Essays on the Empirical Analysis in Seller Behavior in Online Marketplace

Haoyan Sun

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Ming Fan, Chair

Yong Tan

Abhishek Borah

Program Authorized to Offer Degree:

Business Administration

University of Washington

Abstract

Essays on the Empirical Analysis in Seller Behavior in Online Marketplace

Haoyan Sun

Chair of the Supervisory Committee:
Associate Professor Ming Fan
Department of Information System and Operations Management

In this dissertation, I examine problems that related to the phenomena of online marketplace. I specifically study how different forms of digital communications affect consumers and how online sellers can strategically adopt these tools to connect with them. In this dissertation, I investigate the effectiveness of various digital marketing strategies, such as sponsored search, social media endorsement, live chat, and buyer protection program, and how both the platform and marketplace sellers can leverage those tools to improve their businesses. The findings suggest that platforms should keep providing innovation tools to help online sellers to grows, it will not only be beneficial to the sellers, but also be critical to the growth of online platforms.

TABLE OF CONTENTS

Chapter 1: Introduction

In the dissertation I focus on the problems that online sellers facing and investigate what are the strategies that online sellers can utilize in order to help them grow their business. Online marketplaces are increasingly adopting innovative business models that can help sellers to better attract customers and engage customers. Chapter 2 to Chapter 4 have separately looked several tools that are available to the sellers, and how their adoption of the tools can further help the growth of both sellers and the platform.

The rest of the dissertation is structured as follows. In Chapter 2, I explore the two popular advertising tools, sponsored search and social media endorsement, in attracting web traffic from online sellers' perspective. I estimate an endogenous treatment model with latent variable to model online traffic given sellers' decisions on employing different advertising strategies. The evidences show that both sponsored search and social media endorsement are significant in attracting traffic, after controlling for seller self-selection behavior in choosing their strategies. Sponsored search has a higher impact on low-reputation sellers in bringing more traffic, and sellers with more return customers have higher benefit in using sponsored search in generating traffic. The findings are consistent with informative view of the advertising theory, and they provide interesting and import implications for both the platform and market participants.

In Chapter 3, I examine how online sellers' use of live chat influences their sales conversion probability. I argue that live chat can increase conversion probability by performing the functions of "informing" and "persuading". I further explore how the two roles of live chat interact with existing information displayed on product pages, such as product descriptions,

ratings and reviews, as those types of information also help consumers mitigate uncertainties regarding product quality as well as product fit to their needs. I apply a random coefficient model in a Bayesian hierarchical framework, and test the model using a panel dataset from Taboao. The results suggest that live chat has a positive impact on conversion probability, and the positive effect is stronger when product information on webpages is less comprehensive and when consumers perceive higher value of the seller products.

In chapter 4, I examine the behavior effect of both sellers and consumers after a seller joins the return guarantee program launched by the platform. The program is designed to help consumers establish trusts when they shop online. So I test the signaling effect of seller belonging to this program, and find that consumers do perceive this as a positive signal of seller's trustworthiness, which in turn affects their browsing and purchasing behavior. Also test the effect of participating the program on alleviating seller moral hazard problem and find it also reduces the possibility of moral hazard.

Chapter 5 provides a general conclusion of my dissertation.

Chapter 2: An Empirical Analysis of Seller Advertising Strategies in an Online Marketplace[1]

1. Introduction

In online e-commerce platforms, such as Amazon and eBay, sellers have to pay fees, usually a percentage of total transaction amount, to platform organizers in order to participate in the marketplace. Taobao, China's largest online marketplace owned by Alibaba Group, however, adopts a radically different business model by offering the basic platform service to sellers for free and generating revenues by monetizing traffic through advertising tools (Chen et al. 2016). Taobao's business model with little initial barriers to entry naturally attracts a large number of participants to the platform. As of March 2015, Taobao had 350 million active buyers and over 10 million active sellers, generating a combined gross merchandise value of US$394 billion (Alibaba Group 2015).

Although it is easy to become a seller at Taobao, surviving and prospering in this competitive market is difficult (Wong and Chu 2015). When hundreds of sellers are selling some similar products, attracting traffic to a store becomes critical. As a result, many sellers try to make their products more visible by investing in advertising tools (Osawa 2013), which in turn will benefit the platform. Indeed, Taobao's revenue mostly comes from advertising by sellers. According to a recent study, over 60% of Alibaba's revenue comes from its online advertisement (iResearch 2014). The two primary advertising tools at Taobao include sponsored search advertising and social media endorsement. According to one of Alibaba's IPO investment advisory firm, sponsored search contributed about 50-60% of Taobao's advertising revenue, while the rest is generated from seller advertising through different social media channels (China

[1] This paper is coauthored with Ming Fan and Yong Tan.

Renaissance 2013).

Taobao started sponsored search advertising service in 2007. Sellers have to bid keywords to have their product advertisement appear along with relevant search results on Taobao's website (Figure 1). Sellers only pay when a consumer actually clicks on the advertisement. Social media endorsement, another advertising tool offered by Taobao, was launched in 2009. It allows individual endorsers, who can receive commissions from the transactions, to post testimonials for different products so as to direct traffic to the seller website (Figure 2). Usually, sellers need to announce the products they want to promote at Taobao. Then, the endorsers can pick the products they want to endorse, and share the product information and user experience at a social media platform such as blogs and microblogs. Prior literature has examined sponsored search and social media endorsement separately in studying their relationships with firm performance. Empirical work on sponsored search has focused on the effectiveness of different types of search keywords (Blake et al. 2015), and sponsored search as a driver of click-through and conversion rates for an individual advertiser (Yang and Ghose 2010, Agarwal et al. 2015). Studies on paid social media endorsement are quite scant. The goal of this paper is to fill in the gap in the literature by studying the differential effects brought by sponsored search and social media endorsement jointly.

------ Insert Figure 1 and 2 about Here -----

Specifically, we investigate the effectiveness of sponsored search and social media endorsement at Taobao. Despite the popularity of the two advertising tools, their comparable effects remain unclear from an empirical standpoint. The mechanisms of the two advertising models are different. As sponsored search can target consumers who actively searching for similar products, social media endorsement can reach a wider group of consumer, who may or

may not be close to make a purchase decision. Prior studies have examined the impact of a specific marketing tool on firm performance as well as consumer response (e.g. Agarwal et al. 2015; Blake et al. 2015; Yang and Ghose 2010). Much less is known, however, on seller heterogeneity influences both the choices and the effectiveness of different advertising models in attracting traffic. Will a certain advertising model be more effective in boosting traffic for some sellers than others, given they vary in their reputation, clientele, and other dimensions? This inquiry is of great importance, as it can not only provide insights into the understanding on how these two advertising tools work, but also help sellers choose their optimal strategies and advise the platform with value-added services for its participants. We seek to explore the differential effects of the two advertising models. Specifically, we aim to answer the following questions: What type of sellers is more likely to adopt sponsored search and/or social media endorsement? Are those advertising strategies effective in increasing online traffic? Does the effectiveness of those strategies differ across different sellers, and how?

We find evidence that both sponsored search and social media endorsement are significant in attracting traffic, after controlling for seller self-selection behavior in choosing their strategies. Our empirical results, however, suggest that the two advertising mechanisms work quite differently, confirming the predictions based on the informative view of advertising theory (Nelson 1974). Sponsored search has a higher impact on lower reputation sellers in bringing more traffic. For high reputation sellers, the matching effects from sponsored search and from organic listing could be partially substitutive for each other, which leads to an over-exposure of those sellers and lower their marginal return from sponsored search. On the other hand, sellers with more return customers have higher benefit in using sponsored search in terms of generating traffic, which is consistent with the repeated business effect of advertising.

The rest of the paper is structured as follows. Section 2 reviews the relevant literature and discusses our research motivation. Section 3 develops the theoretical background for the empirical model. Section 4 introduces the dataset and illustrates the empirical model and the specification and identification strategies for the analysis. Section 5 provides results from empirical analysis and additional robustness check results. Section 6 discusses results with practical implications. Section 7 concludes the study with limitations and suggestions for future research.

2. Prior Literature and Research Motivation

Our study is related to the growing literature of sponsored search and social media endorsement in an online environment. There are quite a few studies that have looked at sponsored search (Athey and Ellison 2011; D. X. Chan et al. 2011; T. Y. Chan et al. 2011; Jeziorski and Segal 2015; Narayanan and Kalyanam 2015; Xu et al. 2012). Recent empirical work has examined sponsored search as a driver of click-through and conversion rates (Yang and Ghose 2010, Agarwal et al. 2015), and the effectiveness of different types of keywords (Blake et al. 2015). Specifically, Yang and Ghose (2010) used a retailer chain store's advertising data on Google and estimated the interrelationship between organic search listings and sponsored search listing. They found that the click-through rate on organic listings have a positive interdependence with click-throughs on paid ads. All the performance indicators in the presence of both paid and organic search listings are significantly higher than those in the absence of sponsored search advertisements. Agarwal et al. (2015) used data from an online retailer's keyword advertising campaign and examined how competing organic results affect the performance of sponsored search advertisement. They found that increased competition among organic listings leads to a decrease of click-through-rate on the sponsored search ads, but helps increase conversion rate on

those sponsored search ads. Blake et al. (2015) conducted a large-scale field experiment on eBay and found that the effectiveness of sponsored search advertisements differ for brand keyword search and non-brand keyword search.

Literature on social media endorsement has been scant. It is a new mechanism to promote products and services, and it is different from social advertising, in which ads are targeted based on underlying social networks and are highlighted when a friend interacts with the ads (Tucker 2016). Social media endorsement allows advertisers to post their promotion tasks on a platform and interested endorsers can take on the tasks to promote the products for monetary rewards (Peng et al. 2016). Two streams of research are related to social media endorsement. First, it is related to celebrity endorsement, which attaches the fame of a celebrity to a brand or product and influences consumer purchases. Celebrity endorsement has been shown to have positive effects on firm value (Agrawal and Kamakura 1995) and on book sales (Butler et al. 2005). Social media endorsement is different from celebrity endorsement in terms of the level of influences and the network size. The effectiveness of social media endorsement largely depends on the fan base and network connects of the endorser. Second, social media endorsement is also related to a growing body of literature that studied how word-of-mouth (WOM), especially online WOM, affect firm performance (Bickart and Schindler 2001; Clemons et al. 2006; Dellarocas et al. 2007; Duan et al. 2008; Godes and Mayzlin 2004; Gu et al. 2012; Liu 2006) and create social influence (Aral and Walker 2011, 2012; Fang et al. 2013; Iyengar et al. 2011; Moe and Schweidel 2012; Oestreicher-Singer and Sundararajan 2012; Susarla et al. 2012; Trusov et al. 2010). For example, Bickart and Schindler (2001) compared WOM generated from consumer and seller respectively, and found that consumer-generated information has a greater impact on consumer purchase decisions than seller-created information. Gu et al. (2012) compared internal WOM and external

WOM on high-involvement products, their results suggested that a retailer's internal WOM has a limited influence on sales, while external WOM sources have a significant impact. However, those studies on online WOM all focus on vonluntary contributions from consumers, while in our context, endorsors clearly have a profit motivation.

This paper also extends the literature that evaluates different drivers of web traffic. One of the key measures of the effectiveness of different marketing strategies is the amount of traffic that can be attracted since a retailer's financial performance closely relates to its ability to attract traffic (Perdikaki et al. 2011), and those traffic could potentially be converted into sales and revenue (Benbunan-Fich and Fich 2004). Luo and Zhang (2013) investigated the predictive power of web traffic and consumer buzz on firm value, and found that buzz and traffic explain a substantial portion of the total variance of firm value. Extant research has also investigated the drivers and impact of online traffic (Chatterjee et al. 2003; Dewan et al. 2002; Liaukonyte et al. 2015; Rutz et al. 2011; Wu et al. 2005). Among those, many researches focus on web traffic and clicks generated through advertising. For example, Chatterjee et al. (2003) examined the click proneness across consumers and studied the effects of exposure to repeated banner advertising and competing advertisements. Rutz et al. (2011) used search keyword-level data set and found indirect effect of sponsored search on future return traffic. We extend this body of literature and examine the impact on traffic brought by the two advertising strategies.

Moreover, our study is related to the growing literature that studies business models for online platforms and the mechanism of two-sided markets in engaging both buyers and sellers. Bhargava and Choudhary (2004) developed economic models of marketplace as intermediaries to examine how their pricing and product line design strategies could provide value added services to both sellers and buyers. Soh et al. (2006) compared 19 online marketplaces regarding

the strategy, price transparency, and performance, and they found that marketplace should pursue strategies of differentiation in order to succeed. They should provide compensatory benefits for sellers in the case of high price transparency and for buyers in the case of low price transparency. More recently, Chen et al. (2016) compared the two different revenue models adopted by eBay and Taobao and found the advertising model creates more values for buyers and makes sellers better off in most scenarios.

Our study contributes to the extant literature in the following ways. First, our study provides an empirical examination of two important advertising mechanisms on influencing traffic. Given the growing interests in both sponsored search and social media endorsement, empirical insights into the benefit of each strategy and their contingent effects on attracting traffic form an important contribution to theory and practice. Most literature have tended to examine the isolated impact of either advertising model, and overlook the interplay of those advertising strategies on traffic given sellers in different stages of decision making. Second, our study provides an empirical assessment of the consequences of the self-selection behavior in choosing advertising strategies by the sellers. Our results highlight the motivations for the sellers to employ different advertising strategies given their heterogeneous characteristics. Third, to the best of our knowledge, our research is one of the first studies to provide a rigorous quantification of the economic impact of online advertising for a two-sided market platform that relies on advertising as its main revenue source. Notably, by showing the effectiveness of sponsored search and social media endorsement, we could inspire other online marketplaces such as Amazon and eBay to offer similar advertising services to sellers, as these services can not only attract traffic for sellers, but also benefit the platform.

3. Theory

We first investigate the type of sellers who are more likely to adopt sponsored search and/or social media endorsement. Studies in information systems and economics suggest that online sellers are usually identified by their reputation scores (Ba and Pavlou 2002; Cabral and Hortaçsu 2010; Jin and Kato 2006; Resnick et al. 2006). In an online environment with unobserved quality and asymmetric information, seller reputation – shown in observable track performance, user feedbacks, and other indices – is especially important to demonstrate trustworthiness and credibility (Cabral and Hortaçsu 2010; Houser and Wooders 2006; Jin and Kato 2006; Livingston 2005; Resnick et al. 2006). Ye et al. (2014) examined a reputation system of an online marketplace and found that user feedbacks become so important in the eyes of sellers that low-quality sellers coerce buyers to revoke their negative feedbacks through retaliation (Bolton et al. 2012). Since reputation is such an important asset of sellers, a natural question is whether Taobao sellers with higher or lower reputation are more likely to use advertising strategies. From the informative view of the advertising theory (Esteban et al. 2001; Nelson 1974), one main objective of advertising for sellers is to inform buyers and match their products to the preferences of buyers. As low-reputation sellers are less likely to appear in the front of the organic search list, they have stronger needs to advertise. Nevertheless, for high-reputation sellers, they may still be motivated to advertise. In spite of the importance of reputation score, there are reasons to believe that reputation alone is not adequate to signal quality. For example, Bolton et al. (2004) utilized an experimental approach and found that although the feedback mechanism induces quite a substantial improvement in transaction efficiency, it still did not reach the levels as in a repeated business condition. In addition, although high-reputation sellers are more likely to appear in the front of the organic search list,

there is no guarantee it will absolutely happen as the ranking of the search list also depends on price and other factors that are proprietary to the platform. It is not straightforward, from theory, whether sellers with high or low reputations are more likely to advertise, and this becomes an interesting empirical question.

Sponsored search and social media endorsement are likely to have positive effects on boosting traffic for Taobao sellers. As noted earlier, advertising is attractive to firms as a means through which they can convey information to consumers and match products to the needs of buyers (Bagwell 2007). The informative view of advertising theory proposed by Nelson (1974) argues that advertising provide information to consumers through the following two effects: matching effect and repeated-business effect. First, matching effect means that advertising contains information on product attributes and price, which can match potential buyers' heterogeneous tastes and their reservation prices (Anderson and Renault 2006). Second, advertising has the repeated-business effect as it can rekindle memories of experiences with the advertised products (Nelson 1974, Hertzendorf 1993). Both sponsored search and social media endorsement can have the matching effect as well as the repeated-business effect. Therefore, we expect both sponsored search and social media endorsement can significantly contribute to the increase of traffic for the sellers.

In addition to the direct effect of sponsored search and social media endorsement in attracting traffic, we test two predictions that are associated with the matching effect and repeated-business effect based on advertising theory (Nelson 1974, Bagwell 2007). First, we understand, in addition to the matching effect from sponsored search, organic search list, which is generated by the platform's search engine based on consumer keyword input, also provides the same function. Thus, the effectiveness of sponsored search may depend on whether the

advertised products also appear in the organic list or not. High-reputation sellers are more likely to appear in the front of the organic search list generated by the platform's search engine. Such exposure for high-reputation sellers already contains information on horizontal attributes and product prices, and thus can partially substitute the effects of sponsored search. In other words, information contained in sponsored search could be redundant when high-reputation sellers already have advantageous exposure in the search result list due to their high reputation. Hence, their marginal benefit of using sponsored search could be reduced because of over-exposure from both paid list and organic list. Conversely, low-reputation sellers are often in a less advantageous position on the search list, and the favorable spot brought by sponsored search is unlikely to substitute the effect of organic search list, leading to a higher marginal benefit in attracting potential customers. Therefore, we expect that sponsored search has a more significant effect on traffic for sellers with lower reputation.

Second, we test the repeated-business effect for the two advertising mechanisms. Advertising may remind consumers of their previous experience with the product (Nelson 1974, Hertzendorf 1993), and the exposure of either sponsored search and endorsement from social media could be effective in attracting customers who already have interactions with the sellers (Blake et al. 2015). This argument is consistent with marketing literature on mere exposure of banner ads (Chatterjee et al. 2003; Drèze and Hussherr 2003), by which consumers tend to develop a preference for things merely because they are familiar with them. Sellers may induce the return customers by advertising and allow those buyers to go directly to the acquainted stores without further searching. Return customers have a much higher likelihood to click the sponsored search advertisements compared to new customers. Those sellers with higher probability of repeat business have the greatest payoff to improved consumer connection.

Therefore, the more return customers a seller has, the more likely sponsored search advertisements would be clicked on. Thus, we posit that sponsored search has a stronger positive effect on traffic for sellers with more return customers. Similarly, social media endorsement should be effective on return customers as those customers are already familiar with the products, and the remarks in social media such as blogs can remind the experience for the return customers.

4. Data and Econometrics Model

4.1. Data and Variables

We conduct our inquiry on the effectiveness of sponsored search and social media endorsement based on a panel data provided by Taobao. The dataset contains 4,348 randomly sampled sellers in women's fashion category, with monthly data across 10 months from May 2011 to February 2012. This is an unbalanced panel data as some sellers may enter the market during the period. On the next sub-section, we give detailed descriptions on the variables in the dataset. Table 1 provides the definitions of the main variables. Table 2 presents the summary statistics of the variables.

------ Insert Table 1 and 2 about Here ------

Traffic is the dependent variables for our analysis, and it is the number of unique individual visitors that have visited any product pages of a particular seller. Specifically, we use the monthly total of the visits in our analysis. *Traffic* is a count variable and is highly skewed as the standard deviation is much higher than its mean (distribution is shown in Figure 3).

------ Insert Figure 3 about Here ------

Sponsored search (sponsored) is a dummy variable indicating whether a particular seller pay for the service of sponsored search advertisement in a given month. If a seller adopts the

sponsored search, the variable is coded 1; otherwise, it is coded as 0. The average of the variable 0.56, suggesting over 50% of the sellers used the sponsored search.

Social media endorsement (social) is a dummy variable indicating whether a particular seller utilizes social media endorsement service from Taobao in a given month. If a seller adopts the social media endorsement service, the variable is coded 1; otherwise, it is coded 0. The average of the variable, as shown in Table 2, is 0.46, suggesting less than half of the sellers used the service in a given month during our study period.

Reputation is the accumulative feedback scores for a seller. It indicates how well and how long the sellers have established themselves on the marketplace. A buyer can rate the seller positive, neutral, and negative after each transaction. Sellers receive (+1) point if feedback is positive, (+0) point if feedback is neutral, and (-1) point if the feedback is negative. These scores accumulate over time for each seller and the reputation score is simply the cumulative sum of each transaction feedback. Taobao reputation score reveals to the customers how many successful transactions a seller has made and also implies how successful a seller is. Reputation score is also one of the major indices that help customers make purchase decisions. Previous literature suggests that the accumulated reputation score can be a proxy for seller size, as well as for seller trustworthiness (Jin and Kato 2006; Rob and Fishman 2005). Especially for those potential buyers who had never visited the online store before, they would rely heavily on the reputation score to make the purchase decision.

Return customer indicates the number of customers who have made purchases with a particular seller more than one time within past six months. The average number of return customer is 146.97, and the standard deviation is 802.29, higher than the mean, suggesting the distribution of the variable is skewed. Return customers are usually the source of organic traffic.

Average price captures the average price of all the products of a seller in a given month. Different sellers have different business strategy to establish themselves on the marketplace. Average price is a good measurement of sellers' product scope, indicating whether a seller is a high-end luxury brand seller or a budget brand seller.

Product variety indicates the breadth of product categories offered by a seller in a given month. Given different business model, sellers may want to focus on a limited categories of products or they would like to provide a large variety of products for customers to choose from. As shown in Table 2, the average number of product variety is 299.79, and the standard deviation is 719.81, suggesting sellers on the platform have very different business model in terms of satisfying diverse consumer needs.

Geographical region is sellers' physical locations that are shown on their online profiles. We categorize each specific location to the three major regions in China: East, West or Central. The official classification of the three regions in China appeared in the Seventh Five-Year Development Plan passed by the National Congress of China in 1986 (Wei 2013), and we follow the plan's classification to derive these dummy variables. These three economic regions represent different levels of economic growth in China. Eastern China is the most advanced in terms of average income, while western China trails in economic growth. The central region falls between the levels of Eastern and Western China.

The distributions for many of the above variables are highly skewed and we log transform the continuous independent variables in our empirical analysis.

4.2. Empirical Model

Our model is estimated using the endogenous treatment procedure developed by Rabe-Hesketh and Skrondal (2012). The technique is similar to the treatment effect approach

developed by Heckman (1979) but in a more generalized and flexible form. Our model includes two endogenous treatment variables, and we estimate the system equations simultaneously instead of following a two-stage estimation approach of sample selection and treatment model. In the following, we model the traffic equation and the selection equation of advertising strategies respectively.

4.2.1. Traffic Equation

We postulate that sellers' advertising strategies affect web traffic. The empirical model begins with the standard assumption that the sellers make choices about their strategies to increase store traffic, where store traffic is a function of advertising strategies employed and other seller characteristics. Specifically, we assess the impact on traffic based on seller price level, reputation, number of return customers, and product variety, all of which capture different aspects affecting online traffic. The number of unique visits y_{it} is a count variable and can be assumed to follow a Poisson distribution. Because we find evidence of y_{it} being over-dispersed, we consider a negative binomial model as the traffic model, which is a generalization of a Poisson regression model that allows for over-dispersion by incorporating an individual unobserved effect into the conditional mean (Hausman et al. 1984). We have the traffic model as follows:

$$p(y_{it}|X_{it}) = \frac{\Gamma(\alpha + y_{it})}{\Gamma(\alpha)\Gamma(y_{it}+1)}\left(\frac{\alpha}{\alpha+\lambda_{it}}\right)^{\alpha}\left(\frac{\lambda_{it}}{\alpha+\lambda_{it}}\right)^{y_{it}},$$

where α represents the extent of over dispersion. $\Gamma(\cdot)$ is a gamma function. Therefore, the negative binomial model is also a Poisson-gamma mixture model. The conditional expectation λ_{it} is structured as:

$$E(y_{it}|X_{it}) = \lambda_{it} = exp\left(y_{i,t-1}\tau_0 + sponsored_{it}\tau_1 + social_{it}\tau_2 + X_{it}\beta + \vartheta_i + \varepsilon_{it}^1\right).$$

The dependent variable y_{it} is the number of unique visitors seller *i* receives at time *t*. ε_{it}^{1} is the unobserved error and is assumed to follow a log-gamma distribution. $y_{i,t-1}$ is the traffic receipt from last period; $\vartheta_i \sim (0, \sigma_\vartheta)$ is the fixed effect capturing individual seller heterogeneity; *Sponsored* denotes whether a seller has adopted sponsored search; *social* denotes whether a seller has adopted social media endorsement; τ_1 and τ_2 captures the treatment effect of advertising strategy. X_{it} is a vector of variables including *reputation, average price, the number of return customer, and product variety.*

The challenge of analyzing the effectiveness of the strategies is that the choice of the advertising strategies could be endogenous, because the important determinants of seller using those strategies may be unobserved, and that these unobservable effects may be correlated with the random component of the traffic model. It is likely that there could be shared unobserved heterogeneity for adopting the strategies and attracting traffic. For example, mature sellers may be more likely to adopt the strategies with their financial capability, meanwhile they tend to attract more traffic than new sellers do. Our model accommodates the potential endogeneity of the self-selection through direct inclusion of latent variable π_{it} in both specifications. It is different from fixed effect panel model where individual specific unobservable depend on observed covariates in an unspecified way. Instead, we use a parametric approach, models out the dependence between individual specific effects and the covariates, leaving only the common unobservables. Below we establish the traffic model as

$$y_{it}{}^{*} = \log(\lambda_{it}) = y_{i,t-1}\tau_0 + sponsored_{it}\tau_1 + social_{it}\tau_2 + X_{it}\beta + \vartheta_i + \pi_{it}\delta + \varepsilon_{it}^{1},$$

$$\begin{bmatrix} \pi_{it} \\ \varepsilon_{it}^{1} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \sigma^2 \end{bmatrix}\right).$$

4.2.2. Selection Equation of Advertising Strategies

We model sellers' choices of each strategy that is driven by sellers' specific variables indicative of their inclinations to pursue advertising strategies. Since we have two binary choice variables, we therefore use bivariate probit model, which allows more than one equation, with correlated disturbances, in the same spirit as the seemingly unrelated regression model (Greene 2012). Equations for probability of selecting sponsored search and probability of selecting social media endorsement are simultaneously estimated. We model the strategy selections as a function of the following independent variables: *reputation*, *average price* and *product variety*. Although it is desirable to include all the variables that could possibly affect their selection of the strategy, we include those variables that we have information on and use the latent variable and error term to capture unobserved effects. Below we show the selection equation.

$z_{1it} = \omega\gamma_1 + \pi_{it} + \varepsilon_{i\square}^2,\ sponsored_{it} = 1\ if\ z_{1it} > 0, sponsored_{it} = 0\ otherwise\ ,$
$z_{2it} = \omega\gamma_2 + \pi_{it} + \varepsilon_{it}^3,\ social_{it} = 1\ if\ z_{2it} > 0,\ social_{it} = 0\ otherwise\ ,$

$$\begin{bmatrix} \varepsilon_{it}^2 \\ \varepsilon_{it}^3 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right), -1 < \rho < 1.$$

Individual observations on $sponsored_{it}$ and $social_{it}$ are available for all i and t; ω is a vector of variables including reputation, average price, product variety, and geographic region. $\varepsilon_{it} \sim N(0,1)$ is the error term. $\pi_{it} \sim N(0, \psi)$ is a factor representing shared unobserved heterogeneity, and δ is a factor loading. For ease of computation, we restrict $\psi = 1$.

4.2.3. Full model

Putting the above equations together, we have the full model as follows:

$$y_{it}^{*} = y_{i,t-1}\tau_0 + sponsored_{it}\tau_1 + social_{it}\tau_2 + X_{it}\beta + \vartheta_i + \pi_{it}\delta + \varepsilon_{it}^1\ ,$$

$$z_{1it} = \omega\gamma_1 + \pi_{it} + \varepsilon_{it}^2,\ sponsored_{it} = 1\ if\ z_{1it} > 0, sponsored_{it} = 0\ otherwise\ ,$$
$$z_{2it} = \omega\gamma_2 + \pi_{it} + \varepsilon_{it}^3,\ social_{it} = 1\ if\ z_{2it} > 0, social_{it} = 0\ otherwise\ .$$

Where, we assume that the latent variable and all the error terms follow a multivariate normal distribution:

$$\begin{pmatrix} \pi_{it} \\ \varepsilon_{it}^{1} \\ \varepsilon_{it}^{2} \\ \varepsilon_{it}^{3} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sigma^{2} & 0 & 0 \\ 0 & 0 & 1 & \rho \\ 0 & 0 & \rho & 1 \end{bmatrix} \right).$$

Our full model framework is adopted from Rabe-Hesketh and Skrondal (2012). We use a maximum-likelihood approach to maximize the log-likelihood function conditional on the sum $\sum_{t=1}^{T} y_{it}$. This conditional likelihood function does not depend on the unobserved π_{it}, as it is transformed out. Hence the estimator is consistent for the coefficients on the time-varying covariates and it is asymptotically normal.

5. Empirical Analysis

5.1. Results

Table 3 presents the results of the selection model of advertising strategy, in which we investigate whether seller characteristics are related to their choices of sponsored search and/or social media endorsement. We find that seller reputation has positive, significant effects on the choice of sponsored search with estimated coefficient of 1.111 ($p < 0.01$). The coefficient estimation of reputation for social media endorsement is 0.618, and is highly significant ($p < 0.01$). This finding indicates that sellers with higher reputation are more likely to adopt advertising strategies than those with lower reputation.

------ Insert Table 3 about Here ------

In addition, we show the boxplot distribution of seller reputation over the categories of using either one or both strategies (Figure 4). We see a steady increase of mean reputation score from sellers, ranging from those who do not use any of the two tools to those who use both. This further confirms the tendency of adopting advertising strategies when seller reputation increases. Moreover, sellers' choices of strategies also vary according to the price level, product variety, and geographic regions. Product variety has a negative effect on the likelihood of using

sponsored search, with coefficient estimate of -0.298 ($p < .01$), but has no significant effect on social media endorsement. It shows that sellers with less product variety are more likely to use sponsored search, but not necessarily to use social media endorsement. Average price has a positive, significant effect on sponsored search, with a coefficient estimate of 0.489 ($p < .01$), but has no significant effect on using social media endorsement, suggesting that sellers with higher price level are more likely to use sponsored search. In sum, only reputation plays a role in sellers' choices of social media endorsement, while sellers' decisions of using sponsored search are affected by reputation, product variety and average price. Sellers located in the East region are relatively more likely to use sponsored search, while sellers located in the both the East and the West region are more likely to use social media endorsement.

------ Insert Figure 4 about Here ------

Table 4 shows the results of the traffic model. In the base model, after we controlling for other factors, estimated coefficient for sponsored search is 1.378 and it is highly significant ($p < 0.01$). The estimated coefficient of social media endorsement is 1.058 and it is highly significant as well ($p < 0.01$). Thus, sponsored search and social media endorsement are both effective strategies in boosting online traffic. Figure 5 also shows the distribution of traffic according to the usage level of the strategies. Sellers using social media endorsement have lower mean traffic than sellers using sponsored search, which is consistent with the results of the traffic model, suggesting that sponsored search has a higher marginal effect on traffic than social media endorsement. We also find that seller reputation, product variety and the number of return customer are all significant in increasing traffic. The estimated coefficient for reputation is 0.049 and is highly significant ($p < 0.01$). The estimated coefficient for product variety is 0.191, and is

highly significant ($p < 0.01$). The estimated coefficient for return customer is 0.533 and is significant as well ($p < 0.01$).

------- Insert Table 4 about Here ------

Interaction Model (I) (Table 4) shows our investigation of whether the effects of sponsored search and social media endorsement are complementary or substitutionary. The estimated coefficient for the interaction term of sponsored search and social media endorsement is -0.399, and it is highly significant ($p < 0.01$). It indicates that sponsored search has a smaller marginal effect on store traffic when it is used together with social media endorsement. By the same token, social media endorsement has a smaller marginal effect on boosting store traffic when coupled with sponsored search. Therefore, the effects of these two advertising strategies are partially substitutive.

The interaction model (II) in Table 4 shows that sponsored search is a more effective tool in boosting traffic for sellers with lower reputation, with the coefficient estimation of -0.286 ($p < 0.01$) for the interaction term of sponsored search and reputation. Social media endorsement appears less significant when interact with seller reputation. The estimated coefficient for the interaction term of social media and reputation is 0.024, and it is only marginally significant ($p < 0.1$). This finding suggests that, compared to high-reputation sellers, sellers with lower reputation may have more benefit in using sponsored search, yet high reputation could somewhat bring marginal benefit when sellers adopt social media endorsement. Lastly, sponsored search works better in increasing traffic for those sellers with more return customers. The estimated coefficient for the interaction term of sponsored search and return customer is 0.123, and the coefficient is highly significant ($p < 0.01$). The estimated coefficient of the interactive term between social media endorsement and return customer is -0.011, but it is insignificant.

5.2. Robustness Test

Our dataset contains sellers that adopted both strategies, therefore allows us to compare the effectiveness of the two strategies. However, some may argue that our results may be influenced by the correlation of the two strategies or the selection bias of those sellers that adopted both strategies. Hence, we test our empirical model based on a split sample, where we only include sellers who have had adopt only one strategy or none in the whole sample period as a robustness check to ensure that there is no such bias existed in our results. The results for the split sample analysis are qualitatively similar to the results of entire sample, where all the signs, magnitude and significance level remain the same.

To further confirm the main effect of either strategy on increasing traffic, we also conduct matching techniques as robustness checks. We match sellers who resemble each other in all relevant characteristics and then compare the impact on traffic by using each strategy. Since we have two binary choice variables, we have to first match sellers that not only have similar characteristics but also have the same decision on adopting one of the strategies, then we treat the adoption of the other strategy as a treatment. We use log-transformed traffic variable as outcome variable and calculate the after treatment effect for each strategy on the outcome. Given our treatment could possibly change for every time period, we have to do the matching sequentially and calculate the overall after treatment effect (Wooldridge 2010). Although there are several matching techniques, we use propensity score matching and nearest neighbor matching. For propensity score matching, we have after-treatment-effect of sponsored search equals to 0.44, and after-treatment-effect of social media endorsement equals to 0.24. For nearest neighbor matching, we have after-treatment-effect of sponsored search equals to 0.64, and after-

treatment-effect of social media endorsement equals to 0.28. Both matching technique confirms the effectiveness of both strategies.

In additional to the original variables that are used in the main analysis, we check the validity of our results by using an alternative variable of seller reputation: seller tenure. Although seller tenure is usually correlated with seller reputation, whereas seasoned sellers usually have higher reputation, it may not function the same as seller reputation does. This way, we can eliminate the possibility that the impact of those strategies on traffic are caused by seller reputation. We find that the results pertaining to the effect of the strategies using this alternative variable is substantively similar to the main result: seasoned sellers are more likely to adopt both strategies and the effect of both strategies are still significant on increasing traffic after controlling for seller tenure.

6. Discussions and Implications

The evidence we find from the selection model of advertising strategies suggests that sellers with different reputation level have different preferences in choosing their advertising strategies. We find that sellers with higher reputation are more likely to adopt sponsored search. As we mentioned earlier, high-reputation sellers are the ones that have already established themselves on the platform, but they may still be motivated to employ advertising tools. Although high-reputation sellers are more likely to appear in the front of the organic search list, there is no guarantee it will absolutely happen as the ranking of the search list also depends on price and other factors. In addition, sellers with higher reputation may have more resources to spend on advertising, as sponsored search could be expensive. From the perspective of persuasive role of advertising (Bain 1956; Braithwaite 1928; Comanor and Wilson 1974), it is

also possible that high reputation sellers may be motivated to advertise for competitive reasons. However, we do not have data on market structure and it could be a topic for future research.

From our baseline traffic model, we confirm that both sponsored search and social media endorsement play significant roles in increasing traffic. We also find that the two strategies, sponsored search and social media endorsement, are partially substitutive. Although sellers who employ both sponsored search and social media endorsement are able to, on average, attract more traffic than those who only use one tool, the marginal effect of sponsored search (social media) is lower when social media (sponsored search), the other tool, is also used. It is possible that some potential buyers are exposed to both advertising mechanisms, which lowers the marginal effect of each tool.

Our results provide interesting insights on the ways the two advertising models work. Although sellers with higher reputation are more likely to adopt sponsored search, the effect of sponsored search is stronger for sellers with lower reputation. Low reputation sellers are often positioned in a less advantageous position in organic search, and, as a result, buyers cannot find them easily. By adopting sponsored search, lower-reputation sellers have higher chance to be matched with potential customers, and hence, the marginal benefit of sponsored search is higher for them. In addition, high-reputation sellers are more likely to appear in the front end of product search, and, if in the meanwhile their sponsored search advertisements are displayed alongside the organic list, it can create an over-exposure for those sellers and the effect of attracting customer becomes somewhat redundant. In contrast, social media endorsement works differently in reaching the potential customers. Customers who see the advertisement in social media platform usually do not have a specific purchase intention. They may follow an endorser's link and visit a seller's product page when the endorser's post contains certain positive elements that

intrigue them. In this case, a high-reputation score may play a positive role in converting social media exposure into traffic, especially when the seller's reputation is also mentioned in the social media post.

The findings of this study suggest that return customers have a much higher likelihood to click the sponsored search ads compared to new customers, given the repeated exposure. Therefore, the traffic generating effect brought by sponsored search is stronger for sellers with higher number of return customer. However, we find this repeated exposure effect is not significant for social media endorsement. One explanation is that return customers are more likely to be exposed to sponsored search ads because these are targeted ads based on the product keywords. Return customers, because of their prior purchase records and preferences, could be exposed to sponsored search ads more frequently. For social media endorsement, despite its potential positive effect on repeated customers, the likelihood for return customers to see a social media post for a related store or product is low, given that the audience of the social media endorsement is quite general and may have little overlap with repeated customers. Thus, we do not find social media endorsement has significant effect on return customers.

Our study provides insightful managerial implications to practice. In online marketplace, traffic is the one critical factor for sellers to succeed. It constitutes the mass of visitors that can potentially be converted into sales (Benbunan-Fich and Fich 2004). Indirectly, continuous web traffic could help sellers to gather more feedback, gain more return customers, and sustain long term growth. Given the importance of traffic, sellers should be aware of the value in adopting advertising tools. Our results have important implications to both new sellers with low reputation and seasoned sellers. Our results suggest that sellers' self-selection of advertising strategies may not be as effective as they think. Although low-reputation sellers have higher marginal benefit

from sponsored search, they, ironically, are less likely to adopt sponsored search. This could be due to the high cost of search advertising. Sellers with lower reputation and less experience, may lack the strategic understanding as well as the resources for search advertising. Low-reputation sellers may consider increasing their investment on sponsored search for the following reasons. First, as out study finds out, the marginal benefit of their investment is higher than that for high-reputation sellers. Second, bringing more traffic has some long-term, strategic benefits. With more traffic, a seller's reputation and return customers with increase more quickly, and this can create a virtuous cycle for the sellers.

For high-reputation sellers, while the value of sponsored search is more for low-reputation sellers, they may still want to use sponsored search for the following reasons. First, the benefit in using sponsored search is still positive. Second, they are the ones with possibly more return customers, and sponsored search could be an effective way to target those customers and generate more repeated sales. Further, they may keep using sponsored search for competitive reasons. In addition, high-reputation sellers could consider spending resources on social media endorsement, which might attract new customers from a different channel.

From the marketplace perspective, our results confirm the value of the advertising-based business model and provide three fold implications to the platform organizers. To begin with, sellers' preferences for different advertising strategies allow the marketplace to understand the heterogeneous sellers better and to customize services that would benefit different sellers. As new sellers benefit from the freemium model to start their businesses, the platform should educate the sellers to utilize effective tools to gain competitive advantage. For new sellers, the platform should raise their awareness of various marketing tools and coach them to use those tools skillfully. Essentially, it is not an optimal strategy for sellers to only use the free service.

Adopting tools strategically would create long-term benefit for both the sellers as well as for the platform as advertising can create value for both of them. Further, our results can help platform organizers to design a good policy for the marketplace. The long-term success of the marketplace depends on new entrants, who bring in innovative products and help extend the reach of the market. However, those newcomers may lack the resources and knowledge to compete with incumbents. Platform should have policies that introduce various marketing tools to entrants with discounts or rebates. Finally, the platform should design and implement alternative mechanisms to engage sellers. Even though we only discuss the two popular advertising tools, there are alternative ways to satisfy the needs of different types of sellers and avoid direct competition. Prior literature has studied mechanisms such as charity-giving (Elfenbein et al. 2012) and warranty program (Dewally and Ederington 2006), all of which could also be leveraged by the platform to reduce direct competition among the participants and maintain a healthy ecosystem.

7. Concluding Remarks

Using data from Taobao marketplace, we explore what affect sellers' choices of advertising strategy and how those choices affect online traffic for heterogeneous sellers. We find that high-reputation sellers are more likely to adopt advertising strategies compare to low-reputation sellers. We also find that sponsored search is more effective for sellers that have lower reputation and more return customers. Compared to sponsored search, social media endorsement is more general and has similar effect for either new sellers or mature sellers in terms of attracting traffic. We also empirically identify that the two advertising strategies, social media endorsement and sponsored search, are partially substitutionary, and, hence, sellers should maximize their benefit from choosing the right strategy according to their heterogeneous characteristics. Given all those findings, our results produce good managerial insights for both

online marketplace and marketplace sellers. First, our findings provide a good foundation for marketplace to target their service and allow them to design better policy to accommodate the sellers. Our findings indicate that the adoption of those advertising strategies may not be uniformly beneficial to all sellers. Therefore, the marketplace has to provide sellers with the most effective tool, which will benefit both the platform and the sellers. Second, sellers can benefit from choosing the tools strategically in order to take fully advantage of the paid services. This insight serves to help sellers to come up with better planning and resource allocation in future online advertising campaign. Lastly, our paper provides empirical evidence to support the informative view of advertising theory. Our results show that the matching effect and repeat business effect are important in online advertising.

We believe there are several possible extensions of this study. Our sample is limited to one product category, i.e. women's fashion, in the marketplace, in which consumer and seller behavior might be different when it comes to the sales of other product type. Therefore, checking for generalizability of results across different marketplaces or product categories would be interesting for future research. Future studies could also test other perspectives of the advertising theory and provide explanations for sellers' motivations in adopting advertising tools. Further, it will be interesting to see whether advertising models can complement other business models such as transaction-fee based ones.

Chapter 3: Assessing the Effectiveness of Live Chat on Conversion Probability: Evidence from an Online Marketplace[2]

1. Introduction

According to a study by Morgan Stanley, one important factor contributed to the success of Taobao, the biggest e-commerce platform in China, is its live chat system, WangWang, which facilitates buyer-seller communications (MorganStanley Ltd 2005). Over the years, many online retailers, including Nordstrom.com, Kohl's, and Walmart.com, have also come to embrace live chat for the purpose of improving customer service and facilitating transactions on their websites. A survey by J. D. Power in 2013 indicates that live chat has become the leading contact source for the online channel, as 42% of customers use live chat in comparison to 23% use email and 16% use other social media. More importantly, online satisfaction is the highest among customers who use the chat feature (J.D.Power 2013).

Consumers usually start their online shopping by browsing and comparing different products. They typically initiate live chat requests in the middle of the online shopping when they want help from a live person to answer specific questions about product features or delivery (Clarkson 2010). Similar to brick-and-mortar stores, the two fundamental tasks of an online seller are: (i) attracting traffic to the store, and (ii) converting the traffic into purchases (Perdikaki et al. 2011). Consumers are likely to abandon product pages or shopping carts if their burning questions remain unanswered (Clarkson 2010), and live chat can certainly facilitate the second task and improve sales. In this paper, we seek to establish that live chat can substantially influence consumer purchase decision and increase conversion probability.

[2] This paper is coauthored with Ming Fan and Jianqing Chen.

We argue that the live chat in an online marketplace plays two roles in connecting sellers and consumers. The first role is to inform consumers about specific product and service attributes. Most of the consumers start live chat requests when they are interested in a certain product but have questions about the product or transaction. Therefore, live chat service allows a seller to engage visitors with purchase intent and address their burning questions in making purchasing decisions. Second, live chat can guide customers by making suggestions and highlighting the values of the products when the seller and buyer have real-time conversations. Just like in a physical store, where a sales representative can add value by helping customers select the right product, in an online store, the seller can better understand customer needs and reach out to the customers by making recommendations and suggestions. Understanding the relative importance of these two roles can help online seller allocate their resources and serve their customers more efficiently. In the meantime, the effectiveness of the two roles of live chat could depend on the existing information on product pages, such as product descriptions, ratings and reviews, as those types of information also help consumers mitigate uncertainties regarding product quality as well as product fit to their needs (Chen and Xie 2008; Kwark et al. 2013). Our research seeks to address the following interrelated questions: Is there a heterogeneity among different types of sellers in offering live chat service? How does a seller's effort on live chat affect the seller's traffic conversion? How does the effectiveness of live chat depend on existing website information?

Despite the increasing recognition that live chat is critical in facilitating online transactions, there is a lack of empirical work examining such effect. Recently, Ou et al. (2014) find that buyers' effective use of online communication tools cultivates swift guanxi and trust by enhancing the buyers' perceptions of interactivity. As a result, these buyers are more likely to

make repeat purchases. Nonetheless, as Ou et al. (2014) focused on whether online communications would increase repurchases, the sample of their survey study was limited to those who already have made a purchase with the seller, and the effect of live chat on converting store traffic to transactions has not been addressed. In addition, we examine the effect of live chat based on second-hand, objective data from sellers, and thus offer a direct approach in investigating the effect of live chat.

Our empirical findings show evidence of live chat effort in driving online conversion rate. However, the effects of live chat are different given the existing information perceived by consumers. The informative effect of live chat could complement information loss on the product page, while the persuasive effect of live chat is stronger under the situations when consumers give preferences toward the seller/product.

The rest of the paper is structured as follows. Section 2 reviews relevant literature and highlights our theoretical contribution. Section 3 provides descriptions of the research context and develops the theoretical background for the empirical model. Section 4 introduces the dataset and variables. Section 4 illustrates the empirical model and specifications. Section 5 provides results from empirical analysis. Section 6 concludes with a discussion of practical implications, contributions, and limitations.

2. Literature Review

The need for online shopping help is evident from prior studies (Zhu et al. 2009), and a great amount of prior research is formed in the fields of computer-mediated communication (CMC). CMC includes many different types of technologies such as email, online communities, chat and so on. IS literature has focused on the social structures that emerge when people use these applications, as well as the design and implementation issues associated with constructing

communication technologies. Many studies have focused on the interface of online consumer trust and communication technologies. Recently, the effect of CMC technologies has been studied empirically in the context of online communities, such as crowdsourcing platforms (e.g., Blohm et al. 2016; Burtch et al. 2016), communities for user-generated content (e.g., Johnson et al. 2015; Zeng and Wei 2013), and online feedback systems. However, research on the effective use of live chat service in facilitating sales performance in C2C marketplace is still limited. Yadav and Varadarajan (2005) present a conceptualized model of interactivity from a marketplace perspective, and argue that interactivity is what perceived valuable by consumers in the online marketplace. Ou et al. (2014) adopted media synchronicity theory and apply that in the context of online marketplace, they found that CMC technologies can help building swift guanxi between buyer and seller, so that they can facilitate repeat purchasing behavior.

Our work also relates to a large body of literature that examined marketing communication mix. With the rapid growth of Internet, the elements of marketing communication mix have also evolved. live chat is deployed as one new element in the marketing communications mix (Bucklin et al. 2009; Chen and Xie 2008). Numerous research have examined the impact of marketing communication mix implemented outside the website (e.g., display ads, e-mails) in enabling conversions at the website (Chan et al. 2011; Chatterjee et al. 2003; Ghose and Yang 2009; Manchanda et al. 2006; Rutz and Bucklin 2011). In particular, Li and Kannan (2014) integrate the effects of a variety of marketing communications (e.g., search, display ads, e-mails, referral engines) on website visits and conversion, they found that the attribution of each channel to the overall conversion rate is different given different brand strength of the firms. Their work explains the nature of carryover and spillover effects across various online marketing mix elements, yet their data doesn't contain the live chat channel. To

the best of our knowledge, very few studies have examined why and how live chat is effective as an important form of marketing communication mix.

We adopted our theory from economics literature and examined the two roles of live chat. The first is the persuasive role, wherein advertising alters consumer's tastes and creates spurious product differentiation and brand loyalty (Bagwell 2007). The second is the informative role, wherein advertising reduces search costs that otherwise may deter a consumer from learning of each product's existence, price and quality (Bagwell 2007). Both of the two roles have been widely discussed in the context of advertising (e.g. Braithwaite 1928; Hertzendorf 1993; Nelson 1974; Robinson 1969). We contribute to this body of research by demonstrating how live chat could also fulfill the two roles, and we show that it is of special importance to subject the two roles to the empirical evaluation of live chat.

Our study contributes to the growing literature that looks at different driving forces of conversion rate. A number of papers focus on using clickstream data to analyze the conversion effects of website visits (Moe and Fader 2004; Sismeiro and Bucklin 2004) and advertisement clicks (Agarwal et al. 2011; Manchanda et al. 2006; Xu et al. 2014). For example, Sismeiro and Bucklin (2004) used clickstream data to develop a model of online buying by linking the purchae decision to what visitors do and to what they are exposed to while at the shopping site. Their finding suggests that the more time and effort that visitors invest in the site, the more likely they are to eventually buy at the site. They also find that a site's offering of a electronic decision aid (such as comparison matrix tool) does not increase conversion rate. Xu et al. (2014) studied the effects of various types of online advertisements on purchase conversion. Their results show that display ads have a relatively low direct effect on conversion rate but they are likely to stimulate subsequent clicks on other advertisements, which then lead to purchases.

3. Research Context and Theoretical Framework

3.1 Research Context

To address our research questions, we study seller behaviors on Taobao.com, one of the largest online marketplace in the world. In 2004, Taobao launched its live chat service, "WangWang," as a tool to facilitate communication between buyers and sellers. Although it is functionally similar to standalone instant messaging applications such as Skype and MSN, WangWang is imbedded in Taobao website and its icon is displayed on product description pages. When a potential buyer logs in at the marketplace, WangWang is automatically activated and the buyer can see whether a seller is online, indicated by the color of the icon. If a seller is online, the buyer can chat with the seller. WangWang not only supports live chat, but also allows file transfers. As such, sellers can send additional product images or related certificates to interested buyers upon request. Buyers can use the tool to ask a wide range of questions, such as product features, shipping, and return policies.

Industrial reports have argued that live chat service, which is absent in eBay, is an important advantage that makes Taobao more competitive against eBay in China (MorganStanley 2005, Ou and Davison 2009). WangWang improves customer services by facilitating real-time communication between buyers and sellers and also fulfill the needs of many Chinese consumers with haggling habit. However, sellers' dedication to use the tool varies broadly---some sellers reply to buyers' questions promptly and make real-time conversation with buyers most of time, while others do not. The data on the usage of live chat among sellers provides us a unique opportunity to investigate how sellers' responsiveness in live chat influences buyers' conversion probabilities.

3.2 Theoretical Framework

Product Quality and Fit

In making online shopping decisions, consumers have to consider quite a few factors such as brand, product description, price, availability, and shipping. All these factors can be broadly categorized to two dimensions: product quality and fit (Kwark et al. 2014). Product quality refers to the degree of excellence of the product, and consumers prefer high quality rather than low quality, everything else being equal. In the case of women's apparel, how well an apparel has been made and packaged is about the quality. Although quality usually cannot be measured directly, in an online environment, product ratings and reviews can serve as an indicator of the quality and, thus, could affect consumers' purchase decisions.

The *fit* of a product refers to the degree of being in agreement with a consumer's need. Fit is consumer specific, and consumers may have different preferences for the same attribute (Kwark et al 2014, Mendelson and Parlakturk 2008). For instance, for women's apparel, buyers have different preferences for its color and size. On a shopping website, sellers typically provide fit information through detailed product description in text and pictures. In the case of women's apparel, sellers on Taobao typically provide information such as designer, product color, size, and fabric, which is all about the fit. Besides product description, online product review could also provide information regards to the fit. For example, some reviews contain the buying consumer's height and weight and whether the size fits his/her figure after trying it on.

Although information such as product ratings, reviews, pictures, and detailed product descriptions help customers to know different aspects of a product, there are still tremendous uncertainties regarding product quality and fit, which cannot be easily addressed. For example, a certain style of clothing may look good or not that good with other pieces. Those issues require the help from experienced sales associates who understand both the products and customers.

Role of Live Chat

Consumers usually start live chat requests when they have already been attracted to an online store and are browsing various products. If they cannot quickly find answers to their questions, customers are likely to abandon their browsing and shopping carts (Clarkson 2010). Live chat service allows a seller to engage visitors with purchase intent and address their critical questions regarding their purchasing decisions. The use of live chat can not only provide additional information to consumers to reduce product quality and fit uncertainty, but also influence consumer preferences by introducing and promoting certain products. As an important element of the expanded Internet-based marketing communication mix (Bucklin et al. 2009), live chat, just like service representatives in brick-and-mortar stores, can ensure immediate customer service and effectively convert interested browsers into actual buyers. We aim to investigate the role of live chat in improving a store's conversion rate by looking at two major tasks of marketing communication that live chat can deliver: *informative effect* and *persuasive effect* (Keller 1987, 2001).

First, live chat allows sellers to provide customers with timely information, which will reduce overall uncertainties and help customers better assess a product's fit to their needs. This is the *informative effect* of live chat. Compared to shopping in physical stores, online shopping is associated with more uncertainties because customers cannot touch and feel the products (Devaraj et al. 2002; Dimoka et al. 2012; Kim and Krishnan 2015; Zhu and Zhang 2010). When customers are browsing through products across different sellers, they rely on the pictures, product descriptions, and product reviews to make purchase decisions. Nonetheless, customers may often have questions that are not answered by the information on product pages, or the answers may not be easily located in product descriptions and reviews. All those uncertainties

may either impede consumers who would have a good fit with the product from purchasing, or impose search costs on consumers and deter a consumer from learning of the product. When a seller provides information through live chat, the seller can address a buyer's questions directly and can reduce a buyer's uncertainty regarding the product's fit to her/his need. When uncertainties are reduced and consumers possess more product-related information, they will be more likely to purchase a product that fits their needs (Goh, Heng, and Lin 2013). In the case of women's fashion, customers may ask for detailed information on fabrics, size, and color. Such questions, if not answered, could hold customers back from making the purchase. With additional information obtained through live chat, uncertainties can be alleviated. As a result, a potential buyer is more likely to find a product that matches her needs, and has more desire to complete the transaction.

Second, live chat provides sellers an opportunity to engage conservations with customers and "lobby" customers by highlighting the positivity of the products, and thus a customer's perceive quality or valuation of a product might be enhanced, which we name *persuasive effect* of live chat. In contrast to the informative effect of live chat via which buyers learn more about a product's fit to their needs, persuasive effect may occur via sellers' elaboration on the quality or value of the product. For instance, via WangWang sellers on Taobao often quote positive testimonies from buyers to illustrate the great value of a product, highlight recent orders and sales figures to indicate the popularity of a product, elaborate on how well a product has been made or how durable a product is, and reiterate the good rating as a proof of the good quality of a product. All these types of positive statements via live chats can influence consumer evaluation of the seller and the product, and make consumers more likely to make up their minds and make purchase.

Combining the functions of live chat through the informative effect and persuasive effect, we have reasons to argue that sellers with higher live chat responsive rates can better address the customers' needs, enhance customer experience, and reduce the abandonment of product and sales pages. Thus, we hypothesize:

Hypothesis 1: A higher live chat responsive rate is associated with a higher sales conversion rate for sellers.

In addition, we aim to examine how the two functions of live chat interact with the existing information dimensions so that we can further understand this communication mechanism. We argue that the two functions of live chat—informing and persuading—could have differential effects as a result of how consumers perceive product fit and value from the existing information on the product page. While the informative role of live chat—in terms of providing additional desired information—could be a natural supplement to the fit information consumers are seeking, the persuasive role may largely depend on the perceived value of a customer toward a product.

When sellers have already provided detailed information such as product detail-views and size charts, less demand is desired for this type of information via live chat. Therefore, the relationship between live chat and purchase conversion is contingent on the comprehensiveness of product information. When detailed information has already been provided, live chat is of less value in terms of addressing customers' specific concerns and questions regarding product fit. In other words, live chat plays a more important role in converting potential buyers when the product description on the item page is less comprehensive, or live chat is likely to be less useful in the presence of already abundant detailed product information. Therefore, we expect that the

positive effect of live chat on conversion is stronger when sellers have less detailed product information on product pages, and we have the following hypothesis:

Hypothesis 2: Live chat responsiveness has a larger impact on conversion rate when sellers have less detailed product information.

persuasion is more likely to succeed when a potential customer already has a positive impression or feeling toward the product (Ahluwalia 2002); that is, when the customer perceives higher utility from a product, persuading is more likely to work. Researches in social psychology have widely examined the conditions that can make persuasion more likely to succeed, and it is well established that individuals are more likely to be persuaded by someone they like (Cialdini 2001, Cialdini and Goldstein 2004). Prior research indicates that consumers' expected utility of an online product is higher when the product has a higher product rating and/or a lower price (Kwark et al. 1014). Product rating can reflect product quality and has been shown to be a critical factor in consumers' purchase decision (Chevalier and Mayzlin 2006; Clemons et al. 2006; Dellarocas et al. 2007; Dewan and Ramaprasad 2012). Additionally, price also plays an important role in shaping customers' perceived utility of a product. For example, Kim and Krishnan (2015) find that consumers are unlikely to buy expensive products online especially with the presence of product uncertainty. When consumers already have a high perceived value toward a product or a store, it is much easier for the seller to use, for example, positive testimonies or scarcity of the products to promote the products. In summary, we expect that live chat can be more effective in persuading customers and has a more positive effect on conversion when customers perceive higher utility from the product—that is, when either the product rating is higher or when the product price is lower. Therefore, we hypothesize:

Hypothesis 3: Live chat responsiveness tends to have a larger impact on conversion rate when the product rating is higher.

Hypothesis 4: Live chat responsiveness tends to have a larger impact on conversion rate when the product price is lower.

We further investigate with the role of live chat on conversion with the presence of online reviews. Previoius research have found that customer reviews are the most effective user generated social tactic for driving sales(Dhar and Ghose 2010). However, there is a significant amount of noisy information in these kind of social content, and it might have different interaction effect with live chat. On one hand, While online reviews provide plenty of information to consumers, the volume of reviews also signals the popularity of a product (Zhu and Zhang 2010) and is important in driving sales (Duan et al. 2008; Godes and Mayzlin 2004). Reviewers may reveal information in the fit dimension which may work as substitutive to the live chat service provided by sellers. On the other hand, quality signals through volume of reviews can complement the persuasive effect of live chat. If the quality dimension dominates, we expect live chat, in terms of persuading consumers, can be more effective when review volume is higher. If the fit dimension dominates, we expect live chat, as a source of product information, to be less effective when review volume is higher. Combining the two perspectives, we arrive at two competing hypotheses:

Hypothesis 5a: Live chat responsiveness tends to have a larger impact on conversion rate when review volume is higher.

Hypothesis 5b: Live chat responsiveness tends to have a lower impact on conversion rate when review volume is higher.

4. Data and Variables

Our dataset consists of aggregated monthly information of 2,469 sellers on Taobao in the category of women's clothing. We have a unique dataset that records seller's operation and performance for a 10-month period since May 2011. We have data regarding sellers' response rate of their live chat service and their monthly traffic data such as total web traffic and converted traffic. Our dataset also contains seller business information such as number of repeat customers and return product ratio. We also have detailed information for each seller including seller rating, number of reviews, average price, and number of product pictures. Table 5 lists detailed variable description. The summary statistics of the variables are displayed in Table 6. We discuss these variables in detail below.

------ Insert Tables 5 and 6 about Here ------

Total traffic is the number of unique web visitors that have browsed the seller's products. Similarly, *paid traffic* is the number of unique web visitors that ended up making purchases with the seller. *Conversion rate* is calculated as *paid traffic* divided by *total traffic*. From the summary statistics we can see the average conversion rate is about 2%. This number is consistent with current industry average for e-commerce, where a typical conversion rate is 1.33%-3.6%. (Bez 2016)

Response rate is a ratio measuring sellers' responsiveness on live chat. It is calculated based on total number of responded conversation divided by the number of live chat requests from customers. In our sample, the average response rate is about 57%.

Reputation refers to the cumulative product rating score for each seller over time. Prior literature suggests that reputation score can be a good proxy for seller size, as well as overall trustworthiness (Dellarocas 2003; Jin and Kato 2006; Elfenbein et al. 2012). In our dataset sellers'

reputations score ranging from 139 to 347,041, whereas the distribution is highly skewed. The mean of reputation variable is 10,164.56, and the standard deviation is 17,773.66.

Repeat customer refers to the number of customers who have made more than one purchases with the seller during the past 6 month. It varies significantly from seller to seller, ranging from no return customers to as high as 1,976 return customers.

New product refers to the number of new products launched by a seller within each month. The variable is also highly skewed, with the mean equals to 74.19 and standard deviation equals to 211.21.

Return ratio is the percentage of product returns within a month. It is interesting to see that the average return ratio is close to zero, indicating the majority of sellers in our sample have a very low return rate.

Price refers to the average price of all the products for a seller, shown in Chinese Yuan. It represents price levels of the sellers. Different price levels may attract different segments of consumers. Our summary statistics show that seller price ranges widely and have a skewed distribution. The average is 170.76, which is about $27, and the standard deviation is 162.41.

Rating refers to the detailed seller ratings for the sellers. The rating is a moving average of the scores a seller received in the past six months, and a customer can rate the seller on product quality as compared to description, service, and shipping, at a scale from from 1 to 5. Thus, the seller rating is a measure of store product and service quality. The average rating is 4.76, and the standard deviation is 0.17.

Review refers to the total number of newly-added product review for a seller in a month. On average, a seller accumulates 441.15 product reviews every month, with the standard deviation of 806.42.

Picture refers to the average number of pictures listed on sellers' product pages. We operationalize it as a proxy of how much detailed product information a seller provides on the product pages. On average, sellers post 4 pictures on each product.

5. Empirical Model

In this section we discuss how we develop our two stage model in a hierarchical Bayesian framework. Bayesian hierarchical models have been widely used in marketing and is a useful tool for modeling multi-faceted, non-linear phenomena, especially in analyzing (i) individual level decision making, and (ii) the differences of preferences and sensitivities among individuals (Allenby et al. 2005). Furthermore, heterogeneity can be better represented through Bayesian approach. For example, traditional econometrics methods use regression intercepts to model heterogeneity. With random coefficient model, we don't have to confine the differences to intercepts only, and we can model heterogeneity in slope coefficients (Rossi et al. 2005). We apply random coefficient model in the second stage to empirically identify whether live chat effectiveness is contingent on other factors related to consumer decision making. Specifically, we model live chat response rate as a function of seller characteristics, and, in the second stage, we model conversion probability as a function of live chat response rate with a distribution of coefficients across sellers. We describe our model in details below.

5.1. Modeling Live Chat Response Rate

To control for potential endogeneity, we first model live chat response rate as a function of seller characteristics, including reputation, price level, the number of repeated customer, return product ratio, and the number of new products. Basically, sellers' endogenous effort on providing live chat service is closely related to their own characteristics and heterogeneity. Therefore, we choose these instrument variables that are relevant to sellers' tenure and resources,

active level and price level. First of all, reputation is a good measurement of seller, it usually indicates a seller's overall size and how well a seller has developed herself on the marketplace. Compare to low reputation sellers, high reputation sellers usually have more capability and resources to invest in their businesses. Second, sellers with different pricing level and clientele may have different behavior in terms of investing in live chat. Sellers with different price level may attract different types of customers; thus the behavior of sellers would also be different. Third, sellers' active level on launching new products is also a good indictor of seller type. For example, if a seller launches new product quite often, he/she may want to seek more channel to promote these products or make customers aware of their new products. Additionally, we also control seller's return ratio in the model, as people may argue that live chat may happen after sales, because some customers may request chat in order to return a product. However, given our dataset, we see the average return ratio is almost zero, which indicates almost no after-sales chat service needed. The model is estimated through a panel data fixed effect approach, where we included u_i as a fixed effect, $u_i \sim N(0, \sigma_u^2)$, and the unobserved error term $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$. The model works as a control function for modelling conversion probability in the second stage of estimation.

$$Response\_rate_{it} = \gamma_0 + \gamma_1 Repuation_{it} + \gamma_2 New\_Product_{it} + \gamma_3 Price_{it} + \gamma_4 Repeat\_Customer_{it} + \gamma_5 Return\_Ratio_{it} + u_i + \varepsilon_{it}$$

Control function approach is a computationally simple method used in two-stage estimation. It is similar to two-stage least square estimation, but in a more flexible way. Wooldridge (2014, 2015) has shown that the control function could also be used in non-linear function as long as it is in the linear exponential family. In the first stage, we regress the live chat

response rate on a set of exogenous instrumental variables. We use the bootstrap method to obtain valid standard errors from the first stage, with the number of bootstrap replications set to 200. Later in the second stage, the residual from the first stage regression is included as a new covariate in the binomial model to control for endogeneity.

5.2. Modeling Conversion Probability

To model the conversion probability, we adopt a binomial model, in which each paid traffic y_{it} equals to the number of "success" out of "total traffic" trials. Following the binomial regression framework (Gelman and Hill 2006), we model the converted traffic following a binomial distribution:

$$y_{it} \sim Binomial(n_{it}, p_{it}),$$

where, y_{it} denotes the number of converted traffic in the current period t for seller i; n_{it} denotes the total number of traffic visiting a seller store in the period; and p_{it} is the conversion probability. By incorporating individual heterogeneity into the conversion probability, we have the following model on conversion rate:

$$p_{it} = logit^{-1}(x'_{it}\beta_i + \alpha_i + \zeta_{it}),$$

where,

$$x'_{it}\beta_i = \beta_0 + \beta_{1i} Response\_rate_{it} + \beta_{2i} log(Conversion_{i,t-1}).$$

In the above model, $Conversion_{i,t-1}$ denotes the lagged term of conversion rate in the prior period for store *i,* which models the carryover effect to future conversion rate; α is the seller level fixed effect, $\alpha_i \sim N(0, \sigma_\alpha^2)$. This model also accommodates over-dispersion. According to Gelman et al. (2007), a direct way to construct overdispersed binomial distribution is add normal error on the logistic scale. ζ_{it} is a data level error, and has its own normal

distribution $\zeta_{it} \sim N(0, \sigma_\zeta^2)$. The model reduces to binomial when $\sigma_\zeta^2 = 0$; otherwise it is over-dispersed.

To capture the unobserved heterogeneity, we also specify a distribution of coefficients across sellers. It follows a standard hierarchical formulation (Rossi et al. 2005):

$$\beta_i = \Delta Z_i + v_i,$$

where,

$$\mathrm{Z}' = (1, Price, Rating, Review, Picture).$$

This specification of Z and the hierarchical model above allow for a main effect and an interaction for sellers' product characteristics and information. For example, the effectiveness of live chat may be different depending on the seller product rating or price level. The delta matrix (Δ) is the coefficient matrix to be estimated. Furthermore, we model the unobserved error terms to be correlated and to follow a multivariate normal distribution, $[\mathrm{v_{i1}}, \mathrm{v_{i2}}]' \sim \mathrm{MVN}(0, \Sigma)$, where Σ is a 2×2 covariance matrix.

6. Empirical Analyses and Results

To estimate our model, we apply an MCMC methods using a Hamiltonian Monte Carlo algorithm. In particular, we run 4 chains for 1,000 iterations and use the last 2000 iterations to compute the posterior distribution of the model parameters. To ensure convergence, we test our model with different starting values and we find that all the chains are stable and converge to the same value.

We first present the results of the first stage parameter estimation in Table 7. The coefficient of reputation is 0.28 ($p<0.01$), indicating sellers with higher reputations tend to have higher response rates. The coefficient of average price is -0.06 ($p<0.01$), suggesting average price is negatively related to live chat response rate. In other words, sellers with lower price level

usually put more effort on responding to live chat. The coefficient of new product is 0.10 ($p<0.01$), indicating sellers who are active with more new products tend to respond to live chat more diligently. The coefficient of repeat customer is -0.02 ($p<0.01$), suggesting that sellers with more new customers tend to be more responsive to live chat requests. We also find that return ratio is negatively related to live chat response rate, with a coefficient of -0.39, indicating that sellers with lower return ratio have higher response rate. As mentioned before, we include return ratio in the model to control the effect of return on live chat. The results tend to rule out that live chat response is in response to return related issues.

------- Insert Table 7 about Here -------

The second stage coefficient estimates are presented in Table 8. Panel A shows the main results. Live chat response rate has a significant and positive impact on conversion, with a posterior mean of coefficient equal to 0.42. Also, there is significant carryover effect from last period, with a posterior mean of coefficient equal to 27.50. Panel B shows estimation from random coefficients. This hierarchical specification sheds light on how the effect of live chat is influenced by the existing seller level characteristics that are directly available to consumers. On one hand, to access the effect of informative role, we find that the random coefficient of the interaction effect between live chat and *picture* is -0.05, and is statistically significant, therefore hypothesis H1 is confirmed. This indicates that when a seller has on average fewer detailed pictures, live chat has a more positive effect on conversion probability. On the other hand, to access the effect of persuasive role, we find that the posterior mean of coefficient that captures interaction effect between live chat and *rating* is 1.09 and is statically significant. This indicates that the effect of live chat is increased when rating is at a higher level, therefore hypothesis H2 is confirmed; The posterior mean of coefficient that captures the interaction effect between live

chat and *price* is -0.3 and statistically significant, which shows that the effect of live chat is increased when price is at a lower level, therefore hypothesis H3 is confirmed; The posterior mean of coefficient captures interaction effect between live chat and *reviews* is 0.08, and is significant and positive. Therefore, we confirmed hypothesis H4a. It also implies that the volume of review actually signals the quality of the seller, so that the effect of live chat is increased when sellers have a higher review volume. All these findings indicate that the effect of live chat on conversion is more positive when product price is lower, when product rating is higher, and when the review volume is larger.

------- Insert Table 8 about Here -------

7. Discussion and Implications

Discussions

Based on our analysis, there is significant heterogeneity among sellers in providing live chat services. Their effort level on investing in live chat is correlated with the characteristics of online sellers, which are mainly related to sellers' quality (e.g. reputation) and their operational activities (e.g. number of new products, number of repeat customers). Reputation is one good indicator to measure online sellers' business ability or quality. On one hand, each online seller has to establish their reputation from every single transaction they have made. As customers all prefer to shop with high reputation sellers, the need to earn high reputation reduces adverse selection and leads to an increased exit of low quality sellers (Hui et al. 2016). Therefore, with higher reputation, sellers are more dedicated to provide better services. On the other hand, sellers with high reputation are usually the ones that have established themselves on the market. Their ability to spend more resources on customer services is a lot higher than the new sellers. We also find that sellers' effort level is dependent on their operational activities. For example, when

sellers have new products to launch, they will want to promote those products and put more effort on live chat services. Therefore, we observe positive correlation between live chat response rate and the number of new products. Similarly, sellers with less repeat customers will want to focus more on customer service so that to increase their customer retention.

After controlling for seller's self-selection bias, the significant effect of live chat on conversion rate highlights the importance of using live chat. On average, if the live chat response rate increases 10%, the conversion probability could increase by 4.29%, all else equal. Although this number seems small, given the high volume of online traffic, it could still make a big difference for many sellers. Especially for sellers with a large volume of traffic, improving live chat response rate will be beneficial to them.

The results from the random coefficient model confirmed our hypotheses that the effect of live chat does interact with the existing information dimensions. The differential effects of informative role and persuasive role are also confirmed. First, the informative role plays a significant role when product information is not comprehensive. When detailed information has already been provided, the informative role is less played in terms of addressing customers' concerns and questions regarding product matches. Our results show that there are two ways for consumers to access fit information, either through the product page, or through live chat. Therefore, the informative function of live chat could complement information loss on the product page. Second, live chat has a more positive effect on conversion when customers perceive higher value from the product. In online marketplace, review and ratings not only reduces information asymmetry but also enhance consumer valuation of the product. Although the utility function is different for different consumers, higher quality and lower prices are always preferred by consumers. We are able to identify the persuasive effect of live chat under

the situations when consumers give preferences toward the seller/product. In other words, live chat shift its role from informing consumers to influence consumers when there are indictors of high valuation of the product/sellers.

Implications

The results from our empirical analyses provide implications to online marketplace sellers. Many sellers invest in different kinds of advertising tools to increase online traffic, these investment are often wasted if there is no actual conversion being made. Given the average conversion rate for online marketplaces is only 2%, there is a great room for sellers to improve their online sales conversion. Live chat becomes a low cost and value added services to all the online sellers. In particular, it would be beneficial to invest in live chat for sellers that have a high volume of traffic.

Additionally, sellers should strategically utilize live chat service with the information they can provide to consumers. As we identify the informative role and persuasive role of live chat, it would become more efficient for sellers to utilize live chat under different circumstances. When sellers offer live chat service, they should pay attention on the way to approach customers. During a live chat, sellers can compliment a buyer's choice and highlight the positive things about their products, thus cultivating positive interpersonal feeling that is conducive to leading interested customers to make purchase. Ou et al. (2014) have shown that buyers who use instant messages to chat with sellers tend to have more trust and positive personal relationship with the seller. In addition, sellers could indicate how popular a certain item is among other buyers or even celebrities by showing the picture of celebrities using the same product, and the possible low inventory level of the item. In this way, potential customers are more likely to make up their mind and make the purchase.

8. Conclusion

The Internet has significantly expanded the ways that stores and markets communicate and interact with consumers (Sonnier et al. 2011). Internet advertising, word-of-mouth (WOM) reputation systems, and online reviews have fundamentally changed the way consumers search and purchase goods in an online environment (Bucklin et al. 2009). While online advertising and WOM have amassed a large amount of research, studies on the effect of live chat remains scant. In this paper, we focus on investigating the effectiveness of live chat on increasing conversion probability. We take seller heterogeneity into consideration and apply a Bayesian hierarchical model in our analysis. Our study complements the limited set of empirical literature examining the live chat effectiveness from a seller perspective. After accounting for the endogeneity problem of live chat effort by sellers, we demonstrate the positive impact of live chat on conversion probability.

Our study contributes to the traditional marketing communication literature. We stress the important role of live chat in communication to buyers, elaborate its major functions of informing and persuading, demonstrate its positive impact on increasing conversion, and most importantly, show how such impact vary across different sellers. Our results provide important implications for online retailing practices. First, we show that sellers' heterogeneous efforts on using live chat are related to seller attributes. Those sellers that are likely to invest in live chat are also the ones with better reputation and a greater potential to grow. Second, we highlight the importance of coordination between live chat services and information perceived by consumers. On one hand, our findings suggest that when product information is not complete on the item page, the inform function of live chat is even more crucial in increasing conversion. On the other hand, sellers should know that live chat service is more effective when consumers perceive

higher utility from their products, which occurs when the price is lower, the rating is higher, and the review volume is larger.

Although this research improves our understanding of live chat and its implication in improving seller performance, there are still some interesting questions that merit further investigation. First, our data is limited at an aggregated level and we are unable to identify the level of engagement for each live chat session. Additionally, we are unable to separate behavior from post purchase live chat service. Although the post purchase return ratio is quite low for our sample, it is possible that, for some customers, live chat happens after they make purchase, whereas our analysis is limited to pre-purchase behavior. With data availability, future research could use click-stream data to exam consumer behavior when live chat can be evaluated in each session. Second, since consumer perceive product information and form expected utility at the same time, we are unable to disentangle the extent to which live chat effect contributes the most. Perhaps future studies could examine the content of live chat messages and subsequent consumer behavior experimentally.

Chapter 4: The Treatment Effect of Buyer Protection Program in online Marketplace: Evidence from Taobao[3]

1. Introduction

Consumer feedback and reviews are becoming increasingly importantly in online markets. According to a recent study, restaurants that increased their ranking on Yelp by one star raised their revenues by 5 to 9 percent (Johnston 2013). Despite the enormous popularity and success of online reputation systems (Ba and Pavlou 2002, Dellarocas 2003, Resnick et al. 2006, Cabral and Hortacsu 2010), there are a number of problems. First, the ratings and reviews could be subject to fraud and user manipulation. In 2013, New York State regulators caught 19 companies that were writing fake reviews online for various companies (Streitfeld 2013). It was also reported that sellers at eBay engaged in fake transactions to boost their reputation scores (Gomes 2006), and some businesses often threatened and retaliated against customers who wrote bad reviews (Tuttle 2014). As a result, consumers have become more skeptical of the authenticity of online reviews. A survey suggests that 1 in 4 people believe customer ratings are biased or fake (Johnston 2013). Second, reputation systems could be an entry barrier for newcomers because establishing a good reputation could be a lengthy process, and entrants have difficulty competing with established sellers without a good reputation score (Roberts 2011). In addition, sellers who have achieved a high reputation may become insensitive to small changes in their reputation scores and lack incentives to maintain quality service over time (Fan et al. 2003).

Just like the physical world, online markets may have to reply on a multifaceted approach to reduce information asymmetry and build trust. In physical stores, in addition to word-of-mouth and advertisement, retailer also use warranties and money-back guarantees to obtain trusts

[3] This paper is coauthored with Ming Fan and Jianqing Chen.

from consumers. Recently, online companies and platforms have introduced similar consumer protection programs. For example, online marketplaces, such as eBay and Taobao, now provide the buyer protection program. The program covers transactions once purchased products are not delivered or do not match the ones described in listings. If a buyer does not receive the item as described and is unable to resolve the case directly with the seller, then the platform refunds the cost of the item and the shipping charge. Although this emerging mechanism in online marketplaces shows promise, rigorous studies on this subject are scant. This paper aims to fill the gap by empirically investigating the effects of the online warranty and guarantee programs.

We study a specific product guarantee program at Taobao, the largest online marketplace in China. The marketplace launched a consumer protection initiative in 2007 that includes a Return Guarantee Program (RGP). Since its inception, the program has been embraced by sellers to varying degrees across a range of product categories. For instance, in the clothing category, the program has been quite popular: In May 2011, about 50% of the sellers participated in the RGP.

The objective of this study is to examine whether the participation of buyer protection program bring impact on both seller and consumer behavior. We focus on how the market mechanism affects buyers' purchase decisions, and whether there are positive influences on sellers' subsequent behavior. We first look at two different types of consumer behavior when a customer sees the sign indicating the seller joined the buyer protection program: the browsing behavior and their purchasing behavior. Then we test whether the changes from consumer behavior will further influence sellers' behavior, from the perspective of reducing adverse selection and moral hazard.

In addressing these questions, the paper contributes to the rich literature in e-commerce and information systems. We approach the problem from both signaling and moral hazard perspectives. On one hand, we show that the benefit for sellers to join this program is to signal their trustworthiness to the buyers so that to increase more sales. On the other hand, because of the hidden nature of sellers' behavior, empirical work on moral hazard of participants in online markets is scarce. Empirical evidence on the behavior effects of the RGP can help the marketplace organizers better design programs and mechanisms that incentivize market participants to behave in desired ways. Also, our study focuses on an online mechanism that has received limited attention in prior information systems research, and the results of our study contribute to this important research stream.

The rest of the paper is organized as follows: We review the existing literature in Section 2, followed by research context and theories in Section 3. In Section 4, we describe our data, research design, and empirical analysis. We present the results in Section 5 and our discussion and managerial implications in Section 6. Section 7 concludes the paper.

2. Literature

Warranties and guarantees are provided by product sellers, and they serve as less than full insurance against unsatisfactory performance (Cooper and Ross 1985, Heal 1977, Lutz 1989). For example, stores offer money-back guarantees for products and allow a buyer to return the product for a full refund of purchase price within a certain period (Moorthy and Srinivasan 1995). Companies offer product warranties largely because of uncertainties and risks in market transactions. In brick-and-mortar stores, customers may not be able to assess product quality for experience goods, such as appliances and electronics. Even for search goods, whose features and

qualities can be quickly revealed after an inspection, uncertainties exist because consumers are increasingly using the online channel and cannot evaluate products before delivery.

From the sellers' perspective, two theories are relevant to warranty provisions. First, warranties or guarantees can be used by sellers to signal product quality (Lutz 1989, Spence 1977). Second, warranties can serve as an incentive for manufacturers to improve product quality and reduce the moral hazard problem (Cooper and Ross 1985, Lutz 1989). Both theories assume the existence of information asymmetry in the marketplace (Akerlof 1970); however, the nature of the asymmetry is different. The signaling theory assumes that product quality is exogenous and information asymmetry takes place because sellers have better knowledge about the product quality than buyers. When buyers lack product information, they might not be willing to purchase in the market. Product warranties or guarantees can be used as a way for firms to signal their product quality. In Spence's model (1977), high warranties signal high quality. High-quality producers can afford to offer close to full warranties because their products are unlikely to break. Low-quality producers cannot imitate the strategy because of the high warranty cost. Signaling theory, nevertheless, is unable to explain different warranty practices in the real world, where the relationship between warranties and product quality is more complicated. Empirical studies find that high-quality products are not always sold with higher warranties (Bryant and Gerner 1978, Priest 1981).

From the view of the second theory, information asymmetry also exists regarding buyers' and sellers' actions, which can directly affect product performance. Cooper and Ross (1985) developed a model in which moral hazard plays an important role. In contrast to the signaling model, product or service quality is chosen by the seller rather than being exogenously given. Under such an environment, warranty policies act as an incentive mechanism affecting the

behavior of both buyers and sellers. For example, when some sellers have a high cost in providing warranties, they might opt to reduce their warranty coverage and instead raise product quality.

Lutz (1989) also develops a model that is consistent with empirical evidence. The model assumes that product quality is hidden and the seller has to choose the warranty level. Consumers then decide their maintenance effort, which affects the performance of the product. The model shows that high-quality sellers might actually choose a low warranty, particularly when consumers increase their maintenance efforts because they conjecture that the product is of high quality. The studies by Cooper and Ross (1985) and by Lutz (1989) demonstrate the importance of the behavioral effect of product warranties. Product warranties and guarantees can serve as a vehicle to establish economic incentives for sellers to optimize their production efforts (Priest 1981).

Empirical studies in this area have focused on the signaling role of product warranties (e.g., Bryant and Gerner 1978, Chu and Chintagunta 2011, Kelley 1988, Priest 1981, Wiener 1985). The results are mixed. Studies have found both positive and negative relationships between warranty coverage and product reliability. A handful of studies on warranty programs provided by online sellers have recently been conducted (Dewally and Ederington 2006, Li et al. 2009, Robert 2011). The study by Dewally and Ederington (2006) tests the association of different signaling strategies and price differentials. Li et al. (2009) analyze data from eBay auctions and find that signals regarding direct quality indicators and seller credibility (e.g., picture postings, guarantees, reputation scores, and third-party payment method) lead to increased bidder participation. Roberts (2011) uses auction data on tractor sales, and the study finds that the “guaranteed or your money back” promise does not substitute for reputation, either

in determining price or the probability of sales. The study attributes the results to possible delays in buyer response to the guarantee and to skepticism about reimbursement in the event of fraud.

The purpose of this study is to examine the consequences of the use of the RGP on the behavior of both online consumers and sellers. We contribute to the literature and identify implications for the operation of online marketplaces in the following areas. First, some empirical studies (e.g., by Dewally and Ederington 2006, Li et al. 2009, and Robert 2011) have studied buyers' reactions to warranty or guarantee programs, but none of them have studied their browsing and purchasing behavior under the treatment of the guarantee program. With the aid of a unique dataset, we have the information about monthly consumer browsing and purchase decision with each seller, while observing the differences of these and compare their behavior before and after the participation.

Second, the context of the study provides a good test case for mechanisms to improve market efficiency. Theoretical work on signaling (Spence 1977) and moral hazard (Cooper and Ross 1985, Lutz 1989) provide interesting questions that we can study empirically. The online nature of the business has exacerbated the information asymmetry problem in the marketplace. The need for sellers to signal quality is high. Even more important is providing incentives for sellers to invest the efforts necessary to serve their online customers. Investigating empirically how market participants react to such a program—one that provides signaling opportunities, as well as incentives—provides interesting results.

3. Study Context and Theoretical Framework

3.1 Study Context

We study the use of the RGP for sellers in the clothing category of the marketplace. Before the introduction of the Consumer Protection Initiative in 2007, the marketplace we study

used a reputation system similar to that of eBay, where a buyer can rate a seller as positive, neutral, or negative after each transaction. The seller receives +1 point if the feedback is positive, 0 points if the rating is neutral, and -1 point if the rating is negative. Reputation scores for a seller accumulate over time.

The Consumer Protection Initiative comprises several separate programs, including a repair warranty program, authentic product assurance, and the RGP. These programs are meant to be tailored to different products, whose warranty requirements might vary. For example, a repair warranty is useful for electronic products, which might require repair services, while authentic product assurance is valuable for products that have strong brand recognition.

Clothing products in the marketplace are supplied by a large number of vendors, and most of the sellers do not have strong brand identities. Sellers usually must compete on style and price. Many customers shop for the latest designs and also seek bargains. One big concern from customers is that the delivered products can look very different from the images on the website. In addition, proper sizing can be difficult to gauge. Before the introduction of the RGP, return and exchange policies varied widely among the sellers. The return and exchange process could be quite difficult because of the lack of consumer protection in the retail industry in general. After the RGP's implementation, customers could return or exchange purchased products more easily. According to the rules set up by the marketplace organizer, the participating sellers in the RGP are required to unconditionally refund or exchange products for customers within seven days after product delivery. To join the program, sellers have to make a deposit of 1,000 RMB (about $160) in the marketplace. After joining the program, a seller can include the RGP logo for eligible products, and the seller's buyers can be compensated by the marketplace from the funds deposited when they have problems or disputes with the seller in returning purchased products.

Because the program reduces the risks associated with online purchases, sellers might find the RGP useful in attracting more customers.

3.2 Theoretical Framework

Significant information asymmetry exists in the online marketplace between sellers and buyers that could negatively affect the efficiency of the market and the welfare of all market participants. Because customers lack information about product quality and seller trustworthiness, they are reluctant to purchase products from unknown sellers or may have had a bad experience in shopping in the online market. We aim to study the impact of the RGP in terms of both consumer behavior and seller behavior. Our first objective of the research is to investigate how the RGP affects consumer behavior in terms of their browsing and purchasing behavior. Basically, consumers could perceive return guarantee as an alternative indicator of seller trustworthiness. Once the trust is established, consumers may tend to have close look on these sellers' products. Therefore, RGP can reinforce sellers' status no matter whether the seller has high or low reputation. The reason is that reputation score is an imperfect measures of seller trustworthiness. For example, online reputation score can be manipulated. Some strategic sellers who received negative ratings retaliated by giving the buyers negative ratings. One purpose of the retaliation is to induce mutual withdrawal of the negative ratings so that the negative records for the sellers can be removed (Bolton et al. 2009). There are other reported examples that some sellers develop high reputation scores by selling lower value items first and then switching to high-value categories (Gomes 2006). Buyers, who are aware of the potential problems associated with reputations, understand the noisy nature of reputation scores and will discount the reputation scores. The RGP can help reinforce a good reputation because sellers who participate in the RGP should be perceived to be more reliable. As a result, consumers will be more likely

to develop trust and purchase intension with the seller. The direct consequences are that they will like to spend more time on the seller's website, maybe browse more products and eventually purchase more items.

The second objective of the study is to evaluate whether the RGP has any behavioral effects in addition to its signaling function. Reputation concerns could provide incentives for sellers to work hard initially; but as careers progress, these incentives might disappear (Tadelis 2002). How to provide sustained incentives for sellers over time is an important question. If the RGP provides an incentive to sellers, it can complement reputation scores because of the difference in their root sources of motivation. The motivation to work hard under reputation systems comes from internal career concerns, while the incentive to do so under guarantee programs comes externally from the contracts in the market. Whereas internal motivation for sellers might disappear overtime, the contractual commitment still binds. In addition, logic suggests that sellers who offer return guarantees are more likely to adjust their behavior, putting forth more effort to improve their product or service quality. Being proactive is in the sellers' interest to avoid high return and exchange costs later on. Therefore, we expect to see a positive effect from RGP participation on service efforts of the sellers.

4. Data

4.1. Data Source

Our unique dataset includes 3548 sellers in the category of "clothing" products. We pick the clothing category for two reasons: First, it is one of the largest categories in the marketplace being studied, and second, return guarantee is applicable to this category. As suggested, buyers might find that the delivered clothes do not match the online descriptions and images, or that the

fit is wrong and, as a result, might want to return or exchange the products they bought. Our dataset consists observations from May 2011 to March 2012.

4.2. Variables

We have the data about whether sellers participate in the RGP program, and the sellers' current status, including reputation score, dynamic seller rating, and their operational information. We also have aggregated consumer browsing and purchasing data for each seller. The definitions of the variables are presented in Table 9. We discuss these variables in detail in the following paragraphs.

------ Insert Table 9 about Here ------

Guarantee is a dummy variable, indicating whether a seller participates in the RGP or not.

Reputation refers to the cumulative reputation score for each seller. Similar to eBay's reputation system, a buyer can rate the seller positive, neutral, and negative after each transaction. The seller receives +1 point if the feedback is positive, 0 point if the rating is neutral, and -1 point if the rating is negative. These scores accumulate over time for each seller. Prior literature suggests that the accumulated reputation score can be a proxy for seller size, as well as for seller trustworthiness (Rob and Fishman 2005, Jin and Kato 2006).

The variable *DSR* is the average product rating for a seller in the past six months. In addition to its cumulative reputation score, a seller also receives product quality ratings from buyers, ranging from one star to five stars. Note that the product rating is a relative measure of product quality, and a buyer is advised to rate a product based on whether the one received is consistent with the description on the website. Thus, *DSR* is a measure of overall seller quality.

We have several variables related to the sellers' operation. *Product variety* refers to the number of products a seller carries. *Average Price* is the average price of all the products the

seller carries. *Bookmark* indicates how many times the seller has been marked favorite by consumer. *Paid Advertising* and *Social Endorsement* are the two popular tools sellers usually adopt to attract traffic.

We also have outcome variables that measure consumer browsing behavior with each seller. *Customer average page views (CAPV)* measures how many seller product pages averagely a consumer browses. High CAPV often indicates that consumers are highly engaged and willing to explore more of sellers' products. Average order value (AOV) measures a consumer's average purchasing value with the seller, higher AOV usually indicates consumers tend to buy more products or spend more money with the seller, therefore a good indicator of seller growth potential. *Traffic* and *Conversion* are both crucial to the success of online sellers. *Traffic* measures sellers' popularity and their ability to obtain customer. *Conversion* measures sellers' ability to turn traffic into sales. *%PositiveReview* measures the percentage of overall sales that have been reviewed positively. The descriptive statistics of the variables are displayed in Table 10.

------ Insert Table 10 about Here ------

4.4. Research Design and Empirical Analysis

It is possible that sellers' specific unobserved factors may jointly determine sellers' decision to participate the buyer guarantee program. There could be endogeneity issues arising because of self-selection that need to be considered for identification of the effect of buyer guarantee program. To resolve this issues, we utilize propensity score matching in combination with difference in difference analysis. A DID estimator represents the difference in pre and post participation differences between two groups of sellers, the treatment group (sellers who participate) and the control group(sellers who do not participate).

The challenge is to compose the control group that comprises sellers who are very similar to the treatment group sellers in all aspects but for the fact that they have not participate the program. The difference in behavior between the two groups can then be attributed to the participation of the program. Therefore, we conduct the propensity score analysis using a logistic model formulation because the program participation is binary. We then perform the process of matching the treatment and control group using the estimated propensity score. We used "nearest neighbor matching", and we are able to match 466 treated sellers to equivalent control sellers. The logistic regression results are presented in Table 11.

------ Insert Table 11 about Here ------

We check to see if the underlying assumptions of the PSM process hold. We check if there is substantial overlap in the characteristics of the sellers who participate the program and those who do not. Figure 6 also shows the matching result before and after matching, and this shows evidence for the existence of common support.

------ Insert Figure 6 about Here ------

We then perform propensity score matching for the matched control and treatment group, we use four outcome variables to measure consumer and seller behavior respectively. In terms of consumer behavior, we use CAPV, AOV and Conversion to measure their browsing and purchasing behavior after they see the seller they are visiting has participated in RGP. We also measure if participating in RGP could alleviate the behavior of moral hazard. We study changes in seller performance measure: DSR and percentage of positive reviews. We provide the results in the next section.

5. Results

5.1. Results for consumer behavior

We display the difference-in-differences values for different outcome variables for the treatment and control groups before and after the seller join the guarantee program. We exam whether there are significant changes of consumer behavior in terms of their browsing and purchasing behavior. We measure consumer browsing behavior through customer average page view(CAPV) which reflects the depth of consumer visits with the seller. The results are presented in Table 12. As can be seen from the table, the increase for CAPV is not significant for the control group, but this increase is significant for the treatment group (0.071 after participation versus 0.034 before participation). However, the difference-in-difference result is not significant (0.036, t-test=1.07). This suggests that participation in the RGP doesn't positively influence the consumer browsing behavior.

------ Insert Table 12 about Here ------

We later measure consumer purchasing behavior through both average order value (AOV) and conversion. AOV reflects the value of consumer purchases, and conversion shows the effectiveness of signaling. The results are presented in Table 13. As can be seen from the table, the difference-in-difference result of before and after seller participating the program is very significant (1.83, t-test=10.19). This suggests that consumers tend to purchase more valuable items when they know that the seller participates in the RGP.

------ Insert Table 13 about Here ------

The increase for conversion is not significant for the control group and treatment group, but the increase of difference-in-difference is marginally significant for conversion (Table 14). It suggests that there is some signaling effect of RGP that could attract customers to make purchases with the seller.

------ Insert Table 14 about Here ------

5.2. Results for seller behavior

Besides signaling good quality to the consumers, participating in the RGP also could affect seller behavior. We use this matching and difference-in-difference approach to measure the changes in seller performance: DSR and %PositiveReview. Basically, we observe whether there are increases in sellers' DSR and the percentage of positive reviews for both control and treatment group before and after the program participation. We present the results in Table 15 and Table 16.

------ Insert Table 15 and 16 about Here ------

The results in Table 15 show there is significant increase of DSR for the treatment group, while before participation, treatment group has a significantly lower DSR than the control group. This further confirms that participation to the program impacts seller performance rating, or in other words, sellers provides better products or services after they join the program. The results in Table 16 also shows that the increase of positive review is significant. Before participation, the treatment group has lower percentage of positive review compared to the control group, although after the participation there is no significant differences between the two groups, the difference-in-difference is significantly positive. Therefore, both DSR and %PositiveReview has suggested that sellers performance was significantly increased after they participate in RGP.

6. Discussions and Implications for Practice

6.1. Discussions

Our results suggest a positive and significant treatment effect of the RGP on both consumer and seller behavior. We find that the signaling effect on consumer purchase is significant, where there is significant increase in sales conversion after the seller participates in

the program. Although we didn't find significant impact on consumer browsing behavior, we find that consumer order value increases after the participation. It indicates that consumers tend to buy more expensive products after they see the RGP participation. Previous study (Kim and Krishnan 2015) has found that consumers are unlikely to buy expensive products online especially with the presence of product uncertainty. This further confirms our result that the treatment effect of RGP alleviates product uncertainty, so that consumers become less conservative and tend to buy more expensive products.

In studying the effect of RGP participation on sellers' behavior, we find that the program can alleviate moral hazard behavior, where sellers have increased rating and positive review after the participation. It could simply because sellers tend to put more effort on their products and services after they join the program, otherwise consumers will eventually return the products under the guarantee program. The program itself acts like

6.2. Implications for Practice

The results of the study provide interesting implications for improving the performance of electronic marketplaces. RGPs provide several benefits, including reinforcing the existing reputation system and giving sellers incentives to invest on service. First, RGPs provide value to the market by allowing high-reputation sellers to further differentiate themselves from other sellers. Given the problems associated with reputation systems, as identified previously, a good reputation score may be too noisy to prove the trustworthiness and reliability of the sellers. RGPs can complement the reputation scores and help buyers to feel more comfortable with these sellers. Second, RGPs can provide utility to buyers by mitigating the risks in online shopping. In addition, our results also show that RGPs have an added behavioral effect on sellers. They show that market participants could change their behavior after participating in a RGP. Therefore,

market organizers can design programs and mechanisms to encourage certain behavior from participants. Conducting follow-up studies and evaluating both the behavior change of market participants and whether a program achieves its designed goals after it is implemented are important steps to take.

One direct implication for the marketplace is to put the RGP logo in a more prominent position on seller pages. Currently, in the marketplace studied, the reputation scores are presented in a more prominent place than the RGP logo, and they serve as the attention catcher. As a result, buyers might naturally use reputation scores as the first criterion in screening sellers. In addition, a marketplace could explore other ways to allow quality sellers to compete more effectively with incumbents. Prior literature has studied the possibility of tradable reputations (Tadelis 2002, Xu et al. 2008), so that quality sellers can acquire good reputations to jump-start their business. However, this proposal could still be expensive for startup buyers who lack financial resources.

7. Conclusions

Our study examines an RGP in a large online marketplace in China. We try to answer the questions of whether the use of the RGP has any behavioral effect on sellers in improving service quality, as well as on consumers in terms of their browsing and purchasing behavior. Our results show that the RGP has a positive effect on both sellers and consumers. The signaling effect on consumers is significant which leads higher conversion and higher price purchases. Also, we find that participating in RGP could alleviate online sellers' moral hazard problem.

Chapter 5: Conclusion

In this dissertation, I use data from Taobao marketplace, and examine the problems that related to the behavior of online sellers and how they could improve their performance by utilizing different tools that are available to the marketplace platform. In essay 1, I explore what affect sellers' choices of advertising strategy and how those choices affect online traffic for heterogeneous sellers. I empirically identify that the two advertising strategies, social media endorsement and sponsored search, are partially substitutionary, and, hence, sellers should maximize their benefit from choosing the right strategy according to their heterogeneous characteristics. The results produce good managerial insights for both online marketplace and marketplace sellers. First, our findings provide a good foundation for marketplace to target their service and allow them to design better policy to accommodate the sellers. Our findings indicate that the adoption of those advertising strategies may not be uniformly beneficial to all sellers. Therefore, the marketplace has to provide sellers with the most effective tool, which will benefit both the platform and the sellers. Second, sellers can benefit from choosing the tools strategically in order to take fully advantage of the paid services. This insight serves to help sellers to come up with better planning and resource allocation in future online advertising campaign. In essay 2, I focus on investigating the effectiveness of live chat on increasing conversion probability. I take seller heterogeneity into consideration and apply a Bayesian hierarchical model in our analysis. This study complements the limited set of empirical literature examining the live chat effectiveness from a seller perspective. After accounting for the endogeneity problem of live chat effort level by sellers, we demonstrate the positive impact of live chat on conversion probability. As a low cost and value added services to all the online sellers, live chat would be beneficial to sellers that have a high volume of traffic. Also, the importance of coordinating the informative

role and persuasive role highlights that sellers should strategically utilize live chat service with the information they can provide to consumers. In essay 3, we find that the signaling effect of return guarantee program is significant. Consumers tend to establish trust with the sellers who have participate the program. As a result, consumers are more likely to make purchases with the seller and spend on higher price products. There also exists positive correction effect on the sellers. We find significantly higher review ratings and more positive reviews for the sellers after they have participated the program. This indicates that the program has alleviated the moral hazard behaviors among online sellers.

REFERENCES

References for Essay 1 (Chapter 2)

Agarwal, A., Hosanagar, K., and Smith, M. D. 2015. "Do Organic Results Help or Hurt Sponsored Search Performance?," *Information Systems Research* (26:4), pp. 695–713 (doi: 10.1287/isre.2015.0593).

Agrawal, J., and Kamakura, W. A. 1995. "The Economic Worth of Celebrity Endorsers: An Event Study Analysis," *Journal of Marketing* (59:3), pp. 56–62 (doi: 10.2307/1252119).

Alibaba Group. 2015. "2015 annual report of the Alibaba Group.," (available at http://www.alibabagroup.com/en/ir/secfilings; retrieved February 8, 2016).

Anderson, S. P., and Renault, R. 2006. "Advertising Content," *American Economic Review* (96:1), pp. 93–113 (doi: 10.1257/000282806776157632).

Aral, S., and Walker, D. 2011. "Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks," *Management Science* (57:9), pp. 1623–1639 (doi: 10.1287/mnsc.1110.1421).

Aral, S., and Walker, D. 2012. "Identifying Influential and Susceptible Members of Social Networks," *Science* (337:6092), pp. 337–341 (doi: 10.1126/science.1215842).

Athey, S., and Ellison, G. 2011. "Position Auctions with Consumer Search," *Quarterly Journal of Economics* (126:3), pp. 1213–1270 (doi: 10.1093/qje/qjr028).

Ba, S., and Pavlou, P. A. 2002. "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly* (26:3), pp. 243–268 (doi: 10.2307/4132332).

Bagwell, K. 2007. "Chapter 28 The Economic Analysis of Advertising," in *Handbook of Industrial Organization*M. A. and R. Porter (ed.) (Vol. 3), Elsevier, pp. 1701–1844.

Bain, J. S. 1956. *Barriers to new competition: their character and consequences in manufacturing industries*, Harvard University Press.

Benbunan-Fich, R., and Fich, E. M. 2004. "Effects of Web Traffic Announcements on Firm Value," *International Journal of Electronic Commerce* (8:4), pp. 161–181 (doi: 10.1080/10864415.2004.11044312).

Bhargava, H. K., and Choudhary, V. 2004. "Economics of an Information Intermediary with Aggregation Benefits," *Information Systems Research* (15:1), pp. 22–36 (doi: 10.1287/isre.1040.0014).

Bickart, B., and Schindler, R. M. 2001. "Internet forums as influential sources of consumer information," *Journal of Interactive Marketing* (15:3), pp. 31–40 (doi: 10.1002/dir.1014).

Blake, T., Nosko, C., and Tadelis, S. 2015. "Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment," *Econometrica* (83:1), pp. 155–174 (doi: 10.3982/ECTA12423).

Bolton, G. E., Katok, E., and Ockenfels, A. 2004. "How Effective Are Electronic Reputation Mechanisms? An Experimental Investigation," *Management Science* (50:11), pp. 1587–1602 (doi: 10.1287/mnsc.1030.0199).

Bolton, G., Greiner, B., and Ockenfels, A. 2012. "Engineering Trust: Reciprocity in the Production of Reputation Information," *Management Science* (59:2), pp. 265–285 (doi: 10.1287/mnsc.1120.1609).

Braithwaite, D. 1928. "The Economic Effects of Advertisement," *The Economic Journal* (38:149), pp. 16–37 (doi: 10.2307/2224394).

Butler, R. J., Cowan, B. W., and Nilsson, S. 2005. "From obscurity to bestseller: Examining the impact of Oprah's Book Club selections," *Publishing Research Quarterly* (20:4), pp. 23–34 (doi: 10.1007/s12109-005-0045-2).

Cabral, L., and Hortaçsu, A. 2010. "The Dynamics of Seller Reputation: Evidence from Ebay*," *The Journal of Industrial Economics* (58:1), pp. 54–78 (doi: 10.1111/j.1467-6451.2010.00405.x).

Chan, D. X., Yuan, Y., Koehler, J., and Kumar, D. 2011. "Incremental Clicks The Impact of Search Advertising," *Journal of Advertising Research* (51:4), pp. 643–647 (doi: 10.2501/JAR-51-4-643-647).

Chan, T. Y., Wu, C., and Xie, Y. 2011. "Measuring the Lifetime Value of Customers Acquired from Google Search Advertising," *Marketing Science* (30:5), pp. 837–850 (doi: 10.1287/mksc.1110.0658).

Chatterjee, P., Hoffman, D., and Novak, T. 2003. "Modeling the Clickstream: Implications for Web-Based Advertising Efforts," *Marketing Science* (22:4), pp. 520–541 (doi: 10.1287/mksc.22.4.520.24906).

Chen, J., Fan, M., and Li, M. 2016. "Advertising versus Brokerage Model for Online Trading Platforms," *MIS Quarterly* (40:3), pp. 575–596.

China Renaissance. 2013. "Alibaba Group," (available at http://www.chinarenaissance.com/uploads/soft/130114/1-1306241035E.pdf).

Clemons, E., Gao, G., and Hitt, L. 2006. "When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry," *J. Manage. Inf. Syst.* (23:2), pp. 149–171 (doi: 10.2753/MIS0742-1222230207).

Comanor, W. S., and Wilson, T. A. 1974. *Advertising and Market Power*, Harvard University Press.

Dellarocas, C., Zhang, X. (Michael), and Awad, N. F. 2007. "Exploring the value of online product reviews in forecasting sales: The case of motion pictures," *Journal of Interactive Marketing* (21:4), pp. 23–45 (doi: 10.1002/dir.20087).

Dewally, M., and Ederington, L. 2006. "Reputation, Certification, Warranties, and Information as Remedies for Seller-Buyer Information Asymmetries: Lessons from the Online Comic Book Market," *The Journal of Business* (79:2), pp. 693–729 (doi: 10.1086/499169).

Dewan, R. M., Freimer, M. L., and Zhang, J. 2002. "Management and Valuation of Advertisement-Supported Web Sites," *Journal of Management Information Systems* (19:3), pp. 87–98 (doi: 10.1080/07421222.2002.11045737).

Drèze, X., and Hussherr, F.-X. 2003. "Internet advertising: Is anybody watching?," *Journal of Interactive Marketing* (17:4), pp. 8–23 (doi: 10.1002/dir.10063).

Duan, W., Gu, B., and Whinston, A. B. 2008. "The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry," *Journal of Retailing* (84:2), pp. 233–242 (doi: 10.1016/j.jretai.2008.04.005).

Elfenbein, D. W., Fisman, R., and Mcmanus, B. 2012. "Charity as a Substitute for Reputation: Evidence from an Online Marketplace," *Review of Economic Studies* (79:4), pp. 1441–1468 (doi: 10.1093/restud/rds012).

Esteban, L., Gil, A., and Hernández, J. M. 2001. "Informative Advertising and Optimal Targeting in a Monopoly," *The Journal of Industrial Economics* (49:2), pp. 161–180.

Fang, X., Hu, P. J.-H., Li, Z. (Lionel), and Tsai, W. 2013. "Predicting Adoption Probabilities in Social Networks," *Information Systems Research* (doi: 10.1287/isre.1120.0461).

Godes, D., and Mayzlin, D. 2004. "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science* (23:4), pp. 545–560 (doi: 10.1287/mksc.1040.0071).

Greene, W. H. 2012. *Econometric analysis* (7th ed.), Boston: Prentice Hall.

Gu, B., Park, J., and Konana, P. 2012. "The Impact of External Word-of-Mouth Sources on Retailer Sales of High-Involvement Products," *Information Systems Research* (23:1), pp. 182–196 (doi: 10.1287/isre.1100.0343).

Hausman, J., Hall, B. H., and Griliches, Z. 1984. "Econometric Models for Count Data with an Application to the Patents-R & D Relationship," *Econometrica* (52:4), pp. 909–938 (doi: 10.2307/1911191).

Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error," *Econometrica* (47:1), pp. 153–161 (doi: 10.2307/1912352).

Hertzendorf, M. N. 1993. "I'm Not a High-Quality Firm -- But I Play One on TV," *RAND Journal of Economics* (24:2), pp. 236–247.

Houser, D., and Wooders, J. 2006. "Reputation in Auctions: Theory, and Evidence from eBay," *Journal of Economics & Management Strategy* (15:2), pp. 353–369 (doi: 10.1111/j.1530-9134.2006.00103.x).

iResearch. 2014. "With Revenue of $7.5 Billion in 2013, Alibaba Restarts IPO_Views_Insights_iResearch," (available at http://www.iresearchchina.com/content/details7_15533.html).

Iyengar, R., Bulte, C. V. den, and Valente, T. W. 2011. "Rejoinder—Further Reflections on Studying Social Influence in New Product Diffusion," *Marketing Science* (30:2), pp. 230–232 (doi: 10.1287/mksc.1100.0614).

Jeziorski, P., and Segal, I. 2015. "What Makes Them Click: Empirical Analysis of Consumer Demand for Search Advertising," *American Economic Journal-Microeconomics* (7:3), pp. 24–53 (doi: 10.1257/mic.20100119).

Jin, G. Z., and Kato, A. 2006. "Price, quality, and reputation: evidence from an online field experiment," *The RAND Journal of Economics* (37:4), pp. 983–1005 (doi: 10.1111/j.1756-2171.2006.tb00067.x).

Liaukonyte, J., Teixeira, T., and Wilbur, K. C. 2015. "Television Advertising and Online Shopping," *Marketing Science* (34:3), pp. 311–330 (doi: 10.1287/mksc.2014.0899).

Liu, Y. 2006. "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing* (70:3), pp. 74–89 (doi: 10.1509/jmkg.70.3.74).

Livingston, J. A. 2005. "How Valuable Is a Good Reputation? A Sample Selection Model of Internet Auctions," *Review of Economics and Statistics* (87:3), pp. 453–465 (doi: 10.1162/0034653054638391).

Luo, X., and Zhang, J. 2013. "How Do Consumer Buzz and Traffic in Social Media Marketing Predict the Value of the Firm?," *Journal of Management Information Systems* (30:2), pp. 213–238 (doi: 10.2753/MIS0742-1222300208).

Moe, W. W., and Schweidel, D. A. 2012. "Online Product Opinions: Incidence, Evaluation, and Evolution," *Marketing Science* (31:3), pp. 372–386 (doi: 10.1287/mksc.1110.0662).

Narayanan, S., and Kalyanam, K. 2015. "Position Effects in Search Advertising and their Moderators: A Regression Discontinuity Approach," *Marketing Science* (34:3), pp. 388–407 (doi: 10.1287/mksc.2014.0893).

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

Oestreicher-Singer, G., and Sundararajan, A. 2012. "Recommendation Networks and the Long Tail of Electronic Commerce," *MIS Quarterly* (36:1), pp. 65-A4.

Osawa, J. 2013. "How Does Alibaba Make Money? - Digits - WSJ," *Wall Street Journal* (available at http://blogs.wsj.com/digits/2013/09/09/how-does-alibaba-make-money/).

Peng, J., Bulte, V. den, and Christophe. 2016. "Participation vs. Effectiveness of Paid Endorsers in Social Advertising Campaigns: A Field Experiment," SSRN Scholarly Paper No. ID 2702053, , Rochester, NY: Social Science Research Network (available at http://papers.ssrn.com/abstract=2702053).

Perdikaki, O., Kesavan, S., and Swaminathan, J. 2011. "Effect of Traffic on Sales and Conversion Rates of Retail Stores," *Manufacturing & Service Operations Management* (14:1), pp. 145–162 (doi: 10.1287/msom.1110.0356).

Rabe-Hesketh, S., and Skrondal, A. 2012. *Multilevel and Longitudinal Modeling Using Stata, Third Edition*, Stata Press (available at http://www.stata.com/bookstore/multilevel-longitudinal-modeling-stata/).

Resnick, P., Zeckhauser, R., Swanson, J., and Lockwood, K. 2006. "The value of reputation on eBay: A controlled experiment," *Experimental Economics* (9:2), pp. 79–101 (doi: 10.1007/s10683-006-4309-2).

Rob, R., and Fishman, A. 2005. "Is Bigger Better? Customer Base Expansion through Word-of-Mouth Reputation," *Journal of Political Economy* (113:5), pp. 1146–1162 (doi: 10.1086/444552).

Rutz, O. J., Trusov, M., and Bucklin, R. E. 2011. "Modeling Indirect Effects of Paid Search Advertising: Which Keywords Lead to More Future Visits?," *Marketing Science* (30:4), pp. 646–665 (doi: 10.1287/mksc.1110.0635).

Soh, C., Markus, M. L., and Goh, K. H. 2006. "Electronic marketplaces and price transparency: Strategy, information technology, and success," *Mis Quarterly* (30:3), pp. 705–723.

Susarla, A., Oh, J.-H., and Tan, Y. 2012. "Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube," *Information Systems Research* (23:1), pp. 23–41 (doi: 10.1287/isre.1100.0339).

Trusov, M., Bodapati, A. V., and Bucklin, R. E. 2010. "Determining Influential Users in Internet Social Networks," *Journal of Marketing Research (JMR)* (47:4), pp. 643–658.

Tucker, C. E. 2016. "Social Advertising: How Advertising that Explicitly Promotes Social Influence Can Backfire," SSRN Scholarly Paper No. ID 1975897, , Rochester, NY: Social Science Research Network (available at http://papers.ssrn.com/abstract=1975897).

Wei, Y. D. 2013. *Regional Development in China: States, Globalization and Inequality*, Routledge.

Wong, G., and Chu, K. 2015. "The Rat Race to Advertise on Alibaba's Marketplaces," *Wall Street Journal* (available at http://www.wsj.com/articles/the-race-to-advertise-on-alibabas-marketplaces-1425368909).

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

Wu, J., Cook, V., and Strong, E. 2005. "A Two-Stage Model of the Promotional Performance of Pure Online Firms," *Information Systems Research* (16:4), pp. 334–351 (doi: 10.1287/isre.1050.0071).

Xu, L., Chen, J., and Whinston, A. 2012. "Effects of the Presence of Organic Listing in Search Advertising," *Information Systems Research* (23:4), pp. 1284–1302 (doi: 10.1287/isre.1120.0425).

Yang, S., and Ghose, A. 2010. "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?," *Marketing Science* (29:4), pp. 602–623 (doi: 10.1287/mksc.1090.0552).

Ye, S., Gao, G., and Viswanathan, S. 2014. "Strategic Behavior in Online Reputation Systems: Evidence from Revoking," *Mis Quarterly* (38:4), pp. 1033–1056.

References for Essay 2 (Chapter 3)

Agarwal, A., Hosanagar, K., and Smith, M. D. 2011. "Location, Location, Location: An Analysis of Profitability of Position in Online Advertising Markets," *Journal of Marketing Research*, (48:6), pp. 1057–1073.

Ahluwalia, R. 2002. "How prevalent is the negativity effect in consumer environments?," *Journal of Consumer Research*, (29:2) (available at https://experts.umn.edu/en/publications/how-prevalent-is-the-negativity-effect-in-consumer-environments(05c7c221-e33c-4399-86ad-92fade43f775).html).

Bagwell, K. 2007. "Chapter 28 The Economic Analysis of Advertising," in *Handbook of Industrial Organization*, M. A. and R. Porter (ed.), (Vol. 3), Elsevier, pp. 1701–1844.

Bez, R. 2016. "Ecommerce Conversion Rates - 2016 Benchmarks," *Compass Blog*, September 14 (available at https://blog.compass.co/ecommerce-conversion-rates-benchmarks-2016/; retrieved March 16, 2017).

Blohm, I., Riedl, C., Fueller, J., and Leimeister, J. M. 2016. "Rate or Trade? Identifying Winning Ideas in Open Idea Sourcing," *Information Systems Research*, (27:1), pp. 27–48 (doi: 10.1287/isre.2015.0605).

Braithwaite, D. 1928. "The Economic Effects of Advertisement," *The Economic Journal*, (38:149), pp. 16–37 (doi: 10.2307/2224394).

Bucklin, R., Oliver J. Rutz, and Michael Trusov. 2009. "Metrics for the New Internet Marketing Communications Mix," in *Review of Marketing Research*, Review of Marketing Research, (Vols. 1-0, Vol. 5), Emerald Group Publishing Limited, pp. 175–192 (available at http://www.emeraldinsight.com/doi/abs/10.1108/S1548-6435%282008%290000005011).

Burtch, G., Ghose, A., and Wattal, S. 2016. "Secret Admirers: An Empirical Examination of Information Hiding and Contribution Dynamics in Online Crowdfunding," *Information Systems Research*, (27:3), pp. 478–496 (doi: 10.1287/isre.2016.0642).

Chen, Y., and Xie, J. 2008. "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science*, (54:3), pp. 477–491.

Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research (JMR)*, (43:3), pp. 345–354.

Clemons, E., Gao, G., and Hitt, L. 2006. "When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry," *J. Manage. Inf. Syst.*, (23:2), pp. 149–171 (doi: 10.2753/MIS0742-1222230207).

Dellarocas, C., Zhang, X. (Michael), and Awad, N. F. 2007. "Exploring the value of online product reviews in forecasting sales: The case of motion pictures," *Journal of Interactive Marketing*, (21:4), pp. 23–45 (doi: 10.1002/dir.20087).

Devaraj, S., Fan, M., and Kohli, R. 2002. "Antecedents of B2C Channel Satisfaction and Preference: Validating e-Commerce Metrics," *Information Systems Research*, (13:3), pp. 316–333 (doi: 10.1287/isre.13.3.316.77).

Dewan, S., and Ramaprasad, J. 2012. "Research Note—Music Blogging, Online Sampling, and the Long Tail," *Information Systems Research*, (23:3-NaN-2), pp. 1056–1067 (doi: 10.1287/isre.1110.0405).

Dhar, V., and Ghose, A. 2010. "Research Commentary—Sponsored Search and Market Efficiency," *Information Systems Research*, (21:4), pp. 760–772 (doi: 10.1287/isre.1100.0315).

Dimoka, A., Hong, Y., and Pavlou, P. A. 2012. "On Product Uncertainty in Online Markets: Theory and Evidence," *Mis Quarterly*, (36:2), pp. 395–426.

Duan, W., Gu, B., and Whinston, A. B. 2008. "The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry," *Journal of Retailing*, (84:2), pp. 233–242 (doi: 10.1016/j.jretai.2008.04.005).

Godes, D., and Mayzlin, D. 2004. "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, (23:4), pp. 545–560 (doi: 10.1287/mksc.1040.0071).

Godes, D., and Mayzlin, D. 2009. "Firm-Created Word-of-Mouth Communication: Evidence from a Field Test," *Marketing Science*, (28:4), pp. 721–739 (doi: 10.1287/mksc.1080.0444).

Hertzendorf, M. N. 1993. "I'm Not a High-Quality Firm -- But I Play One on TV," *RAND Journal of Economics*, (24:2), pp. 236–247.

Hui, X., Saeedi, M., Shen, Z., and Sundaresan, N. 2016. "Reputation and Regulations: Evidence from eBay," *Management Science*, (62:12), pp. 3604–3616 (doi: 10.1287/mnsc.2015.2323).

J.D.Power. 2013. "2013 U.S. Wireless Customer Care Full-Service Performance Study," Text, (available at http://www.jdpower.com/press-releases/2013-us-wireless-customer-care-full-service-performance-study-volume-2-and-2013-us).

Johnson, S. L., Safadi, H., and Faraj, S. 2015. "The Emergence of Online Community Leadership," *Information Systems Research*, (26:1), pp. 165–187 (doi: 10.1287/isre.2014.0562).

Keller, K. L. 1987. "Memory Factors in Advertising: The Effect of Advertising Retrieval Cues on Brand Evaluations," *Journal of Consumer Research*, (14:3), pp. 316–333.

Keller, K. L. 2001. "Mastering the Marketing Communications Mix: Micro and Macro Perspectives on Integrated Marketing Communication Programs," *Journal of Marketing Management*, (17:7/8), pp. 819–847.

Kim, Y., and Krishnan, R. 2015. "On Product-Level Uncertainty and Online Purchase Behavior: An Empirical Analysis," *Management Science*, (61:10), pp. 2449–2467 (doi: 10.1287/mnsc.2014.2063).

Kwark, Y., Chen, J., and Raghunathan, S. 2013. "Platform or Wholesale? Different Implications for Retailers of Online Product," Working Paper No. 13–14, , NET Institute (available at https://ideas.repec.org/p/net/wpaper/1314.html).

Li, H., and Kannan, P. K. 2014. "Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment," *Journal of Marketing Research*, (51:1), pp. 40–56 (doi: 10.1509/jmr.13.0050).

Manchanda, P., Dubé, J.-P., Goh, K. Y., and Chintagunta, P. K. 2006. "The Effect of Banner Advertising on Internet Purchasing," *Journal of Marketing Research*, (43:1), pp. 98–108.

Moe, W. W., and Fader, P. S. 2004. "Dynamic Conversion Behavior at E-Commerce Sites," *Management Science*, (50:3), pp. 326–335.

MorganStanley Ltd. 2005. "Creating Consumer Value in Digital China," MorganStanley Ltd (available at http://down.cenet.org.cn/upfile/10/200510621638189.pdf).

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

Ou, C., Pavlou, P. A., and Davison, R. M. 2014. "Swift Guanxi in Online Marketplaces: The Role of Computer-Mediated Communication Technologies," *MIS Quarterly*, (38:1), pp. 209-A24.

Perdikaki, O., Kesavan, S., and Swaminathan, J. 2011. "Effect of Traffic on Sales and Conversion Rates of Retail Stores," *Manufacturing & Service Operations Management*, (14:1), pp. 145–162 (doi: 10.1287/msom.1110.0356).

Robinson, J. 1969. *The Economics of Imperfect Competition*, Springer.

Sismeiro, C., and Bucklin, R. E. 2004. "Modeling Purchase Behavior at an E-Commerce Web Site: A Task-Completion Approach," *Journal of Marketing Research*, (41:3), pp. 306–323.

Sonnier, G., McAlister, L., and Rutz, O. 2011. "A Dynamic Model of the Effect of Online Communications on Firm Sales," *Marketing Science*, (30:4), pp. 702–716 (doi: 10.1287/mksc.1110.0642).

Xu, L., Duan, J., and Whinston, A. 2014. "Path to Purchase: A Mutually Exciting Point Process Model for Online Advertising and Conversion," *Management Science*, (60:6), pp. 1392–1412 (doi: 10.1287/mnsc.2014.1952).

Yadav, M. S., and Varadarajan, R. 2005. "Interactivity in the electronic marketplace: An exposition of the concept and implications for research," *Journal of the Academy of Marketing Science*, (33:4), pp. 585–603 (doi: 10.1177/0092070305278487).

Zeng, X., and Wei, L. 2013. "Social Ties and User Content Generation: Evidence from Flickr," *Information Systems Research*, (24:1), pp. 71–87 (doi: 10.1287/isre.1120.0464).

Zhu, F., and Zhang, X. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing*, (74:2), pp. 133–148.

Zhu, L., Benbasat, I., and Jiang, Z. 2009. "Let's Shop Online Together: An Empirical Investigation of Collaborative Online Shopping Support," *Information Systems Research*, (21:4), pp. 872–891 (doi: 10.1287/isre.1080.0218).

References for Essay 3 (Chapter 4)

Akerlof, George A. 1970. The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84 (3), 488–500.

Ba, Sulin and Paul A. Pavlou. 2002. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly* 26 (3), 243–269.

Bolton, G., B. Greiner, A. Ockenfels. 2009. Engineering trust: Reciprocity in the production of reputation information. Working Paper, Harvard Business School.

Bryant, W.K. and J. Gerner. 1978. The price of a warranty: The case for refrigerators. *Journal of Consumer Affairs* 2, 30-47.

Cabral, L. and A. Hortacsu. 2010. The dynamics of seller reputation: Evidence from eBay. *The Journal of Industrial Economics* 58 (1), 54-78.

Chu, Junhong and Pradeep K. Chintagunta, 2011. An empirical test of warranty theories in the U.S. computer server and automobile markets. *Journal of Marketing* 75, 75-92

Cooper, R. and T.W. Ross, 1985. Product warranties and double moral hazard. *Rand Journal of Economics* 16, 103-113.

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

Dewally, Michael and Louis Ederington. 2006. Reputation, certification, warranties, and information as remedies for seller-buyer information asymmetries: Lessons from the online comic book market. *Journal of Business* 79 (2), 693-729.

Fan, Ming, Yong Tan, and Andrew B. Whinston. Evaluation and design of online cooperative feedback mechanisms for reputation management. *IEEE Transactions on Knowledge and Data Engineering*, 17 (2), 244-254.

Gomes, Lee. 2006. Talking tech: How sellers find ways to trick eBay 'reputation' rating system. Wall Street Journal (December 19).

Heal, Geoffrey. 1977. Guarantees and risk-sharing. *Review of Economic Studies*, 44 (3), 549–60.

Jin, G. and A. Kato, 2006, Price, Quality and Reputation: Evidence from an Online Field Experiment, *Rand Journal of Economics*, 37 (4), 983–1005.

Kelley, Craig A. 1988. An investigation of consumer product warranties as market signals of product reliability. *Journal of the Academy of Marketing Science* 16 (2), 72-78.

Kim, Y., and Krishnan, R. 2015. "On Product-Level Uncertainty and Online Purchase Behavior: An Empirical Analysis," Management Science, (61:10), pp. 2449–2467 (doi: 10.1287/mnsc.2014.2063).

Li, S., K. Srinivasan and B. Sun, 2009. Internet auction features as quality signals. *Journal of Marketing* 73 (1), 75-92.

Lutz, N. 1989. Warranties as signals under consumer moral hazard. *Rand Journal of Economic*s 20 (2), 239-255.

Priest, G.L. 1981. A theory of the consumer product warranty. *Yale Law Journal* 90 (6), 1297-1352.

Resnick, Paul, Richard Zeckhauser, John Swanson, and Kate Lockwood. 2006. The value of reputation on eBay: A controlled experiment. *Experimental Economics* 9 (2), 79-101.

Rob, R. and A. Fishman. 2005. Is bigger better? Customer base expansion through word-of-mouth reputation. *Journal of Political Economy* 113 (5), 1146-1162.

Roberts, James. 2011. Can warranties substitute for reputations? *American Economic Journal: Microeconomics* 3 (3), 69–85.

Spence, A.M. 1977. Consumer misperceptions, product failure, and product liability. *Review of Economic Studies* 44, 561-572.

Tadelis, S., 2002. The market for reputations as an incentive mechanism. *Journal of Political Economy* 110 (4), 854–882.

Wiener, Joshua L. 1985. Are warranties accurate signals of product reliability? *Journal of Consumer Research* 12 (2) 245-250.

Xu, Hong, Jianqing Chen, and Andrew B. Whinston. 2008. Audited reputation. *Economics Letters* 100 (3), 359-362.

LIST OF TABLES

TABLE 1
Main Variable Descriptions

Variable	Variable Definition
Traffic	Total number of unique visitors that visited any product page of a seller website in a given month.
Reputation	Reputation score accumulated through ratings over time.
Product Variety	Total number of different products listed by a seller within a month.
Average Price	The average price of all the products within a month.
Return Customer	The number of customers who have made purchases with the seller more than one time within past six months.
Sponsored	Binary variable indicating whether sellers have adopted sponsored search advertising or not in a given month; sponsored=1 if sponsored search is used; otherwise, sponsored=0.
Social	Binary variable indicating whether sellers have adopted social media endorsement or not in a given month; social=1 if social media endorsement is used; otherwise, social=0.
East	A dummy variable indicating whether the seller is located in the East region of China; East=1 if a seller is located in the East region; otherwise, East=0.
West	A dummy variable indicating whether the seller is located in the West region of China; West=1 if a seller is located in the West region; otherwise, West=0.
Central	A dummy variable indicating whether the seller is located in the central region of China. Central=1 if a seller is located in the central region; otherwise, Central=0.

TABLE 2
Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Traffic	15,857.25	64,925.05	0	2,389,328
Reputation	8,820.69	49,470.78	0	2,218,195
Product Variety	299.79	719.81	1	90,787
Average Price	154.30	193.19	0.733	6,648.57
Return Customer	146.97	802.29	0	30,182
Social	0.46	0.50	0	1
Sponsored	0.56	0.50	0	1
East	0.87	0.34	0	1
West	0.04	0.19	0	1
Central	0.10	0.29	0	1

Note: Total number of observation N= 28619

TABLE 3
Probit Estimates for Selection Model

Variable	*Sponsored*	*Social*
Reputation	1.111***	0.618***
	(0.018)	(0.016)
Product Variety	-0.298***	0.031
	(0.021)	(0.020)
Average Price	0.489***	0.029
	(.028)	(0.026)
East	0.287***	0.211**
	(0.088)	(0.083)
West	-0.289	0.520***
	(0.159)	(0.147)
Constant	-8.484***	-5.243***
	(0.215)	(0.194)

Note: Total number of observation N= 28619
All the continuous variables are log-transformed
Standard error in parentheses
**p < 0.10, **p < 0.05, ***p < 0.01*

TABLE 4
Traffic Equation Estimates for Advertising Strategies

Variable	Base Model	Interaction Model (I)	Interaction Model (II)
Reputation	0.049***	0.041***	0.376***
	(0.008)	(0.008)	(0.011)
Product Variety	0.191***	0.188***	0.221***
	(0.008)	(0.008)	(0.008)
Return Customer	0.533***	0.537***	0.470***
	(0.007)	(0.007)	(0.013)
Sponsored	1.378***	1.543***	0.996***
	(0.027)	(0.031)	(0.086)
Social	1.058***	1.326***	-0.862***
	(0.024)	(0.035)	(0.080)
Sponsored Social*		-0.399***	
		(0.037)	
*Sponsored*Reputation*			-0.286***
			(0.014)
*Social*Reputation*			0.024*
			(0.013)
*Sponsored*Return Customer*			0.123***
			(0.015)
*Social*Return Customer*			0.001
			(0.013)
Constant	3.691***	3.677***	3.344***
	(0.057)	(0.057)	(0.061)
ϑ_i	-0.343***	-0.342***	0.365***
	(0.009)	(0.009)	(0.009)
π	0.600***	0.598***	-0.463***
	(0.014)	(0.014)	(0.014)
lnalpha	-0.683***		
	(0.021)		

Note: Total number of observation N= 28619
All the continuous variables are log-transformed
The model also contains lag term of traffic variable
Standard error in parentheses
**p < 0.10, **p < 0.05, ***p < 0.01*

TABLE 5
Variable Definition

Paid traffic	Total number of unique web visitors that made purchases
Total traffic	Total number of web visitors
Conversion	Paid Traffic divided by total traffic
Response rate	Average live chat response rate for each seller
Reputation	Reputation score accumulated through ratings over time
New product	Total number of new products launched within a month
Price	The average price level of all the products
Repeat customer	The number of customers who have made more than one purchase with the seller during the past 6 month.
Return ratio	Percentage of products being returned
Rating	Dynamic product rating, moving average of past 6 month.
Picture	Total number of product pictures posted on product pages
Review	Total number of monthly review seller received

TABLE 6
Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Total Traffic	19170	36539.06	45	991000
Paid Traffic	311.9	587.56	0	14500
Conversion	0.02	0.02	0	0.80
Reputation	10,164.56	17,773.66	139	347041
New Product	74.19	211.21	0	9829
Price	170.76	162.41	6.65	2638
Repeat Customer	148.02	222.18	0	1976
Return Ratio	0.00	0.01	0	1
Response Rate	0.57	0.26	0.002	0.99
Rating	4.76	0.17	1	5
Picture	17.65	15.05	0	192.97
Text	501.62	435.52	0	3122.59
Review	441.15	806.42	0	15836

Note: Number of observations (10-month period): 22,220

TABLE 7
FIRST STAGE PARAMETER ESTIMATION

Reputation$^{(L)}$	0.28***
	(0.01)
New product$^{(L)}$	0.10***
	(0.00)
Price$^{(L)}$	-0.06***
	(0.01)
Repeat customer$^{(L)}$	-0.02***
	(0.00)
Return ratio	-0.39***
	(0.10)
CONSTANT	-1.56***
	(0.12)
$\boldsymbol{\sigma_u}$	0.29
$\boldsymbol{\sigma_\varepsilon}$	0.23

Note: Total number of observation N= 22220
P<0.1, P < 0.05**, P < 0.01****
$^{(L)}$The natural logarithm form of the variable.

TABLE 8
SECOND STAGE PARAMETER ESTIMATION

Panel A: Posterior Distribution of $\boldsymbol{\beta}$

	Mean (s.e)
Intercept	-4.81*** (0.02)
Response rate	0.42*** (0.03)
Conversion_{t-1}	27.50*** (0.55)
σ_α	0.23*** (0.01)
σ_ζ	0.35*** (0.00)

Panel B: Posterior Distribution of Δ

	intercept	Picture$^{(L)}$	Rating	Price$^{(L)}$	Review$^{(L)}$
Response rate	0.23*** (0.03)	-0.06** (0.02)	1.04*** (0.14)	-0.30*** (0.02)	0.08*** (0.01)
Conversion_{t-1}	27.23*** (0.56)	0.68 (0.56)	-7.04 (3.21)	2.60*** (0.63)	0.70 (0.34)

Note: Total number of observation N= 22220
P<0.1, P < 0.05**, P < 0.01****
(L) The natural logarithm form of the variable.

TABLE 9

Variable Definitions of Essay 3

Variable	Definition
Guarantee	A dummy variable indicates whether a seller uses the RGP or not. It equals 1 when a seller uses the program and 0 otherwise.
Reputation	Cumulative reputation score of a seller.
DSR	Average dynamic seller rating for a given seller in a month
Traffic	Total incoming traffic for a seller in a given month
Average Price	Average price of all the products a seller carries.
Product Variety	Total number of products a seller carries during the month.
Bookmark	Number of times the seller has been favorited by consumers
Paid Advertising	A dummy variable indicates whether a seller uses paid search advertising or not. It equals 1 when a seller adopts it and 0 otherwise.
Social Endorsement	A dummy variable indicates whether a seller uses social endorsement or not. It equals 1 when a seller adopts it and 0 otherwise.
CAPV	Customer average page views for a given seller in a given month
AOV	Customer average order value for a given seller in a given month
Conversion	Total conversion rate during a month for a given seller
%PositiveReview	Percentage of positive reviews the seller has received in a given month

TABLE 10

Descriptive Statistics of Essay 3

	Mean	Std. Dev.	Minimum	Maximum
Guarantee	0.27	0.44	0	1
Traffic	2033.88	11577.77	0	413,844
Reputation	835.69	3257.55	1	59,868
Average Price	135.81	164.81	0.73	3,810
Product Variety	184.62	390.81	1	12,526
Bookmark	63.39	868.80	0	37,867
Paid Advertising	0.09	0.29	0	1
Social Endorsement	0.10	0.30	0	1
CAPV	2.73	10.75	0	492.93
AOV	173.20	358.07	0	6724
Conversion	0.038	0.084	0	1
DSR	3.62	2.15	0	5
%PositiveReview	0.96	0.19	0	1

Note: $N = 3,548$

TABLE 11
Logit Regression Model of Sellers' Participation Of Buyer Protection Program

$\boldsymbol{Reputation}^{(\boldsymbol{L})}$	-0.090***
	(0.02)
$\boldsymbol{ProductVariety}^{(\boldsymbol{L})}$	-0.043
	(0.032)
$\boldsymbol{AvgPrice}^{(\boldsymbol{L})}$	0.080*
	(0.04)
$\boldsymbol{Bookmarked}^{(\boldsymbol{L})}$	0.283***
	(0.03)
SponsoredSearch	0.828***
	(0.14)
SocialEndosement	0.567***
	(0.13)
$\boldsymbol{Traffic}^{(\boldsymbol{L})}$	0.050*
	(0.03)
Constant	-1.65***
	(0.23)

Note: Total number of observation N= 4910
P<0.1, P < 0.05**, P < 0.01****
(L) The natural logarithm form of the variable.

Table 12
Difference in Difference: Customer Average Page View of Treatment Vs. Control Groups

	Treatment	Control	Difference	t-test
Before Participation	1.217	1.184	0.034	1.03
After Participation	1.178	1.107	0.071	3.26
Difference	-0.039	-0.076	0.036	1.07

Table 13

Difference in Difference: Average Order Value of Treatment Vs. Control Groups

	Treatment	Control	Difference	t-test
Before Participation	3.420	3.989	-0.569	-3.77
After Participation	4.749	3.478	1.270	9.71
Difference	1.328	-0.511	1.83	10.19

Table 14

Difference in Difference: Conversion of Treatment Vs. Control Groups

	Treatment	Control	Difference	t-test
Before Participation	0.035	0.045	-0.010	-1.81
After Participation	0.045	0.042	-0.003	0.52
Difference	0.010	-0.003	0.014	1.94

Table 15

Difference in Difference: Dynamic Seller Rating of Treatment Vs. Control Groups

	Treatment	Control	Difference	t-test
Before Participation	3.100	4.002	-0.902	-6.45
After Participation	4.657	3.261	1.396	11.86
Difference	1.557	-0.741	2.298	11.25

Table 16

Difference in Difference: Percentage Positive Review of Treatment Vs. Control Groups

	Treatment	Control	Difference	t-test
Before Participation	0.864	0.950	-0.087	-4.70
After Participation	0.987	0.988	-0.001	-0.17
Difference	0.123	0.038	0.086	4.86

LIST OF FIGURES

Figure 1
Illustration of Sponsored Search

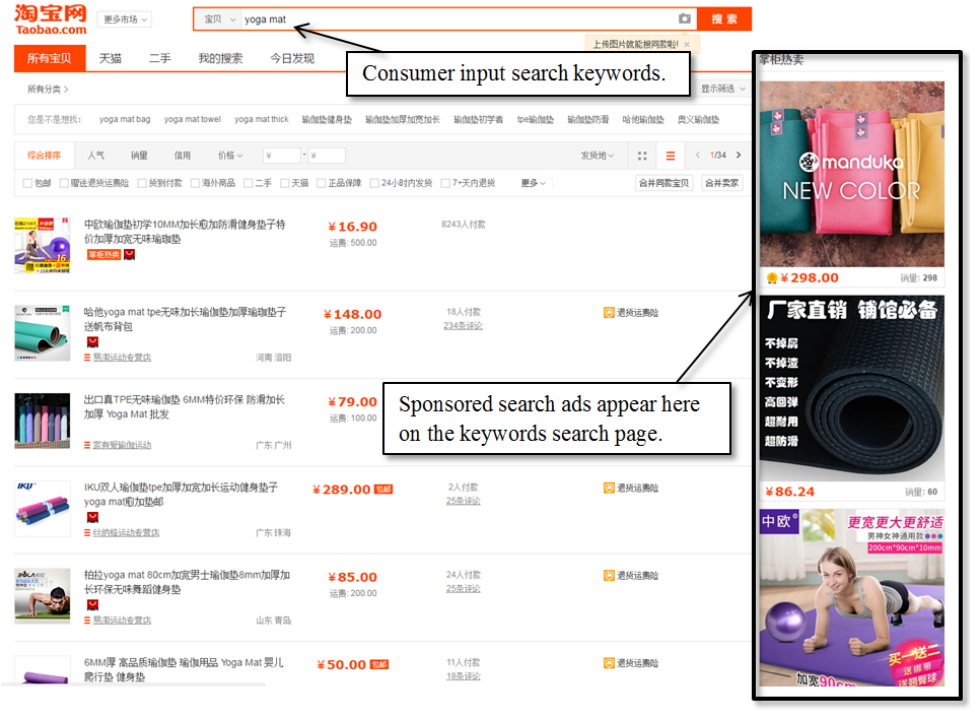

Figure 2
Illustration of Social Media Endorsement

Figure 3
Distribution of Online Traffic

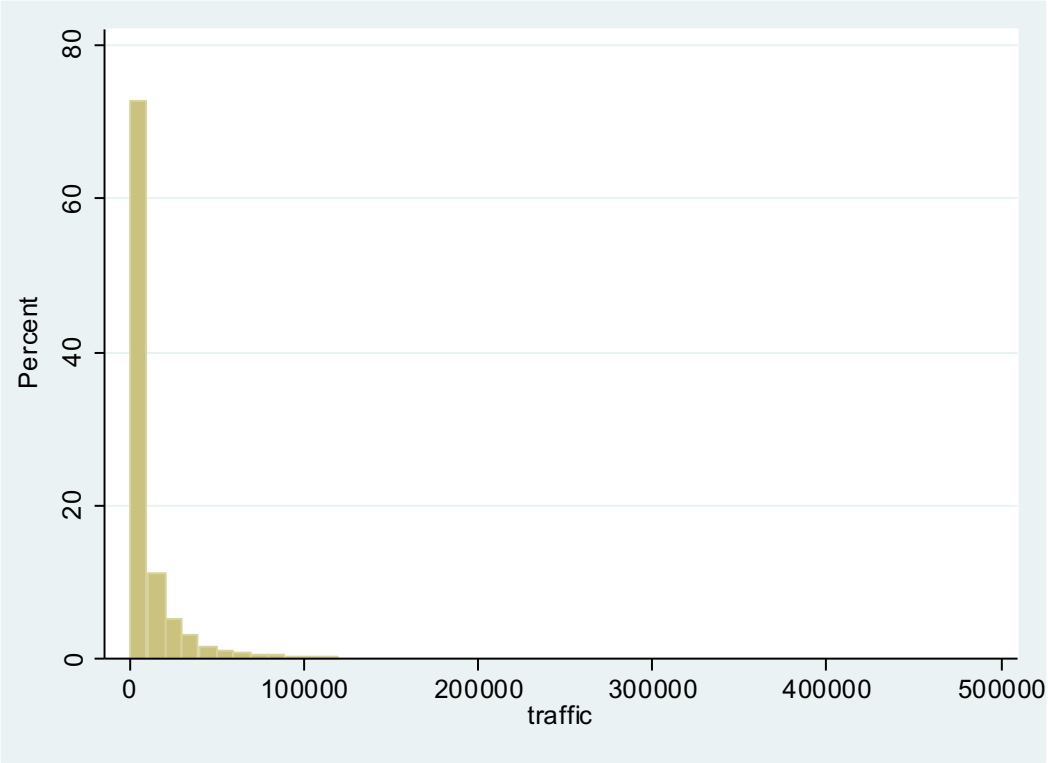

Figure 4

Distribution of Reputation by Strategies

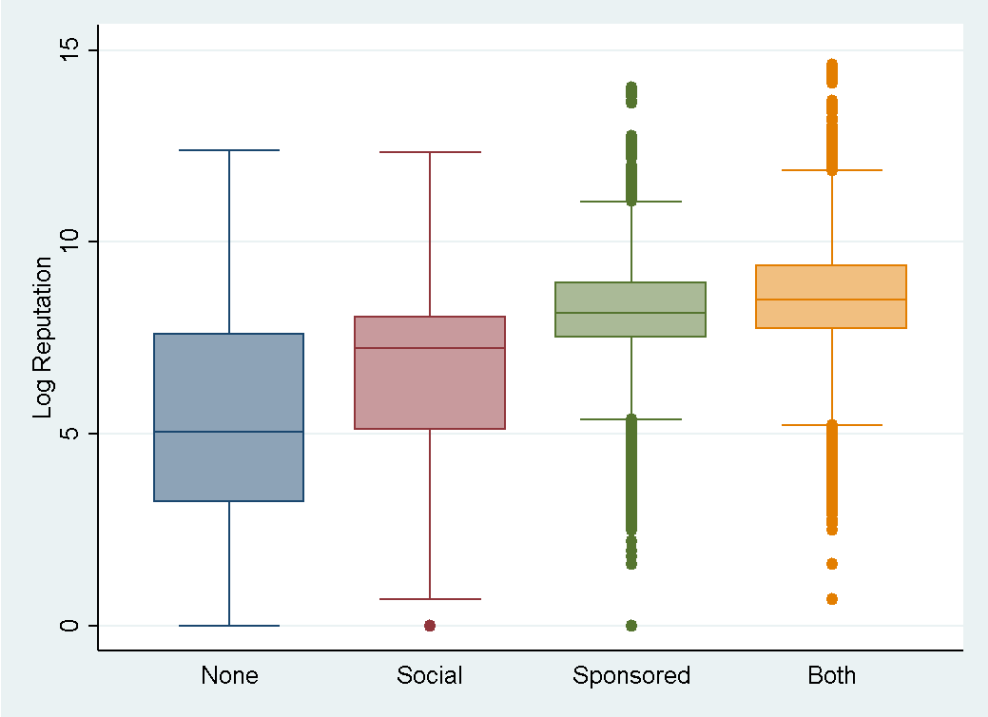

Figure 5
Distribution of Traffic by Strategies

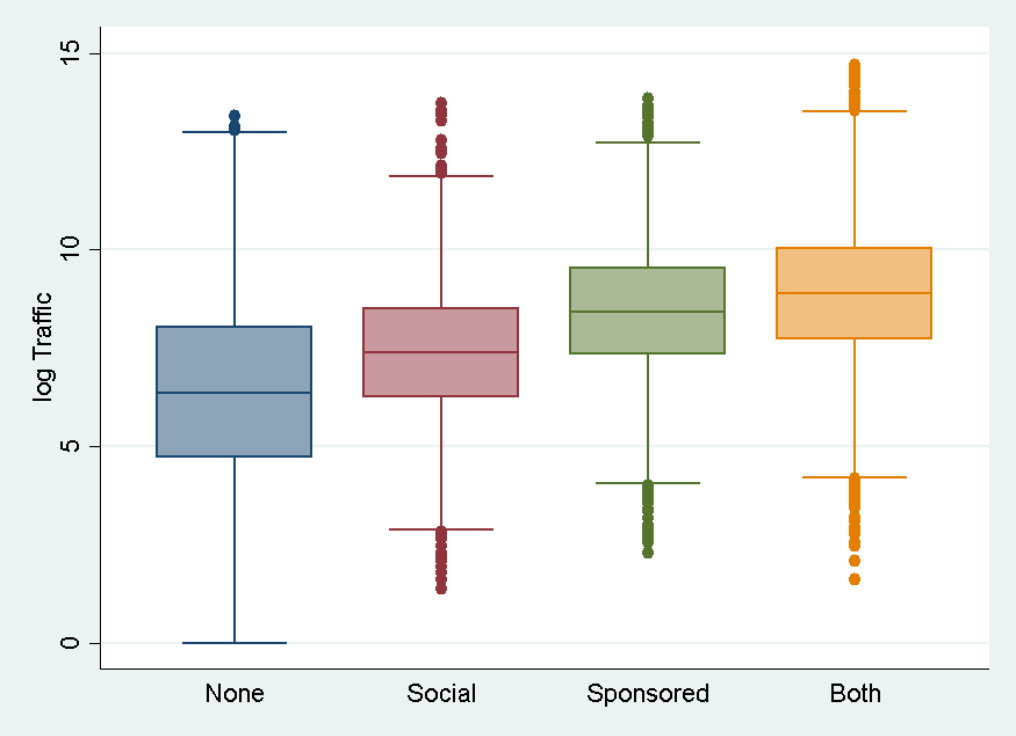

Figure 6

Distribution of Propensity Score Before and After Matching