.0007)			(0.0071)
Dummy if a rebate for jeans, t			-0.0019**			-0.0018***			-0.0941***
			(0.0008)			(0.0005)			(0.0062)
Item fixed effect	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Month fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations (item-month)	189,774	189,774	189,774	189,774	189,774	189,774	189,764	189,764	189,764

Note: Standard deviations are in parentheses.
Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).
We add 1 to all values where logs are taken to avoid $ln(0)$.
All models include product and seller characteristics (same as models in table 3) as control variables.

Figure Appendix-1: Taobao's announcement of a new feedback reward system (translation)

The purpose of the "rebate for feedback" scheme is to:

- Increase the ratio of non-automatic to automatic seller ratings.
- Increase the quality of buyers' comments.
- Increase feedback for new products and thus reduce buyers' hesitation to purchase.

Benefits for buyers:

- Receive cash or a coupon as a reward for feedback.
- Become opinion leader as the display of their feedback is prioritized over others' feedback.

Benefits for sellers:

- Increase ratio of non-automatic to automatic ratings, thus attracting more future buyers.
- Increase buyer incentives to write detailed comments, thus increasing word-of-mouth marketing power.

Sellers can set:

- Reward for 1st high-quality feedback on newly listed products.
- Reward for any products, conditional on feedback being of high quality (and regardless of whether it is positive or negative).

Alternative form of rewards:

- Cash rewards.
- Discount coupon.

Figure Appendix-2: An example of automatic/ zero-word/effective feedback

Table Appendix-1: Seller rating grade

Seller rating score	Seller rating grade	Seller rating category
Below 4 points	0	none
4-10	1	1 heart
11-41	2	2 hearts
41-90	3	3 hearts
91-150	4	4 hearts
151-250	5	5 hearts
251-5000	6	1 diamond
501-1,000	7	2 diamonds
1,001-2,000	8	3 diamonds
2,001-5,000	9	4 diamonds
5,001-10,000	10	5 diamonds
10,001-20,000	11	1 crown
20,001-50,000	12	2 crowns
50,001-100,000	13	3 crowns
100,001-200,000	14	4 crowns
200,001-500,000	15	5 crowns
500,001-1,000,000	16	1 gold crown
1,000,001-2,000,000	17	2 gold crowns
2,000,000-5,000.000	18	3 gold crowns
5,000,001-10,000,000	19	4 gold crowns
Above 10,000.000	20	5 gold crowns