NBER WORKING PAPER SERIES

BUYING REPUTATION AS A SIGNAL OF QUALITY: EVIDENCE FROM AN ONLINE MARKETPLACE

Lingfang (Ivy) Li
Steven Tadelis
Xiaolan Zhou

Working Paper 22584
http://www.nber.org/papers/w22584

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2016

We thank Alibaba Group for providing us access to the data. We are grateful to to 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) and Alibaba Huoshui Research Plan. Xiaolan Zhou is grateful for support from the National Natural Science Foundation of China (grant #71603159). This paper was previously presented under the title "Does Money really talk? Evidences from Reward-for-Feedback Mechanism in Alibaba's C2C Platform". The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Buying Reputation as a Signal of Quality: Evidence from an Online Marketplace
Lingfang (Ivy) Li, Steven Tadelis, and Xiaolan Zhou
NBER Working Paper No. 22584
September 2016
JEL No. D47,D82,L15,L86

ABSTRACT

Reputation is critical to foster trust in online marketplaces, yet leaving feedback is a public good that can be under-provided unless buyers are rewarded for it. Signaling theory implies that only high quality sellers would reward buyers for truthful feedback. We explore this scope for signaling using Taobao's "reward-for-feedback" mechanism and find that items with rewards generate sales that are nearly 30%higher and are sold by higher quality sellers. 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.

Lingfang (Ivy) Li
Fudan University
School of Management
Shanghai, China 200433
lingfangli@fudan.edu.cn

Steven Tadelis
Haas School of Business
University of California, Berkeley
545 Student Services Building
Berkeley, CA 94720
and NBER
stadelis@haas.berkeley.edu

Xiaolan Zhou
Shanghai University of Finance and Economics
xiaolan.z.zhou@gmail.com

1 Introduction

The growth of trade in online marketplaces such as eBay, Amazon Marketplace, Taobao, Etsy, and others in the past two decades is remarkable not only because buyers are purchasing items that they cannot inspect, but purchasing is 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 purchasing in these large anonymous markets. Many studies provided evidence that buyers indeed respond to a seller's reputation in intuitive ways.[1]

A challenge to user-generated feedback is that leaving feedback is time consuming, implying that unless buyers enjoy the pro-social aspect of helping other future buyers, feedback is a public good that is most likely under-provided (Lafky, 2014). Hence, buyers may need to be incentivized to leave feedback. A market design question naturally arises: should marketplaces offer rewards for feedback or should they encourage sellers to do this? And if sellers offer rewards for feedback, how will this affect market outcomes?

Established theory can shed some light on these questions. Online goods are a form of "experience goods" for which quality can only be assessed after the purchase. Nelson (1974) argued that sellers with high quality goods will advertise their goods as a signal of quality, an idea that was formalized by Kihlstrom and Riordan (1984) and Milgrom and Roberts (1996). Intuitively, sellers of high quality goods will be willing to spend on advertising because they will benefit from repeat purchases by happy buyers, which is not the case for sellers with poor quality goods. We argue that, in a similar way, only high quality sellers who expect to receive positive feedback will be those who will pay buyers to leave feedback.

In this paper we use a unique dataset to empirically investigate this idea and shed light on several related questions. First, do sellers choose to signal their high quality by rewarding buyers for leaving feedback? Second, do buyers seem to infer sellers' quality from theirs 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?

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

Our data were obtained from Taobao, the world's largest online customer-to-customer marketplace with nearly 500 million registered users. On March 1, 2012, Taobao introduced a feedback reward mechanism called "Rebate-for-Feedback" (RFF).[2] 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 for 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 measured 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, however, because the RFF encourages buyers to leave feedback, high quality sellers will gain in the long run from being able to attract more buyers without having to reward them for feedback.[3]

Our proprietary panel data set consists of 6,992,131 transactions purchased from 12,857 randomly selected sellers who sold at least one unit between September 2012 and February 2013 on Taobao.com in four distinct categories: cellphones, memory cards (referred to as "TF cards"), cosmetic masks, and jeans. Our main findings are first, that RFF feature is chosen by higher quality sellers, implying that it can serves as a signal of quality. Second, sales of an item are about 30% higher when the seller chooses the rebate option, suggesting that buyers

[2]Taobao.com, run by Chinese e-commerce giant Alibaba Group Holding Ltd., is one of the world's largest shopping sites. 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).

[3]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).

understand the signaling effect of the RFF feature and act on it. Third, RFF induces buyers to write more detailed feedback but does not bias the feedback toward positive feedback.[4]

Our paper contributes to a growing literature that empirically studies the workings of reputation systems in online markets with asymmetric information. Though reputation systems seem to be critical to foster trust and online trade, a series of recent studies have shown that because feedback is user generated, reputation scores can be biased (Dellarocas and Wood, 2008; Nosko and Tadelis, 2015), inflated (Horton and Golden, 2015; Zervas et al., 2015) and where anyone can leave feedback, reputations may be manipulated by market players (Mayzlin et al., 2014). Our analysis suggests that RFF mechanisms can help promote honest and informative feedback, while at the same time offering the added benefit of a signaling mechanism that identifies high quality sellers. This further diminishes the inefficiencies of adverse selection in markets with asymmetric information.

Several papers have focused on the public good nature of feedback, proposing two ways to incentivize buyers to leave more feedback. One is that the online marketplace provide incentives to buyers to leave feedback (Miller et al., 2005; Fradkin et al., 2015), and another is that sellers provide the incentives (Li, 2010). The RFF mechanism implemented on Taobao follows the second approach. Li (2010) proposes and theoretically analyzes an unbiased RFF mechanism in an online auction market. She finds that, in equilibrium, both good and bad sellers choose to offer rebates as long as they care about future gains, and buyers avoid sellers who do not choose the rebate option and incorporate the rebate amount into their bids. Li and Xiao (2014) extend this idea to listed-price online markets and test the rebate mechanism in lab experiments. They find evidence consistent with the theoretical prediction 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 with more frequent and more favorable feedback, but the price of goods does not change with the rebate. The results from lab and field experiments provide mixed answers to

[4]Some sellers offer monetary rewards only for "positive" feedback using an ad in the delivered packages. There even exist sites that provide a service to "buy" positive feedback. See Xu et al. (2015).

whether rebate mechanism can lead to more feedback and more sales as well as whether such rebates will bias 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. This helps solve the classic problem of asymmetric information in online markets. Second, we provide compelling empirical evidences on the role of signaling in online markets, showing that sellers send credible signals and that buyers respond rationally.[5] This in turn results in matching consumers with better sellers or with higher quality goods, providing support to the isights of Nelson (1974). Last but not least, we contribute to the growing market design literature with respect to managing asymmetric information in online markets. Unlike Nosko and Tadelis (2015) and Masterov et al. (2015), who emphasize how marketplaces can manage the asymmetric information problem by inferring seller quality and intervening themselves, RFF mechanisms allow the marketplace to leverage the signaling abilities of high quality sellers.

The paper proceeds as the follows. In Section 2 we describe Taobao's RFF mechanism. In section 3 we lay out our hypotheses and describe our empirical identification approach. In section 4 we describe the data and present our analyses and results, which include a series of robustness checks. In Section 5 we conclude and discuss some limitations of our analysis.

2 The Taobao Reward-for-Feedback Mechanism

Launched in 2003, Taobao Marketplace (www.taobao.com) is the most popular consumer-to-consumer (C2C) online marketplace in China with nearly 500 million registered users.[6] On an average day, more than 60 million visitors have access to more than 800 million product listings, and an average of 48 thousand products are sold every minute.

[5]Two recent studies that use online marketplaces to provide evidence of signing equilibria are Backus et al. (2015) and Kawai et al. (2013).

[6]These are not unique registered users. The number of Chinese online-shoppers was estimated to be 187 million in 2011, with a forecast of 363 million by 2015.

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. A "transaction" on Taobao refers to a product-seller-buyer combination and represents a buyer's purchase (one or multiple units) of an item from a seller. After a transaction is completed, the buyer will receive a reminder to leave a rating for the item in her Taobao account page. The buyer can choose not to leave any rating, choose to rate positive/neutral/negative without leaving comments, or can choose to rate positive/neutral/negative and write comments.

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. 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. Second, if a buyer does not leave any feedback after the seller confirms that she received the item, the the system will leave an automatic positive rating for the seller and display "This is an automatic feedback left by the system on behalf of the buyer" (18 Chinese characters in the message) in the comment area for the transaction.[7] Third, an item's feedback history is recorded separately and can be easily seen when browsing the item. When a buyer leaves feedback for an item, the feedback is recored in the item's rating profile as well as in the seller's feedback profile. When a future buyer searches for a product, without browsing the seller's feedback profile, she can easily see all the ratings (including comments) for the item on the item's information page.

On March 1, 2012, Taobao launched a "Rebate-for-Feedback" (RFF) feature for sellers. 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. If a seller chooses this option then Taobao guarantees that the rebate is transferred from the seller's account to a buyer who leaves high-quality feedback. Feedback quality only depends on how informative it is, rather

[7]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.

than whether the feedback is positive or negative. Taobao measures the quality of feedback with a machine learning algorithm that examines the comment's content and length, finds out whether key features of the item are mentioned, etc.

Figure 1 describes Taobao's announcement of the new online service (our translation). Some notes are in order. One of the announced goals is to "increase the ratio of non-automatic ratings for sellers." Second, the goal of increasing "the quality of buyers' comments" is related to Taobao's use of a 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.[8] Figure 2 shows a Taobao.com page with the RFF feature.[9] 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 that this buyer has been awarded 0.50 RMB rebate for her feedback and reads, "This comment is informative, so it is rewarded with 0.50 RMB."

3 Hypotheses and identification

3.1 Hypotheses

To study the effect of the RFF mechanism on sales and feedback, we are interested in investigating the following two questions: First, do sellers choose RFF as a signal of high quality transactions in lieu of having an established good reputation? Second, how do sellers

[8]http://bbs.taobao.com/catalog/thread/513886-256229600.htm, accessed June 24, 2012.
[9]http://bbs.taobao.com/catalog/thread/513886-256229600.htm, accessed March 29, 2012.

benefit from RFF in terms of sales and feedback? To answer these questions we propose five hypotheses as described in detail below.

S1: *A seller is more likely to choose RFF when he is inexperienced. (Reputation Building Hypothesis)*

The theoretical literature suggests 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 "burn money" to promote brand awareness in its early stages (Milgrom and Roberts, 1996; Bagwell, 2007). Therefore, a seller will choose RFF as a signal to attract buyers when he or she is inexperienced, which will also be an investment in reputation. After accumulating more sales and building a higher reputation score, the seller will become less likely to choose RFF because, at the margin, more feedback becomes less valuable.

S2: *A seller who provides a high quality transaction is more likely to choose RFF. (Seller Signaling Hypothesis)*

As explained earlier in the introduction, the advertising literature (Nelson, 1974; Kihlstrom and Riordan, 1984; Milgrom and Roberts, 1996) implies that only a high quality seller who is confident enough that they will receive positive feedback will commit to reward feedback, assuming that consumers express their experiences truthfully. Therefore, a seller who plans to provide high quality transactions is more likely to choose RFF than a seller who plans to provide low quality transactions.

B1: *Choosing RFF increases an item's sales. (Buyer Belief Hypothesis - A)*

Choosing RFF may have both short run and long run effects on sales. In the short run, RFF acts as a signal because, as described above, choosing RFF shows a seller's confidence in his or her performance, and serves as a signal for high-quality. This is because a rebate is rewarded if it is informative, no matter whether it is positive or negative. As RFF attracts more buyers in the short run, it will increase the number of sales and detailed feedback, which, if positive, will attract more buyers in the long run.[10]

B2: *Choosing RFF increases an item's number of zero-word ratings. (Buyer Belief Hypothesis - B)*

If buyers infer that an item with the RFF feature has higher quality, then its sales will increase (hypothesis B1) but buyers for whom the cost of leaving feedback is high will ignore the rebate itself for its monetary value and only care about its signaling value. Because the requirement for a rebate to be given is an informative enough feedback, a buyer will not get the rebate if she only rates positive/neutral/negative and doesn't write any comments. If an item's number of zero-word ratings increase when the RFF feature is chosen, then it suggests that buyers are attracted by the signal of quality and *not* by the rebate. Hence, if buyers understand the signaling effect of RFF, then the RFF feature should increase an item's number of zero-word ratings.

B3: *Choosing RFF increases the average length of an item's ratings (Long Ratings Hypothesis).*

This hypothesis is quite trivial because informativeness is highly correlated with length, and Taobao's automatic algorithm offers the reward only for what they deem informative. Hence, any buyer who is interested in receiving a rebate must leave longer feedback.

3.2 Identification

To test hypotheses **S1** and **S2**, we run logit regressions where the left-hand side variable is whether or not a seller adopts the RFF feature and test whether adopting RFF is associated with a seller's experience and performance. The regressions include item fixed effects so that

[10]RFF can be regarded as a kind of price promotion conditional on leaving informative feedback. If the nominal price does not increase during the period of the rebate, then getting a rebate for feedback is, for buyers who leave feedback, a price discount. As we show below, the magnitude of the discount is very small compared to the sale price, implying that a serious "sales" effect is unlikely to be behind the results.

the decision to adopt RFF is identified *within* sellers-products. We also run the regressions without item-fixed effects to compare adoption across different sellers.

To test the four hypothesis regarding buyers' responses (**B1-B4**), we regress the variables of interests on rebate variables including item fixed effects, which control for product and seller characteristics. Recall that an item is a product-seller pair, and Taobao assigns a unique item ID to each item. The dataset represents a six month panel, which allows us to use item fixed effect to control for unobserved product and seller characteristics. During this period, a seller may vary his adoption of a rebate for a product. 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.

The tests for hypotheses **B1** (Buyer Belief Hypothesis - A) and **B3** are straightforward by testing whether item sales and the ratio of long ratings are significantly higher when the RFF feature is adopted. To test for hypotheses **B2** (Buyer Belief Hypothesis - B), we test whether the number of ratings with zero Chinese characters is significantly increased with RFF adoption. In the main body of the paper, we use our entire sample data.

3.3 Endogeneity Concerns

There are two sources of endogeneity in seller decisions: prices and RFF adoption. To address the potential problem of price endogeneity, we use an instrumental variables approach in our analysis. Commonly adopted instrumental variables for prices include cost variables, observed exogenous product characteristics, and the number of products in the same market.[11] 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

[11] Berry et al. (1995) suggest 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.

quantiles, and then select items with non-zero sales in the same product category, same price quantile, and the same province.[12]

Turning to the seller's choice of adopting the RFF feature, this is our endogenous (strategic) variable of interest. 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 changes in behavior as a response to the RFF is identified off of variation *within* seller and *within* product by the definition of the item ID. This is akin to a variant of the kind of "quasi-experimental" approach first used by Elfenbein et al. (2012) who study how sellers on eBay use charity as a substitute for reputation.[13] Hence, if product characteristics may drive some results, then our use of item fixed effects within our panel structure should alleviate any such concerns.[14]

4 Data Description

Our data consist of 6,992,131 transactions purchased from 12,857 randomly selected sellers who sold at least one unit in our four chosen categories (cellphones, TF cards, cosmetic masks, and jeans) between September 2012 and February 2013 on Taobao. Theses categories represent both search goods and experience goods in various price ranges. Cellphones and TF cards are search goods because product quality is readily observable prior to purchase (as

[12]We use the whole sample period as the window to select non-zero sales items, instead of using a single month as the window, to better include all competitive items in the same market. We use the same province because many buyers sort search outcomes by the province in which sellers are located and purchase from sellers nearby to reduce shipping time and costs.

[13]They identify a "quasi-experiment" 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) similarly use "matched listings" with the same title and subtitle, and rely on variation within matched listings to investigate various sale strategies in eBay.

[14]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. Again, this is basically the same approach taken by Elfenbein et al. (2012).

long as the seller provides authentic products.) 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 or value can only be truly determined by using them. In the data set, there are 107,620 items in the four categories with positive sales in the sample period.[15] Among the 12,857 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 refers to a product-seller combination, the same product sold by different sellers will be assigned different item IDs. For each item, our dataset contains all transactions during the sample period. A *transaction* is defined as a sale of an item between a buyer and a seller (one or multiple units). 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 (number of Chinese characters), time stamp, and whether it is a first- or second-time rating.[16] Table 2 provides summary statistics of item attributes at the item-month level. This is restricted to include only observations with positive sales of an item in a month because zero sales may mean that an item is out of stock or is not on sale at all.

We define an "effective rating" as a rating that includes real information from the buyer who left it. These exclude automatic Taobao ratings as described in Section 2 (with 18 Chinese characters) as well as ratings with no comments (zero Chinese characters). By excluding ratings with 18 characters we are discarding some ratings that are from genuine buyers but these are no more than one percent of ratings (see Figure 4). We define a "long rating" to be a rating with at least 24 Chinese characters. We pick this threshold because it is the 75 percentile of all the ratings that exclude 18 character ratings.

For each seller our dataset contains information about his location (province), his service promises (e.g., whether he accepts any returns in 7 days after the sale, etc.) and reputation

[15] The number of transactions includes all sales by sellers in our sample, not limited to the four categories.

[16] 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.

data at a daily level, including his seller rating score, his corresponding seller rating grade (see table 1), and the ratio of positive ratings.[17]

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 last 30 days. A seller's ratio of positive ratings is reported for the last month and last six months. For each item, we summarize its product characteristics as monthly sales, monthly average price, 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 frams as the period duration, and the main findings still hold.

There are two types of *rebates*, depending on whether a seller chooses a rebate in the form of cash or coupon.[18] For all transactions with the RFF feature we know whether the buyer received a cash or coupon rebate. About 20.11% of rebates are in the form of cash and the average value of cash rebates is RMB 1.411. Since the conditions of using a coupon are missing in the data, we only use the value of cash rebates. Also, since the start time of a rebate is recorded in our 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 times in the month, otherwise it is defined as a coupon rebate.

Figure 4 shows the distribution of the length of ratings in Chinese characters with and without a rebate, and how the ratio of these two distributions changes over time. For *any* rating length there are more with rebates than without rebates, and the ratio between the

[17] We have the sellers' reputation information on March 3rd, 2013, and we combine it with transactions data to calculate each item and seller's daily reputation during the sample period.

[18] 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.

two rises as the rating length increases. This suggests that rebates play a role in motivating people on the extensive margin to write more comments and on the intensive margin to write longer comments. Figure 5 provides a first glance of the relationship between rebates and sales, as well as the relation 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 monthly ratings of zero Chinese characters. The average monthly ratio of effective ratings of an item with a rebate is higher than without a rebate, the same applies to long ratings, however, positive ratings of an item with a rebate are almost the same as without a rebate.

The raw data suggests that the sale of items, the length of ratings and effective ratings are positively correlated with offering a rebate. To further investigate the sellers' rebate choices and the buyers' responses, we conduct a more nuanced empirical analysis.

5 Empirical analysis

5.1 Rebate Adoption

To study a seller's strategic decision to signal his quality and build a reputation we start with investigating a seller's adoption of a rebate on a particular item. 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:

$$y_{i,s,t} = \alpha + \pi \cdot Product_{i,s,t-1} + \gamma \cdot Seller_{s,t-1} + \beta Positive\_Ratio_{s,t} + \delta_i + \mu_t + \varepsilon_{i,s,t}, \qquad (1)$$

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; $Seller_{s,t-1}$ is a vector of seller characteristics, including seller rating grade in period $t-1$; $Positive\_Ratio_{s,t}$ refers to the seller's ratio of positive ratings in period t, which is the number of positive ratings divided by the number of all ratings in month t. Unlike the analysis of the impact of a rebate in the next subsection, we use the ratio of seller positive 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.[19] δ_i and μ_t are item and month fixed effects to control for unobserved attributions of items and months.

Table 3 shows the results of the regression in equation (1). Column 1 uses item fixed effects and identifies the effects within seller-product pairs, while column 2 shows the results excluding item fixed effects in order to compare behavior across different products and sellers.

As column 1 shows, a seller tends to choose RFF for an item when his rating grade is low, which confirms the *Reputation Building Hypothesis* (**S1**). The higher a seller rating grade, the lower the probability of adopting a rebate. 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 "seller rating grade, $t-1$" is -0.1787, so if the seller's rating grade last month is one level higher, then R changes by a factor of $\exp\{-0.1787\} = 0.836$. Hence, the likelihood of choosing a rebate declines if the seller gets a higher rating grade (identified by within-seller changes in ratings.) Similarly, the estimated coefficient for "Ratio of seller positive ratings, t" is 3.9279 means that a 1% increase in the seller's ratio of positive ratings this month implies that R changes by a factor of $\exp\{0.0393\} = 1.04$, and the likelihood of choosing a rebate increases. We consider the seller's ratio of positive ratings in month t as a proxy for the quality of his transactions in month t. The positive coefficient on this variable suggests that sellers use the RFF feature as a signal of quality. This confirms the *Seller Signaling Hypothesis* (**S2**).

Column 2 identifies effects across items and sellers. Compared to Product characteristics in column 1, the coefficient on the number of negative/neutral ratings in the past month

[19] We have used the period $t-1$ ratio as a robustness test, and the results are almost identical. Also, we do not include the item's ratio of positive ratings for two reasons. First, it is highly correlated with other product characteristics, such as sales, and number of negative and neutral ratings. Second, it tends to fluctuate a lot if there are few ratings. One single negative rating from a deliberate buyer can damage the item's ratio of positive ratings significantly.

becomes significantly negative and the coefficient on the number of positive long ratings in the last month for an item becomes significantly positive. 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 signaling story. Namely, sellers choose their *best* items to further guarantee a higher likelihood of making their customer happy so that she leaves a positive review.

Like in column 1, column 2 shows that the coefficient on a seller's ratio of positive ratings in the same month is still significantly positive. However, the coefficient on a seller's past rating grade reverses sign and becomes significantly positive. As explained above, the within seller-item regression of column 1 suggests that sellers use the RFF feature to signal their quality and improve their reputation, while the result in column 2 suggests that sellers with a higher seller rating grade may be more savvy about using this feature. 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, the excluded category is jeans and all item coefficients are negative and significant, and of these, the least negative on is for cosmetic masks. All else equal, the likelihood of RFF adoption from largest to smallest is jeans, cosmetic mask, cellphone and memory cards, respectively. This suggests that a seller is more likely to choose the RFF feature for *experience goods* and *expensive goods*. This further provides support for the signaling hypothesis in the spirit of the original insights in Nelson (1974).[20]

5.2 The Impact of Rebates on Sales and Ratings

To examine 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,

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

[20] As mentioned earlier, we have used the period $t-1$ ratio of positive ratings as a measure of seller quality as a robustness test, and the results are almost identical as shown in columns 3 and 4.

where $y_{i,s,t}$ is the dependent variables of interest; $Rebate_{i,s,t}$ is a vector that indicates the rebate status of item i sold by seller s in period t, which includes an indicator that takes the value of 1 if item i of seller s offers any rebate in period t, an indicator that takes the value of 1 if the rebate is in the form of cash, and the average value of a cash rebate that month; 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.

For product and seller characteristics, we select the variables that are observable to buyers who view an item page. A buyer can choose to sort searched items by Taobao's default ranking, popularity, sales (in the past 30 days), ratio of positive ratings, province, or price. An item page shows its price, whether shipping is free, number of past sales, number in stock, all ratings (positive, neutral, or negative) with comments for the item, including both first and second time ratings, the seller's name, and the seller's rating. A customer can also choose to see all ratings, including ratings with no comments or automatic positive ratings. We pool an item's neutral and negative ratings together because a lot of buyers tend to give a neutral rating rather than a negative rating in order to avoid retaliation even if a buyer has an unhappy shopping experience.[21]

Because long ratings are more informative than short ratings for buyers to make decisions, we also include the number of positive long ratings (at least 24 characters) and the number of negative and neutral long ratings of an item in product characteristics. For seller characteristics, we include seller rating grade and ratio of seller positive ratings in the previous month because only the lagged value is observed by buyers.

5.2.1 Buyer Responses to Rebates

Table 4 uses ln(sales) as the dependent variable in equation (2) above. Column 1 does not use instruments for prices, while columns 2 through 5 do. Comparing columns 1 and 2, when using instruments for prices, the coefficient on ln(price) changes from an inelastic (-0.4059)

[21] On Taobao, a seller knows the buyer's contact information including cell phone, so some sellers would call buyers and repeatedly ask them to change negative feedback.

to an elastic (-3.1802) demand. Given the large amount of competition on the Taobao site, this confirms that the IV strategy we chose for prices was effective.

The estimated coefficient on the rebate dummy is large and significant, showing that a rebate increases the quantity sold of an item by 28.73% on average. This supports the *Buyer Belief Hypothesis - A* (**B1**). We also estimated the effects of different types of rebates and different values of a cash rebate. Column 3 shows that on average, a coupon rebate increases the quantity sold of an item by 31.34% while a cash rebate increases the quantity sold by 21.42%. We conjecture that this is probably because a coupon usually has a higher value compared to the average value of a cash rebate.[22] Column 4 shows that sales increase in the value of a cash rebate. A cash increase of RMB 1 increases the quantity sold by 31.48%.

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, all attract more sales. 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. The upshot is that the effect of a rebate on sales is large. Using the estimates of column 2 we find that offering a rebate is, on average, as effective as (1) reducing price by about 9% (calculated by dividing the coefficients, $0.2873/3.1802 = 0.0903$); (2) increasing last month's sales by 389.30% ($0.2873/0.0738 = 3.8930$); or (3) increasing the number of positive long ratings in the last month by 705.90% ($0.2873/0.0407 = 7.0590$).

Table 5 has the number of ratings with zero characters as the dependent variable in equation (2). By the nature of the RFF feature, buyers whose ratings contain no Chinese characters can never obtain the rebate according to Taobao's conditions. Hence, as we explain in Section 3.1, buyers who understand the signaling value of rebates but are uninterested in obtaining the rebate may still flock to sellers who offer rebates because of their signaling content. Column 2 of Table 5 (instrumenting for prices) shows that rebates increase ratings

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

with zero Chinese characters by 14.24%. Column 3 shows that a coupon rebate generates more ratings with zero Chinese characters than a cash rebate (16% versus 9.2%.) Like for sales, a coupon rebate seems more effective than a cash rebate as a signal. Column 4 shows that increasing a cash rebate by RMB 1 raises the number of ratings with zero Chinese characters by 15.76%. These results confirm the *Buyer Belief Hypothesis - B* (**B2**).

We further consider the heterogeneous effect of rebates across product categories. Column 5 in tables 4 and 5 reports estimates from dummies for each product category. We find that a rebate has the largest effect on cellphones and jeans, and the lowest on SD cards. For example, in table 4, the estimated coefficient for a rebate on cell phones is 0.3992, and for jeans is 0.3931, while for masks is only 0.1326, and for SD cards is 0.0667. The same differences appear in column 5 in table 5. 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 experienced good (e.g., jeans, masks).

5.2.2 Effect of Rebates on Ratings

Tables 6 and 7 use a variety of rating measures as the dependent variable in equation (2). We define the *ratio of effective ratings* as the number of effective ratings divided by the number of all transactions. Columns 1 and 2 in Table 6 show that offering a rebate raises an item's ratio of effective ratings by 6.11%, and that coupon rebates have a larger effect that cash rebates (6.82% versus 4.09%). As column 3 shows, the effects are strongest for jeans and cellphones, weaker for masks, and nonexistent for SD cards, similar to previous patterns.

We define the *ratio of long ratings* as the number of long ratings divided by the number of all transactions. Columns 4 and 5 in Table 6 show that offering a rebate raises an item's ratio of long ratings by 7.35%, and that coupon rebates have a larger effect that cash rebates (8.02% versus 5.42%). This confirms the obvious *Long Ratings Hypothesis* (**B3**). Again, as column 6 shows, the effects are strongest for jeans and cellphones, weaker for masks, and nonexistent for SD cards. Columns 7-9 in Table 6 consider the log of the number of long ratings and show very similar results.

Columns 1-6 in Table 7 explore whether rebates change the likelihood of receiving positive feedback. 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 a positive effect of feedback on the ratio of positive ratings. One may be concerned that offering rebates causes buyers to feel a need to reciprocate and offer more positive feedback than the seller deserves.[23] 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.

Columns 4-6 of Table 7 drop item fixed effects, allowing the quality of sellers to vary with and without rebates. As the results in columns 4-6 show, all but one of the coefficients (on masks in column 6) are positive, implying a positive correlation between the ratio of positive ratings and the use of RFF. This is consistent with the results in Table 3 and further confirms the *Seller Signaling Hypothesis* (**S2**).

Interestingly, as columns 7-9 in table 7 show, offering a rebate gives the seller earlier ratings, shortening the time to rating by 6.08%. A coupon shortens the period by 7.02% while a cash rebate shortens it by 4.10%. 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 cellphones, weaker for masks, and nonexistent for SD cards

5.3 Robustness Checks

5.3.1 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 good with more risk involved in the purchase, like experience goods and more expensive goods, we

[23] Cabral and Li (2015) find some support for reciprocal behavior in that a buyer's feedback is positively biased if a rebate is offered by the seller in a field experiment on eBay. Unlike in our study, in Cabral and Li (2015) there is no enforcement by the marketplace to pay the rebate so buyers may believe that they will not get the rebate if they leave negative feedback. Our study is similar to the experimental setup as in Li and Xiao (2014) where the experimenter enforces the payment of rebates.

estimate the average impact of a rebate for each category as a robustness check. We divide our data into four categories: cellphones, TF cards, masks, and jeans, and run the panel regressions within each category of sellers and items again.

Using sellers' fixed effects as controls, 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 *S1*. Dropping seller fixed effects, we find that sellers with a high ratio of positive ratings are more likely to adopt a rebate in each category, which proves the Signaling Hypothesis *S2*.

Similarly to our findings with the full sample of goods, we find for each category that rebates increase sales, the ratio of long ratings, and the number of long ratings, while they shorten the days between transaction and rating. A rebate is usually accompanied with a price reduction for masks and jeans. A rebate has almost no effect on a seller's ratio of positive ratings for each category. Figure 6 shows the 95% confidence intervals of the estimates of changes in our main variables of interest. The effect of a rebate on sales from the highest to the lowest are cellphones, Jeans, TF cards, and mask. A rebate increases sales of cellphones, TF cards, masks, and jeans by 33.89%, 3.02%, 27.96%, and 32.25% respectively.[24] The effect of a rebate on the ratio of long ratings from the highest to the lowest are masks, jeans, TF cards, and cellphones. A rebate increases the ratio of long ratings of cellphones, TF cards, masks, and jeans by 7.14%, 2.99%, 7.99%, and 7.74% respectively.[25]

5.3.2 Alternative Period Windows

Recall that we only observe when the RFF offering started and do not observe when it ended. In the analysis in section 5.2 and section 5.3, we use "month" as a period window and user item-month as the observation measurement. One concern is that if an item's rebate period

[24] All the results for for these regressions, as well as for differing time periods described below, are omitted for space considerations and available upon request.

[25] The coefficient on TF cards is not statistically significant. The sample size of TF cards is 5,442 in the regressions for sales, which is much smaller than other products (22,673 cellphones, 86,135 masks, and 99,830 jeans). Also, the sales of TF cards are affected by factors not included in our model and have large variations, because a TF card is usually purchased together with a cellphone rather than independently.

is less than a month, our results may be biased. To address this concern we used 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 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. The regression results of table 4-6 are similar to the results obtained by using two-weeks as a period.

Another concern may be that the periods are adjacent, implying that the first days of period t are a lot close to the last days in period $t-1$ then they are to most days in period t. To address this concern we conducted 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 change.

6 Conclusion

Te burgeoning growth of online marketplaces, and the increased access to data from them, offers 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 a RFF mechanism. Our empirical evidence suggests that higher quality sellers use RFF to send signals, and that buyers respond to these signals rationally, which in turn alleviates some of the adverse selection 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 others 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 signaling content of RFF. We also show that buyers respond to the RFF signal in ways that are consistent with our signaling story. We find that sales of an item are about 30% higher when the seller chooses a RFF. It is as effective as increasing last month's sales by four times, or increasing the number of positive long ratings in the last month by nine times. Furthermore, the use of RFF do not create bias in feedback, and they afford the seller the opportunity to build up a good reputation faster, creating a sort of "flywheel." 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.

Turning to market design consequences of our study, the results of our analyses suggest that marketplaces can help reduce the asymmetric information problem by letting sellers engage in RFF signaling practices. This suggests that online marketplaces can rely on the strategic sophistication of both sellers and buyers to manage the asymmetric information problem by leveraging the signaling abilities 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 a marketplace's quality, relaxing some of the need for external forms of regulation.

References

Backus, M., Blake, T., and Tadelis, S. (2015). Cheap talk, round numbers, and the economics of negotiation. NBER working paper w21285.

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.

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.

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.*

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.

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. (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).

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

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. *Working paper, Boston University*.

Table 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

Table 2: Summary of Statistics (item-month)

	Obs.	Mean	Std.	25%	Median	75%.
Dummy if a rebate	268,245	0.184	0.388	0	0	0
Dummy if a coupon rebate	268,245	0.139	0.345	0	0	0
Dummy if a cash rebate	268,245	0.0457	0.209	0	0	0
Amount of cash rebate (RMB) conditional on a cash rebate	6,255	1.411	2.235	0.5	1	1.1
Monthly sales (quantity)	268,245	41.09	709.6	1	3	11
Monthly sales (RMB)	268,245	4,377	78,930	75	198	746
Average transaction price (RMB)	268,245	172.6	899.4	28.4	68	116.6
Average transaction price: Cellphone	28,281	963.6	2229	226.5	510	1150
Average transaction price: TF card	6,789	133.3	1898	28	40.47	79.55
Average transaction price: Mask	105,823	47.84	504.7	5.6	16.92	55.24
Average transaction price: Jeans	127,352	102.61	199.7	60.39	80	114
Number of ratings with zero Chinese characters	268,245	5.998	149.5	0	1	2
Number of positive ratings	268,245	18.29	434.7	1	2	5
Number of negative and neutral ratings	268,245	0.304	17.02	0	0	0
Number of positive long ratings	268,245	4.003	99.54	0	0	1
Number of neutral and negative long ratings	268,245	0.146	9.962	0	0	0
Ratio of effective ratings	223,122	0.558	0.378	0.25	0.571	1
Ratio of positive ratings	223,122	0.989	0.077	1	1	1
Average number of days before leaving ratings	225,068	9.786	6.141	5	8.429	13.67
Ratio of long feedback	223,122	0.208	0.305	0	0	0.333
Number of long ratings	268,245	4.149	105.3	0	0	1
Seller rating grade*	258,129	9.866	3.505	8	10	12
Ratio of seller positive ratings*	267,588	0.990	0.0193	0.988	0.995	1

Note: The observations are at the item-month level.

An item in a month is excluded if the sales is zero.

A long rating is an evaluation with the number of Chinese characters no less than 24.

An effective rating is an evaluation with the number of Chinese characters not equal to 0 nor 18.

* These two statistics are calculated including sales made by the seller that are not used in our analyses.

Table 3: Adoption of a rebate for an item

Dependent variable: Logit indicator = 1 if adopting a rebate	(1)	(2)	(3)	(4)
Seller characteristics				
Seller rating grade, $t-1$	-0.1787***	0.0585***	-0.1828***	0.0646***
	(0.0125)	(0.0015)	(0.0128)	(0.0015)
Ratio of seller positive ratings, t	3.9279***	5.8356***		
	(0.5839)	(0.3446)		
Ratio of seller positive ratings, $t-1$			4.7729***	5.6680***
			(0.6413)	(0.3764)
Product characteristics				
ln(sales, $t-1$)	0.2941***	0.2389***	0.3018***	0.2409***
	(0.0103)	(0.0052)	(0.0103)	(0.0052)
ln(number of negative & neutral ratings, $t-1$)	-0.0777	-0.2424***	-0.0517	-0.2283***
	(0.0564)	(0.0364)	(0.0568)	(0.0367)
ln(number of positive long ratings, $t-1$)	-0.1005***	0.4710***	-0.1068***	0.4669***
	(0.0179)	(0.0101)	(0.0180)	(0.0102)
ln(number of negative and neutral long ratings, $t-1$)	0.0942	0.0564	0.0917	0.0537
	(0.0745)	(0.0485)	(0.0748)	(0.0487)
Item category				
Dummy if cellphone		-0.5034***		-0.5154***
		(0.0192)		(0.0193)
Dummy if TF card		-1.2403***		-1.2542***
		(0.0516)		(0.0522)
Dummy if mask		-0.2964***		-0.3007***
		(0.0116)		(0.0117)
Item fixed effect	Yes	No	Yes	No
Month fixed effect	Yes	Yes	Yes	Yes
Observations (item-month)	136,932	502,289	135,937	505,467

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 4: Impact of rebate on sales of an item

	Dependent variable:	ln(sales (quantity), t)			
	(1)	(2)	(3)	(4)	(5)
Rebate					
Dummy if a rebate, t	0.3145***	0.2873***			
	(0.0069)	(0.0091)			
Dummy if a coupon rebate, t			0.3134***	0.3125***	
			(0.0102)	(0.0105)	
Dummy if a cash rebate, t			0.2142***		
			(0.0166)		
Value of a cash rebate, t				0.3148***	
				(0.0168)	
Dummy if a rebate for cellphone, t					0.3992***
					(0.0289)
Dummy if a rebate for TF card, t					0.0667
					(0.0825)
Dummy if a rebatene for mask, t					0.1326***
					(0.0139)
Dummy if a rebatene for jeans, t					0.3931***
					(0.0114)
Product characteristics					
ln(price, t)	-0.4059***	-3.1802***	-3.1638***	-3.3367***	-2.8623***
	(0.0102)	(0.2436)	(0.2428)	(0.2586)	(0.2300)
ln(sales, $t-1$)	0.0718***	0.0738***	0.0738***	0.0739***	0.0730***
	(0.0026)	(0.0033)	(0.0033)	(0.0034)	(0.0032)
ln(no. of ngtv & ntrl ratings, $t-1$)	-0.0659***	-0.0675***	-0.0678***	-0.0686***	-0.0684***
	(0.0145)	(0.0183)	(0.0182)	(0.0189)	(0.0175)
ln(no. of pstv long ratings, $t-1$)	0.0407***	0.0374***	0.0371***	0.0402***	0.0375***
	(0.0046)	(0.0058)	(0.0058)	(0.0060)	(0.0056)
ln(no. of ngtv & ntrl long ratings, $t-1$)	-0.0060	-0.0079	-0.0081	-0.0099	-0.0062
	(0.0186)	(0.0236)	(0.0235)	(0.0243)	(0.0226)
Seller characteristics					
Seller rating grade, $t-1$	-0.0315***	-0.0150**	-0.0160**	-0.0149**	-0.0168***
	(0.0039)	(0.0052)	(0.0051)	(0.0054)	(0.0049)
Ratio of seller positive ratings, $t-1$	0.9880***	1.2406***	1.2497***	1.1968***	1.1831***
	(0.1692)	(0.2153)	(0.2148)	(0.2244)	(0.2061)
IVs	No	Yes	Yes	Yes	No
Item fixed effect	Yes	Yes	Yes	Yes	Yes
Month fixed effect	Yes	Yes	Yes	Yes	Yes
Observations (item-month)	214,080	214,080	214,080	209,846	214,080

Note: Standard deviations are in the parentheses and are clustered on seller in models without IVs
Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).
F-statistics of the 1st stage regressions for ln(prices) and amount of cash rebate are 56.78 and 11,474 respectively.
We add 1 to all the number of ratings in order to avoid $ln(0)$.

Table 5: Impact of rebate on ratings with zero Chinese characters

Dependent variable:	ln(ratings with zero Chinese characters, t)				
	(1)	(2)	(3)	(4)	(5)
Rebate					
Dummy if a rebate, t	0.1643***	0.1424***			
	(0.0094)	(0.0073)			
Dummy if a coupon rebate, t			0.1604***	0.1595***	
			(0.0082)	(0.0084)	
Dummy if a cash rebate, t			0.0920***		
			(0.0134)		
Value of a cash rebate, t				0.1576***	
				(0.0134)	
Dummy if a rebate for cellphone, t					0.1861***
					(0.0236)
Dummy if a rebate for TF card, t					0.0263
					(0.0673)
Dummy if a rebatene for mask, t					0.0544***
					(0.0113)
Dummy if a rebatene for jeans, t					0.2055***
					(0.0093)
Product characteristics					
ln(price, t)	-0.1718***	-2.4079***	-2.3966***	-2.4886***	-2.2176***
	(0.0207)	(0.1964)	(0.1958)	(0.2065)	(0.1875)
ln(sales, t-1)	0.1771***	0.1787***	0.1787***	0.1792***	0.1782***
	(0.0040)	(0.0027)	(0.0026)	(0.0027)	(0.0026)
ln(no. of ngtv & ntrl ratings, $t-1$)	-0.0381**	-0.0394**	-0.0396**	-0.0389**	-0.0399**
	(0.0157)	(0.0147)	(0.0147)	(0.0151)	(0.0143)
ln(no. of positive long ratings, $t-1$)	0.0569***	0.0543***	0.0540***	0.0559***	0.0543***
	(0.0068)	(0.0047)	(0.0047)	(0.0048)	(0.0045)
ln(no. of ngtv & ntrl long ratings, $t-1$)	0.0359*	0.0344*	0.0342*	0.0318	0.0353*
	(0.0212)	(0.0190)	(0.0190)	(0.0194)	(0.0184)
Seller characteristics					
Seller rating grade, $t-1$	-0.0201**	-0.0068	-0.0075*	-0.0076*	-0.0078*
	(0.0067)	(0.0042)	(0.0042)	(0.0043)	(0.0040)
Ratio of seller positive ratings, $t-1$	0.2884	0.4920**	0.4983**	0.4510**	0.4606**
	(0.1868)	(0.1735)	(0.1732)	(0.1792)	(0.1680)
IVs	No	Yes	Yes	Yes	No
Item fixed effect	Yes	Yes	Yes	Yes	Yes
Month fixed effect	Yes	Yes	Yes	Yes	Yes
Observations (item-month)	214,080	214,080	214,080	209,846	214,080

Note: Standard deviations are in the parentheses and are clustered on seller in models without IVs. Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).

Table 6: Impact of rebate on effectiveness and length of ratings

Dependent variable	Ratio of effective ratings, t			Ratio of long ratings, t			ln(number of long ratings, t)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rebate									
Dummy if a rebate, t	0.0611***			0.0735***			0.2898***		
	(0.0036)			(0.0027)			(0.0072)		
Dummy if a coupon rebate, t		0.0682***			0.0802***			0.3167***	
		(0.0041)			(0.0030)			(0.0081)	
Dummy if a cash rebate, t		0.0409***			0.0542***			0.2145***	
		(0.0067)			(0.0050)			(0.0132)	
Dummy if a rebate for cellphone, t			0.0652***			0.0790***			0.3803***
			(0.0119)			(0.0090)			(0.0223)
Dummy if a rebate for TF card, t			0.0182			0.0305			0.0974
			(0.0350)			(0.0264)			(0.0636)
Dummy if a rebatene for mask, t			0.0392***			0.0596***			0.1679***
			(0.0056)			(0.0042)			(0.0107)
Dummy if a rebatene for jeans, t			0.0795***			0.0853***			0.3759***
			(0.0047)			(0.0036)			(0.0088)
IVs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Item fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations (item-month)	177,411	177,411	177,411	177,411	177,411	177,411	214,080	214,080	214,080

Note: Standard deviations are in the parentheses.
Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).
Standard deviations are clustered on seller in models without IVs.
We add 1 to all the number of ratings in order to avoid $ln0$.
All models include product and seller characteristics (same as models in table 4) as control variables.

Table 7: Impact of rebate on bias and time of 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**			0.0139***			-0.0608***		
	(0.0007)			(0.0028)			(0.0051)		
Dummy if a coupon rebate, t		-0.0014*			0.0054**			-0.0702***	
		(0.0008)			(0.0018)			(0.0057)	
Dummy if a cash rebate, t		-0.0022*			0.0409***			-0.0343***	
		(0.0013)			(0.0076)			(0.0093)	
Dummy if a rebate for cellphone, t			-0.0028			0.5628**			-0.0800***
			(0.0024)			(0.1733)			(0.0167)
Dummy if a rebate for TF card, t			-0.0004			0.1103**			-0.0593
			(0.0069)			(0.0392)			(0.0482)
Dummy if a rebate for mask, t			-0.0009			-0.1867**			-0.0259**
			(0.0011)			(0.0586)			(0.0079)
Dummy if a rebate for jeans, t			-0.0019**			0.0793**			-0.0864***
			(0.0009)			(0.0252)			(0.0066)
IVs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
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)	177,411	177,411	177,411	177,411	177,411	177,411	180,359	180,359	180,359

Note: Standard deviations are in the parentheses.
Asterisks indicate significance at 10% (*), 5% (**) and 1% (***).
Standard deviations are clustered on seller in models without IVs.
We add 1 to all the number of ratings in order to avoid $ln0$.
All models include product and seller characteristics (same as models in table 4) as control variables.

Figure 1: Taobao's announcement of a new feedback reward system

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 2: Taobao.com page with feedback reward scheme

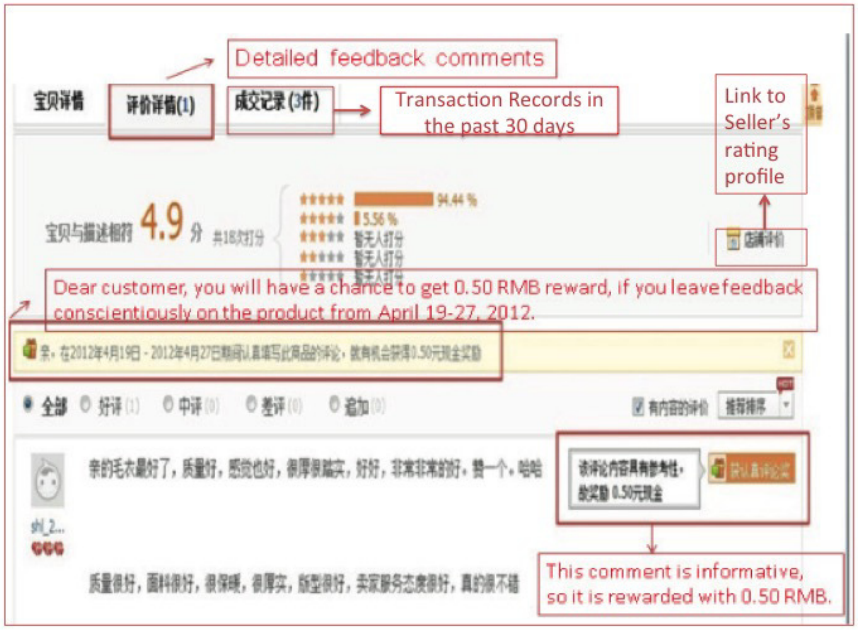
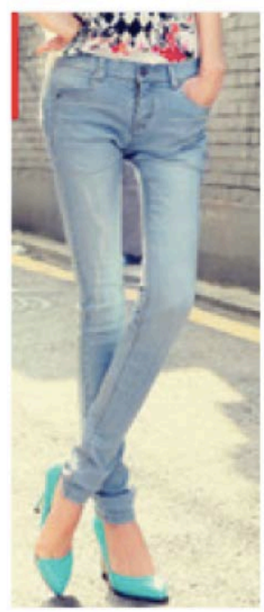


Figure 3: Four Item Categories in our Data

Cell Phones

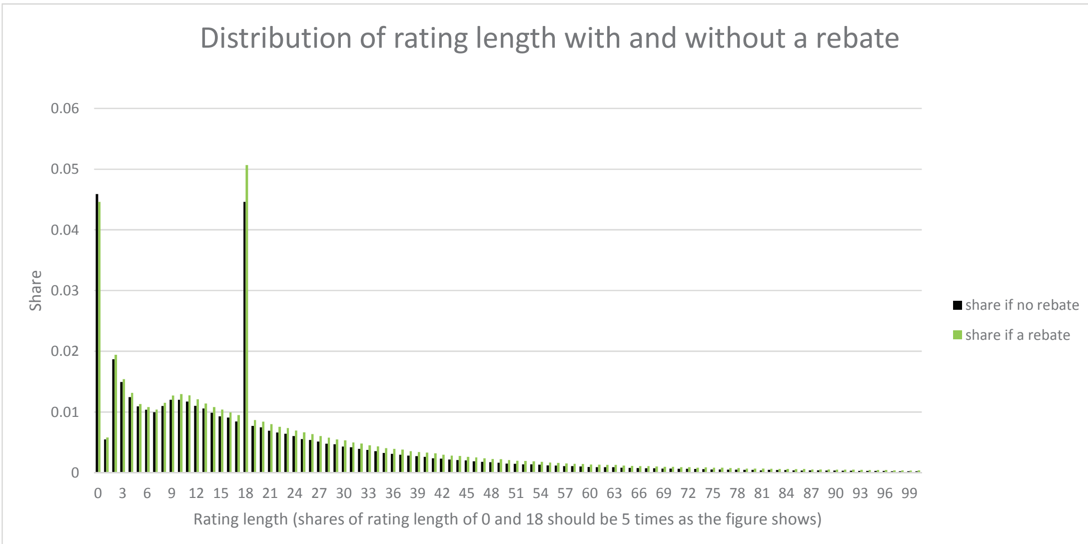
Memory Cards

Cosmetic Masks



Jeans

Figure 4: Ratings with and without a rebate



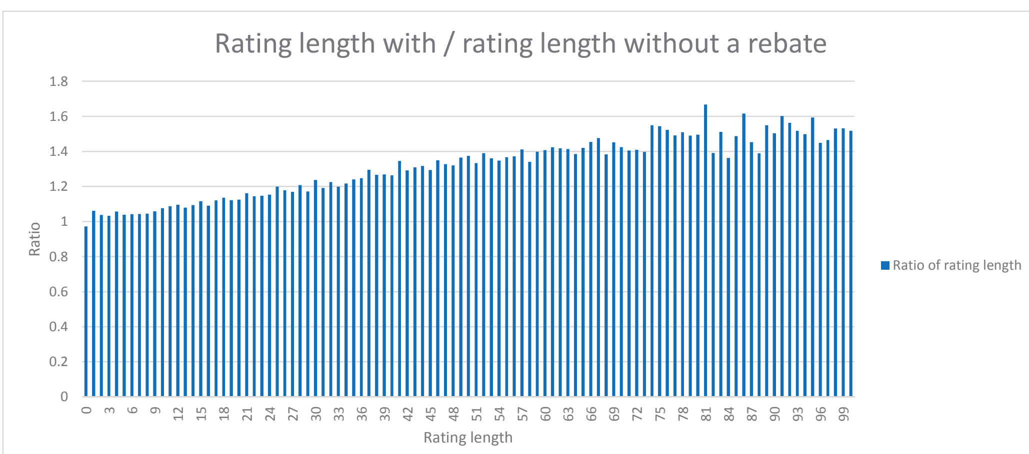
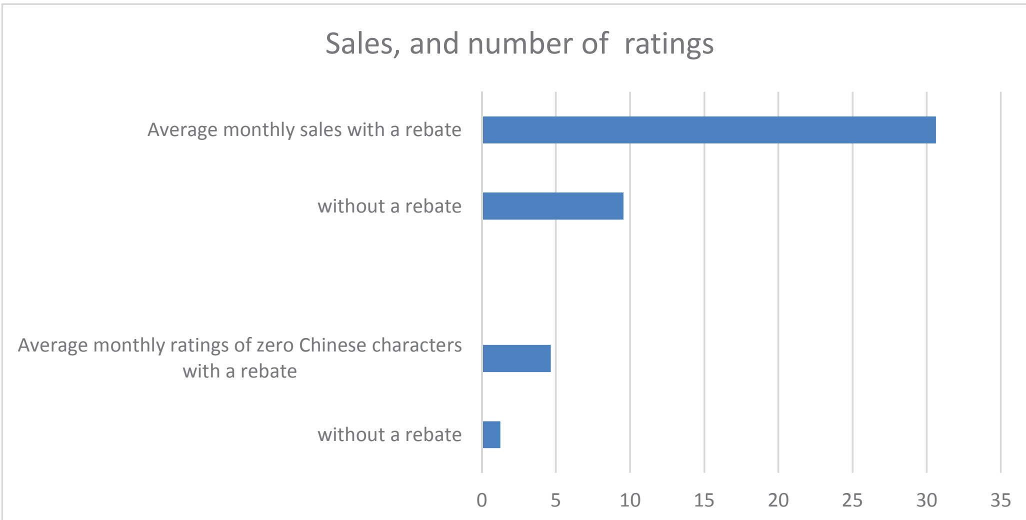


Figure 5: Sales and ratings with and without a rebate



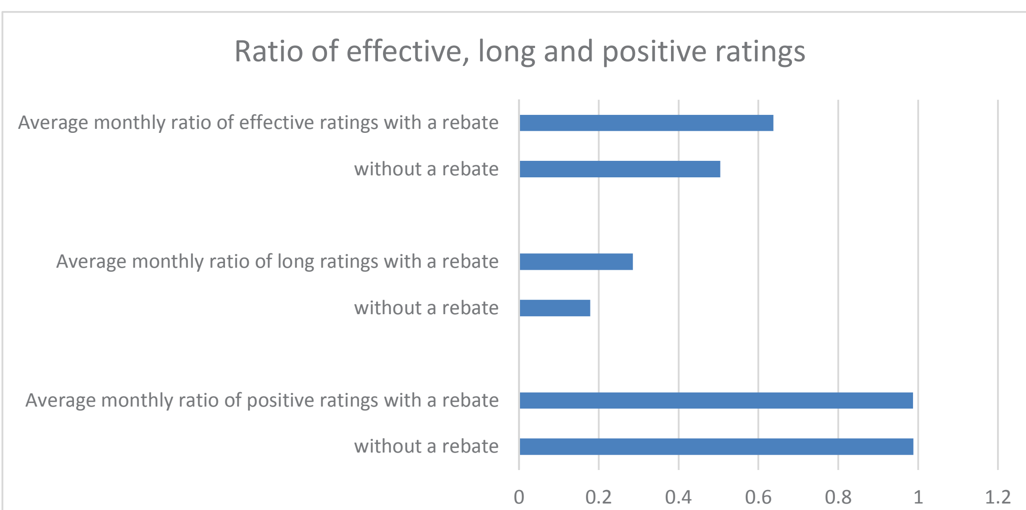


Figure 6: 95% confidence interval of estimates of change in main variables of interest