Social Networks and Restaurant Choice

Timothy J. Richards, Ph.D., Arizona State University
Ashutosh Tiwari, M.S., American Express

Selected Paper prepared for presentation at the 2014 AAEA / EAAE / CAES Joint Symposium "Social Networks, Social Media, and the Economics of Food" Montreal, Quebec, Canada, May 28-30, 2014.*

April 4, 2014

1 Introduction

Food purchased at restaurants constitutes a major share of the household food budget, and yet we know very little about why some restaurants succeed, and others fail. Fully 26 % of the restaurants fail within the first year, and around 50 % within the first three years (Parsa et al. 2005). Restaurants seem to be either extremely successful or they struggle to survive, which suggests that there is some form of non-linearity or bandwagon effect driving restaurant demand (Becker 1991). Banerjee (1992), Cai, Chen, and Fang (2009), and Anderson and Magruder (2012) each find that diners rely on information derived from social networks to inform restaurant choices. Social learning, in turn, implies a "social multiplier" effect that would explain the observed bi-modal nature of restaurant success (Manski 1993, 2000). In this study, we use experimental methods to test for social learning effects in a restaurant environment.

Restaurant meals embody multiple attributes, many of which are either experience or credence attributes in the sense of Nelson (1974). As such, consumers face considerable a priori uncertainty in choosing where to go. Uncertainty concerns not only the food offered in the restaurant, but the overall dining experience as restaurant meals are archetypical multi-attribute experiences. Attributes such as food taste, food quality, ambiance, service quality, location of the restaurant, menu choices and price, all contribute to the overall dining experience. Diners face uncertainty when they have limited or no prior experience when choosing among available restaurants. To resolve this uncertainty, diners seek various sources of information, which include both marketer-controlled and marketer-uncontrolled sources.

Marketer-uncontrolled sources such as word-of-mouth (WOM) are generally more credible and influential than marketer-controlled sources such as paid advertising (Buttle 1998; Mangold et al. 1999; Buda and Zhang 2000). It is well-understood that word of mouth (WOM) has a strong effect on consumer decision making process (Herr, Kardes and Kim 1991; Maxham 2001; De Bruyn and Lilien 2008), but traditional WOM takes place in small social groups and the conversations are ephemeral (Hu and Li 2011). In the last decade, increasing user-based online interaction has eliminated some of the limitations of traditional peer-to-peer communication, and yet has created a sharper distinction between WOM in peer and anonymous networks.

There are two categories of online social networks: peer networks and anonymous networks. In peer networks every member is connected to other members by a primary connection (friend), secondary connection (friend's friend) or tertiary connection (secondary friend's friend) and so on. Watts and Strogatz (1998) show that there is a maximum of six degrees of separation in any peer network – a phenomenon known as the "small world" effect. Examples of online peer networks are Facebook, Linkedin, Twitter and Instagram. Anonymous networks consist of online communities, where members are past users of different products services, who share their experiences with other members. Yelp, Tripadvisor and Citiguide are examples of few popular anonymous networks. In this study, we compare the relative effect of each type of WOM in driving the demand for restaurants. More generally, we study the role of both anonymous social media and social peer networks in shaping trends in restaurant demand, which often portend more general changes in food consumption.

Peer and anonymous WOM differ in several important ways. While peer networks have a trust advantage over anonymous networks (Hilligoss and Rieh 2008), anonymous networks include a far deeper well of knowledge, and different perspectives that may be valuable for potential customers (Cheung and Lee 2012). Web-based interaction or electronic word of mouth (e-WOM) can take place among distant individuals and, more importantly, does not require individuals to send and receive messages at the same time. Moreover, in most cases the messages are stored in the medium and available for a future reference (Bhatnagar and Ghose 2004; Godes and Mayzlin 2004; Duan, Gu and Whinston 2008). At the same time, consumers rely on peer networks for similar information on services they may have limited experience with. With the rise of web 2.0 technology in the last decade and its two way interactive power, online social networking is a ubiquitous phenomenon. While peer social networking websites such as Facebook.com, Twitter.com, Myspace.com, and Instagram.com enable customers to obtain feedback and recommendations for products and services based on peer user experiences, anonymous networking websites such as Yelp.com, Traveladvisor.com and CitiGuide.com use customer reviews to disseminate e-WOM. Which category of social networks, anonymous or peer, is more effective in increasing demand, therefore, is an empirical question.

Empirical social learning effects are well-documented in investment decisions (Hong, Kubic and Stein 2004), new product purchase (Mayzlin 2006; Godes and Mayzlin 2004, 2009) and retirement plan participation (Duflo and Saez 2002, 2003). Reviews and recommendations from members of a consumer's peer network have a strong impact on choice (Narayan, Rao and Sanders 2011; Cai, Chen and Fang 2009; Trusov, Bodapati and Bucklin 2010). These studies, however, focus on peer networking and not anonymous social networks. Luca (2011) and Anderson and Magruder (2012), on the other hand, show that positive ratings from anonymous Yelp reviewers can raise the apparent demand for restaurants, but neither compare the value of anonymous and peer networks to consumers and, thereby, to restaurant owners. Both of these studies also focus on the Yelp star-rating system for their identification strategy, while we use the nature of the review itself.[1] We aim to compare the relative effect of each type of social network on demand, and quantify the importance of each in driving restaurant success or failure.

The lack of research comparing peer and anonymous social networks is primarily due to a lack of data. While this observation seems paradoxical, given the ubiquity of each, the fact that each represents a fundamentally different concept of social learning means that there is no source of revealed-demand data from both. Therefore, we conduct an economic experiment to compare the effectiveness of anonymous versus peer networks as tools for marketing restaurant meals. We directly compare the impact of publicly available user

[1] Our implicit assumption, relative to these other studies, is that Yelp users actually read reviews and do not simply rely on star-rating values to evaluate prospective restaurants.

reviews from a customer review website (Yelp) to that of peer reviews on restaurant demand.

In any empirical model of social learning, identification is always an issue because the individual is also part of the group. Manski (1993) describes this as the "reflection problem": How can a researcher infer the effect of the group behavior on the behavior of an individual, when the individual contributes to some of the observed group behavior? When behavioral effects of a peer group on an individual, who is a peer member of that group herself are modeled, the results obtained are biased. Reflection is best mitigated through appropriate controlled experimental design, which generates rich data and hence mitigates the reflection problem. We conduct a two-stage group-subgroup experiment under strict uniform network size restrictions to tackle the identification problems associated with social learning. We randomly assign members of each peer-group into sub-groups and do not allow peers to decide their subgroup. Such random assignment ensures that peers do not choose subgroups of similar preferences and thus correlation between observed peer attributes and the error term in the restaurant choice regression equation is limited by design. Recommendations based on restaurant visits in the first-stage by one sub-group are given to members of the other sub-group prior to visiting same restaurant. We then aggregate the data during econometric estimation to incorporate group level heterogeneity in a manner similar to Georgi et.al (2007) and Bramoullé et.al (2009). By dividing each peer group into two sub groups, we avoid the reflection problem.

Other than the reflection problem, peer networks are typical to have endogeneity problems. Manski (1993) formed three hypotheses for peers as why they behave in similar fashion: (1) endogenous effects, which explain existence of a herd behavior, in that peers behave as other members in the peer group, (2) contextual effects, which are similarities with respect to the exogenous factors such as similar demographics or psychographics within a peer group, and (3) correlated effects, which are similar environmental factors under which peers within a network reflect similar behavior. Brock and Durlauf (2002, 2007) demonstrate that peer effects are identified in a discrete choice model, even in the presence of correlated effects with binary or multinomial choice models. In this paper, we use an ordered probit model to estimate the importance of network effects in restaurant preference because demand is expressed in terms of a five-point rating scale measuring whether the consumer would visit the restaurant again. We create full-information adjacency matrices for each group that gives us complete information about how well a peer knows other members in the same network. Using this information, along with individual demographic and behavioral attributes, we are able to identify peer effects at individual level (Bramoullé, Djebbari and Fortin 2009). Further, we address endogeneity of network membership through an instrumental-variables estimation approach. Specifically, we estimate social learning effects with an ordered probit model, estimated using a control function approach (Park and Gupta 2009; Petrin and Train 2010). While our experimental design allows us to exclude the individual from the group for whom we want to test the peer effect, the control function modelling approach helps me handle endogeneity problem. Combining these two features mean that our experimental design and modelling approach is both unique and appropriate to study peer effects.

In addition to the importance of each source of social learning, the relative strength of positive and negative reviews is also of some question. Chevalier and Mayzlin (2006) found some evidence that negative reviews have more powerful impact than that of positive reviews in case of book reviews using secondary online data. Asymmetric responses to positive and negative information is a natural implication of Prospect Theory (Kahneman and Tversky 1979) wherein consumers are more likely to respond to stimulii that have negative consequences (in the "domain of losses") than to stimulii that have positive consequences (in the "domain of gains"). We test for this manifestation of Prospect Theory in our experimental social network data by comparing the strength of subjects' responses to positive and negative restaurant ratings by both peers and reviewers on anonymous social networks.

We find that information obtained from peer networks has a stronger influence on restaurant choice than information derived from anonymous social media. Online rating websites, especially restaurant review websites such as Yelp, contain numerous reviews from the past users with detailed user stories and experience. Despite the large pool of specific information about particular restaurants, which likely helps decision makers, peer reviews are regarded as a more trustworthy source of information. In most cases, peer networks are generally smaller than anonymous networks, and a small proportion of peers provide information on a single restaurant due to different tastes and preferences within the peer group. Yet, our findings show that they likely have more influence on other peers in the network. We also take into account the fact that individuals in peer networks do not have equal influence on other members and, similar to Godes and Mayzlin (2009), the most interconnected member is not necessarily the most influential. Other than interconnectedness, the level of influence also depends upon other factors such as strength of connections (Weimann 1983) and communication frequency (Zenger and Lawrence 1989). We also find that negative reviews have a far stronger effect on preferences than do positive reviews. This finding is consistent with Prospect Theory, and suggests that preferences are shaped by a greater aversion to potential loss, than to potential gain.

Our research has both managerial importance and a more general contribution in providing a better understanding how social media effects demand. By understanding the relative role of peer and social networks, restaurant managers may be able to avoid the boom-or-bust dynamic typical of startups in the foodservice industry. The research may also help guide foodservice managers to develop effective online social media marketing strategy and helps optimize their marketing budget. More generally, we identify the relative importance of online rating sites to peer networking sites. To the extent that firms in other industries share the same type of uncertainty faced by restaurant owners, our findings are suggestive of how social media strategies may be designed for maximum effect. Finding that consumers are particularly sensitive to the possibility of a utility-loss as opposed to a utility-gain suggests that managers should prioritize consistency in each attribute that restaurant users find to be important, rather than excellence in some, and mediocrity in others. Doing so minimizes the likelihood that a potential reviewer has a negative experience, and communicates this fact to either her friends or the broader public.

The next section describes a conceptual model that we use to formulate hypotheses that follow from the theory of social learning through peer and anonymous networks. In the third section, we explain the social dining experiment design and execution, while fourth section describes a spatial econometric model that accounts for the unique nature of social learning. We present the results from estimating the econometric model in section five, and conduct a number of specification tests to establish the validity of our approach. Section six summarizes our findings, and suggests some limitations.

2 Economic Model of Social Network Effects

Restaurant offerings are fundamentally different than other service offerings as they have an aesthetic and emotional component to them (Johns 1999). Restaurants face a challenge to offer variety on their menus and at the same time to standardize the experience for the same menu item over time. The restaurant market in the U.S is mature, having developed in response to diverse consumer preferences for multiple dining options (Mack et al. 2000). While the fast food market offers more standardized products and services, fine-dining restaurant offerings are generally more complex, with each restaurant offering a unique combination of various desired dining attributes. Other than the attributes of food served, factors such as customers' sense of style, ambiance and service play vital role in diners' decision making process (Muller and Woods 1994). The combination of all these observable and unobservable factors makes restaurant offering a complex product, combining both goods and service attributes. This complexity implies a high degree of uncertainty with respect to quality, or the general level of satisfaction with the experience. Consumers resolve this uncertainty by obtaining information. Among the various sources of information available, word of mouth (WOM) is particularly important.

Individuals do not live in isolation but are part of various social communities. Members of these communities interact during social gatherings, formal or informal meetings, social events or even day to day unplanned encounters. These interactions induce a two-way flow of information exchange. When this information is particular to any product or service, it is commonly spread through WOM. Consumers are more receptive to WOM from members of their social networks than other marketer controlled sources of information such as advertisements and promotions (Goldenberg et.al 2009 and Domingos 2005). When a consumer dines at a restaurant and then shares her experience with other members in the social network, perceptions of, and preferences for, the restaurant can be affected either positively or negatively (Chevalier and Mayzlin 2006; Nam, Manchanda, and Chintagunta 2010).

The power of WOM depends on both the nature of the message and its source. Social network members who disseminate WOM can be either known peers or unknown experts, and each has a different source of credibility, and influence. Peers are more highly trusted because individuals have already formed a bond, whether through kinship, friendship, shared-interest, or common values revealed through, for example, membership in a church group, political party, or service organization (Hilligoss and Rieh 2008). The existence of a trust relationship, however, does not necessarily imply subjective preferences are the same, or that trust in a general sense extends to the value of information in an area that requires specific expertise. When knowledge in a certain area is required, then consumers are likely to reach out beyond their network of peers to find expertise from anonymous sources (Cheung and Lee 2012). For example, if an individual seeks information on weight-loss programs, and all of his or her peers are also overweight, then even trusted peers may not be regarded as the best source of information on how to lose weight. Consequently, because the trust and expertise attributes of peer and anonymous networks are equally credible, we cannot form a specific hypothesis as to which is likely to be stronger, and leave the result as an empirical question.

No matter the source, there are two dimensions of WOM effects: Magnitude and direction. The magnitude of the WOM effect will depend upon influential power and information dissemination power of the source within the network, and the strength of his or her connection with other members (Dierkes, Bichler and Krishnan 2011). The direction of WOM, however, depends on the nature of the message. It is intuitive that positive WOM will have a positive effect and negative WOM will have a negative effect on demand, but whether the effect is asymmetric is an empirical question. Despite the evidence in the affirmative provided by Chevalier and Mayzlin (2006), they do not offer a theoretical explanation as to why this might be the case. Richards and Patterson (1999) found that negative media reports have a greater effect on prices than positive reports after a foodborne disease outbreak in strawberries, and explain this result using Attribution Theory (Mizerski 1982) and its implications for the shape of a consumer's utility function. However, if fabricated reviews remain a small proportion of online reviews (Ong 2012), then there is no reason to believe that positive reviews are any less credible than negative reviews from either peers or anonymous reviewers. On the other hand, Prospect Theory (Kahneman and Tversky 1979) maintains that consumers evaluation the utility implications of uncertain bets according to their own, subjective frame of reference. Marginal changes in utility above a reference point, or when the consumer is in the domain of gains, are smaller than the negative changes in utility that lie below the reference point, or when the consumer is in the domain of losses. Therefore, according to Prospect Theory, we expect the absolute value the response to negative ratings from either peers or anonymous reviewers to be larger than the absolute value of the response to positive ratings. More formally, if utility is concave (increases at a decreasing rate) in the nature and amount of information received, then the response is likely to be asymmetric with negative information providing a larger negative impact than positive information provides in the opposite direction. According to this theory, the incremental loss of utility by receiving negative reviews is greater than the gain in utility by receiving positive reviews. We test for the implications of Prospect Theory using the experimental responses to anonymous reviews, simply because we are better able to control the nature of the message than from ones' peers. We do so using an econometric model designed to test for heterogeneity in subjects' responses to positive and negative reviews.

Identifying the relative importance of peers versus anonymous reviews, or estimating the relative magnitude of positive versus negative reviews is inherently confounded by the perceived quality of each source of information. Clearly, within each social network, not every member has an equal effect on consumer choice. In peer networks, those who have strong connections, who frequently communicate with the consumer, and who are central to the consumer's network likely have more influence than their counterparts (Weimann 1983; Zenger and Lawrence 1989; Ibarra and Andrews 1993). However, the most connected member is not necessarily the most influential as network members vary in their individual persuasiveness (Goldenberg et al. 2009). Similarly, reviewers with more experience and more followers are likely to be viewed as more credible, so may be more influential. In the absence of specific data describing the expertise of each potential reviewer, our econometric model below is designed to control for the quality of information as much as possible.

In summary, the experiment is designed to examine two empirical questions. First, we test whether peer WOM is more influential than WOM from anonymous networks in restaurant choice. Second, we test the hypothesis that negative reviews from anonymous sources have greater marginal impact than positive reviews. In the next section, we describe the two stage social dining experiment used to provide the data necessary to address these questions.

3 The Social Dining Experiment

We design a two-stage experiment that is intended to compare the influential power of anonymous and peer networks, and the relative effect of positive and negative reviews in the case of restaurant choice. In general, two-stage models are necessary to properly identify the influence of information (Urberg et al. 2003). Our experiment allows for a direct comparison of the influence of information from two different sources. In the first stage, restaurant reviews were provided to the subjects from anonymous sources, and in the second stage reviews were generated from known peers. The first-stage, anonymous reviews consisted of some positive and some negative reviews. In this way, we are able to both compare the effect of anonymous and peer influences, and test for any asymmetric response to positive and negative reviews.

Others demonstrate the value of using a two-stage structure to conduct social networking experiments. Narayan, Rao, and Saunders (2011) use a two-stage experiment to test for the effect of peer information on the choice of a technology product. Ours differs from theirs in three ways. First, they consider peer effects, while we compare peer and anonymous reviewer effects. Second, their interest is in preference revision, while ours lies in comparing two different sources of information. Narayan, Rao, and Saunders (2011) provide no information in the first stage, and information from peers in the second stage and compare the change in preferences that results. Instead, we use a control group to identify the relative effect of each information source at each stage, and within-subject variation to control for heterogeneity in response to either peers or anonymous reviewers. Third, we classify the information provided to respondents into positive and negative ratings in order to test for the existence of an asymmetric response to WOM.

Our experiment is designed to account for the reflection problem that is inherent in any social learning experiment (Manski 1993). The reflection problem raises the question of how to infer the influence of a group on an individual, when the individual is herself part of the group. Our experimental design helps avoid the problem of reflection as we randomly divide each peer group into two equal subgroups. We provide peer reviews from one subgroup to the other subgroup, and *vice versa*. This way, the controlled nature of our experiment helps avoid the identification problem persistent in social network experiments (Jackson 2008) because we control the assignment of information across sub-groups, and the subjects do not.

In the first stage, ten individuals were recruited to serve as "hubs." Hubs are individuals who were selected as organizational nodes for each network, but are not necessarily the most influential people in each network. The purpose of choosing hubs was to assemble a set of networks in which we can be assured that individuals know each other to varying degrees. That is, some members of the network organized by the hub may be best friends, while others may be only rare acquaintances. Each hub was asked to recruit a group of ten individuals. These ten groups formed independent peer networks (the groups are pre-selected to consist of ten individuals who know each other and are connected through a primary, secondary or tertiary connection). We then randomly divided each peer group into sub-groups of five members each: The "A" subgroup and the "B" subgroup. In the first stage, "A" subgroup members visited a restaurant in one suburb of a large, U.S. metropolitan area for lunch (rated 2.5 stars on Yelp). Five A subgroups were provided with positive Yelp ratings information and 5 A groups with negative Yelp ratings. At the same time, B subgroup members visited a similar type of restaurant in a different suburb of the same city for lunch (rated 4.0 stars on Yelp) following a similar procedure. We also recruited 37 individuals as control group members who visited both the restaurants separately in stage 1 and stage 2 without any prior reviews or information about the restaurants. These control group members were individually recruited following a random selection process in a popular shopping complex located near both restaurants. The final sample consisted of 136 respondents for each of two restaurants.

The positive and negative reviews were randomly selected from the set of all available Yelp reviews. Specifically, we collected all reviews available on Yelp.com for both restaurants and, after evaluating the reviews individually, we included for consideration only those that were clearly positive or clearly negative. Restaurant Two has more reviews available than Restaurant One, but the restaurants were chosen so that each had a sufficient number of reviews to yield at least five positive and at least five negative. We randomly compiled sets of five reviews for each positive and negative review types for both restaurants. We then randomly sent these sets of reviews to the corresponding positive and negative groups of respondents in stage one.

After visiting the assigned restaurant in round one, each respondent was asked to provide a rating (on a scale of $1-5$) on each of the following seven attributes of the restaurant experience: (1) taste of the food, (2) quality of the food, (3) availability of healthy menu choices, (4) ambience of the restaurant, (5) quality of the service, (6) price, and (7) ease of locating the restaurant. We measure restaurant preference by asking respondents to rate (on a scale of 1-5) their likelihood of revisiting each of the restaurants, and whether they would recommend a friend to visit each restaurant. At the end of the first stage, all subjects were asked to write a Yelp-style review about their dining experiences. These reviews are read and assigned a rating on a scale of 1 - 5 in a manner similar to the Yelp reviews for analysis purposes, but the entire review serves as a peer review for each member of the other subgroup in stage two.

We allowed approximately ten days for each round to be completed, and then a week to fill out the surveys. The data was collected using an online survey service, Network-Genie (https://secure.networkgenie.com). While the stage one survey had five sections: demographic information, behavioral information, network information, eating out preferences and stage one restaurant experience; stage two has only one section, gathering data on the nature of each respondent's restaurant experience. In the behavioral information section, we asked respondents about their online activity level, involvement with online social media (both anonymous and peers networking websites) and use of online social media as a product/service information tool. In the network information section, we asked respondents to rate all the peers in their network on a scale of 1 to 5 based on how well they do they know other members. This way, we obtained complete 10 x 10 social adjacency matrices for each peer network.

Members of a peer group are connected to each other through primary, secondary, or a tertiary connection. If a member is familiar with another member in the same group, their relationship is defined as a primary connection. If two members in the same network, A and B, don't know each other directly but have a common friend C, then the connection between A and B is deemed a secondary connection. Similarly, there can be various tiers of connections within a peer network. We allowed multiple connection tiers in recruiting the peer groups as network intransitivity may strongly affect the quality of the peer-effect estimates (Bramoullé, Djebbari and Fortin 2009). Following Bramoullé, Djebbari, and Fortin (2009), intransitivity is formally defined as the ratio of the number of intransitive triads over the number of triads. In a more general sense, networks are intransitive when there exist numerous levels of connections among peers. Due to the presence of intransitivity in peer networks, peer effects will likely include direct effects, generated from primary connections as well as indirect effects, generated from secondary, tertiary, or higher degree connections. Transitivity is a sufficient condition but not a necessary condition for indirect effects. To fully understand the peer network dynamics, it is important to understand direct as well as indirect network effects.

The resulting sample is broadly representative of the general population. The mean age of sample subjects is 37.27 years. Interestingly, 95 percent of the respondents have recommended a new restaurant to their peers, and 80 percent of all respondents have used online reviews in the past. Respondents with some college degree/trade school (non-bachelor and non-master degree) and bachelor's degree were the two most prevalent groups in the sample. Most respondents are in the middle of the income distribution as 69 percent of respondents had an annual income between $25,000 and $125,000. We summarize the experiment data in table 1.

[table 1 in here]

Summarizing the data by overall rating by restaurant provides some evidence that reviews have the expected impact on ratings for both restaurants. In table 2, the entries are interpreted as showing the mean rating across the seven attributes described above, for the control, negative review, and positive review groups, by round. If the entry is significantly greater than control, then the review has a positive effect on the reported restaurant rating, and vice versa. In general, positive reviews raise the rating level above the control group, and negative reviews are associated with ratings that are below control. There is, however, one exception in the case of negative reviews for the second restaurant as the expected negative effect is, in fact, positive (table 2). The data in this table provide summary evidence that support the influence anonymous social networks have on restaurant preference, but further econometric estimation is required to control for other, potentially confounding, factors (for example, demographic attributes, dining habits, or unobserved heterogeneity).

[table 2 in here]

A second summary provides evidence on the potential role played by peer preferences. In table 3, we report cross-tabulations of subjects' the likelihood of revisiting each restaurant (table 3) by stage and the type of review received. In the left three columns, we cross-tabulations refer to the type of Yelp review received in the first stage. Among those who received a review, a greater proportion said they would return if they received a positive review relative to those who received a negative review. In the right two columns, we show the summary results for those who did not receive a peer review relative to those who did. Recall that we did not classify the reviews received in the second round as we did for the first-round reviews because we cannot control experimentally for what constitutes a positive and negative review. We do find, however, that subjects who received a review reported that they were "very likely" to return to the restaurant more often than those who did not receive a review. This observation may suggest two potential behaviors. First, whether positive or negative, the mere fact of receiving a review provides some information, and resolves some of the choice uncertainty they may have had. Diners with more realistic expectations are less likely to be disappointed, which is a primary implication of Prospect Theory. Second, it may also be the case that not all peers have an equal effect on others in the network. Accounting for the extent of network interconnectedness using a formal spatial econometric model of social influence may help resolve this question.

[table 3 in here]

4 Econometric Model

In this section, we describe two econometric models. The first model compares the marginal impact of information obtained via anonymous networks on restaurant preference with the marginal effect of similar information obtained through peer networks. The objective of this model is to discover which network has a greater impact on restaurant choice. The second model tests whether negative reviews have a greater marginal impact (in absolute value) than positive reviews in determining restaurant preference. Because our measure of preference is an ordinal ranking metric, an ordered probit framework is used throughout. We explain each model in turn.

In the first model, we use proximity as a measure of network connectedness in a spatial econometric model of network influence. Proximity is a commonly used measure in the social network studies to identify the relative location of individuals in a social network (Freeman 1979; Hanneman and Riddle 2005; Opsahl, Agneessens and Skvoretz 2010). Although there are other measures of network location, such as betweenness and centrality, others find that estimates of network influence are not sensitive to these different definitions (Richards, Allender, and Hamilton 2013), so we focus on proximity.

4.1 Anonymous versus Peer Networks

There are two features of the experimental data that determine the appropriate econometric model. First, subjects provide an indication of their restaurant preference through a five-point, ordinal rating scale, where the rating measures their likelihood of returning to the restaurant in question. Because underlying preferences are assumed to be continuous, our data represent an ordinal manifestation of an underlying, unobserved, or latent utility variable. Therefore, an ordered probit model is necessary. Second, the data are likely to contain significant unobserved heterogeneity, or differences in preference that are due to purely idiosyncratic factors. In order to control for unobserved heterogeneity, we use a random coefficients version of the ordered probit model.

We define a subjects' assessment of the likelihood of revisiting each restaurant as an indicator of their strength of preference for each restaurant. Each subject was asked to rate the likelihood that they would revisit each restaurant on a 5-point Likert scale, each response reflecting their expected utility from revisiting the restaurant in question.[2] The two-stage experiment generated 274 observations over two rounds

[2] We use a 5-point Likert scale throughout the experiment as a 5-point Likert scale and a 7-point Likert scale provide comparable results (Dawes 2008), where 5 represents the maximum level of satisfaction and 1 represents the lowest level of

for 137 respondents, which is appropriate for a 5-point Likert scale (Hinkin 1995). Although this sample size may seem small relative to other studies that use primary data, it is important to keep in mind that social networking experiments necessarily involve some combination of small samples, or group-wise structures similar to ours. When subjects are asked to describe their relationship with each other member of the group or experiment, the task quickly becomes too difficult to complete accurately.[3]

An ordered probit model estimates the probability of moving from value of an ordinal variable to another. In the current context, the utility derived from visiting a particular restaurant is latent, or unobserved, but we do observe an ordinal-valued indicator variable that measures each subject's willingness to return to each restaurant (1 = not likely to 5 = very likely). We define the utility for subject i belonging to peer network j visiting restaurant k as u_{ijk}, or the vector $\mathbf{u}$. Instead of assuming that utility of individual i, $(i = 1, 2, \ldots .n)$ depends upon the mean rating provided by his or her peers in network j as in a more typical social-learning model (Duflo and Saez 2002, 2003), we follow a more general spatial approach. When using explicitly spatial methods to analyze social network models, it is possible to dyad-based information (relationships between two individuals) rather than simply a relationship between and individual's behavior and a peer-mean. Bramoulle, Djebbari, and Fortin (2009) argue that this approach also helps resolve the reflection problem as another's opinions are clearly separate from any calculation of a group-mean response. In this approach, utility (u_{ijk}) depends on the combination of information provided about the restaurant k by the peers in network j, and the strength of the relationship of individual i with other peers in the network j.

Restaurant reviews, whether by peers or anonymous reviewers, provide information that helps resolve some choice uncertainty and, thereby, raises utility. In order to test the relative effect of peer and anonymous reviews, we pool the data from both rounds and estimate a single model across both restaurants. We define the reviews provided by other members of network j as the vector $\mathbf{y}$ which is numerically coded on a 5-point Likert scale where 5 represents the highest and 1 represents the lowest rating. Each individual in network j, however, is assumed to be influenced by each other member of the network. Therefore, we weight each member's review by their location (or degree of connectivity) in the network. We use the NxN adjacency matrix, $\mathbf{G}$, for this purpose. Recall that the adjacency matrix is constructed using each individual's perception about his or her relationship strength with other peers in the network. If all members of a network know each other equally well, and have equal connection strength, our model becomes an average-peer effect model similar to Sacerdote (2001), and Duflo and Saez (2002, 2003). Utility also depends on the information provided by the anonymous reviewers for restaurant k, which we represent as the vector $\mathbf{x}$. We account for observed heterogeneity in utility by including a matrix (Nxm) of other subject-specific variables $(\mathbf{Z})$, and unobserved heterogeneity by allowing each of the information-response coefficients to vary randomly over

satisfaction.

[3] Narayan, Rao, and Saunders (2011) use a convenience sample of 70 MBA students.

sample subjects. With these assumptions, utility is written as:

$$\mathbf{u} = \beta_1 \mathbf{Gy} + \beta_2 \mathbf{x} + \gamma \mathbf{Z} + \boldsymbol{\varepsilon}, \tag{1}$$

where ε is an iid normal error term. We account for unobserved heterogeneity by assuming the β_k parameters are normally distributed such that: $\beta_k \tilde{} N(\bar{\beta_k}, \sigma^2_{\beta_k})$. In our model, unobserved heterogeneity derives from variations in tastes which are not accounted for observed differences among individuals. Define the observed restaurant-rating variable as q_m, which assumes a value of $m = 1, 2, ..., M-1$, only if a threshold utility level (α_m) is exceeded (Hausman, Lo, and MacKinlay 1992), or:

$$q_m = \left\{ \begin{array}{l} 1 \; if \; u \in (-\infty, \alpha_1], \\ 2 \; if \; u \in (\alpha_1, \alpha_2], \\ 3 \; if \; u \in (\alpha_2, \alpha_3], \\ 4 \; if \; u \in (\alpha_3, \alpha_4], \\ 5 \; if \; u \in (\alpha_4, \infty), \end{array} \right\} \tag{2}$$

where each of the α_i parameters are estimated from the experimental data. The probability of observing each rating value is, therefore, dependent upon others' ratings, anonymous reviewer ratings, and observed demographic variables according to: $\Pr(q_m = m | \mathbf{Gy}, \mathbf{x}, \mathbf{Z})$ where:

$$\Pr(q_m = m) = \left\{ \begin{array}{ll} \Pr(\beta_1 \mathbf{Gy} + \beta_2 \mathbf{x} + \gamma \mathbf{Z} + \boldsymbol{\varepsilon} \leqslant \boldsymbol{\alpha}_1), & if \; m = 1, \\ \Pr(\alpha_{m-1} < \beta_1 \mathbf{Gy} + \beta_2 \mathbf{x} + \gamma \mathbf{Z} + \boldsymbol{\varepsilon} \leqslant \boldsymbol{\alpha}_m), & if \; 1 < m < 5, \\ \Pr(\alpha_4 < \beta_1 \mathbf{Gy} + \beta_2 \mathbf{x} + \gamma \mathbf{Z} + \boldsymbol{\varepsilon}), & if \; m = 5. \end{array} \right\}$$

Assuming the error term, ε, is normally distributed, the ordered probit model is then written as:

$$\Pr(q_m = m) = \left\{ \begin{array}{ll} \Phi\left(\frac{\alpha_1 - \beta_1 \mathbf{Gy} - \beta_2 \mathbf{x} - \gamma \mathbf{Z} - \boldsymbol{\varepsilon}}{\sigma}\right), & if \; m = 1, \\ \Phi\left(\frac{\alpha_m - \beta_1 \mathbf{Gy} - \beta_2 \mathbf{x} - \gamma \mathbf{Z} - \boldsymbol{\varepsilon}}{\sigma}\right) - \Phi\left(\frac{\alpha_{m-1} - \beta_1 \mathbf{Gy} - \beta_2 \mathbf{x} - \gamma \mathbf{Z} - \boldsymbol{\varepsilon}}{\sigma}\right), & if \; 1 < m < 5, \\ 1 - \Phi\left(\frac{\alpha_4 - \beta_1 \mathbf{Gy} - \beta_2 \mathbf{x} - \gamma \mathbf{Z} - \boldsymbol{\varepsilon}}{\sigma}\right), & if \; m = 5, \end{array} \right\}$$

where Φ is the standard normal cumulative distribution function with standard deviation σ.

Estimating the model above assumes that each type of network effect is exogenously formed. However, it is well understood that unobservables in each subject's utility function are likely to be correlated with their location in the network. Unobservables, for example, may consist of advertisements the subject may have seen, discussions with others outside of the network, or any other information that may influence his or her decision. Therefore, the ordered probit model above is likely to suffer from endogeneity bias *a priori*. In order to correct for endogeneity, we use a two-stage control function estimation method (Park and Gupta 2009; Petrin and Train 2010). The control function method accounts for the bias likely to arise from the endogeneity of the peer-effect term using a two-stage estimation method based on the sample-selection models of Heckman (1978) and Hausman (1978). In the first stage, we from the control function by regressing the endogenous peer-effect term ($\mathbf{Gy}$) on a set of variables likely to serve as valid instruments.

In our experiment, demographic and behavioral variables are all likely to be correlated with peer effects, but mean independent from the error term. We then include the residuals from this regression as a control function in the ordered-probit model, which is then estimated using simulated maximum likelihood (Train 2003).

More formally, the control function is written as:

$$\mathbf{Gy} = \lambda Z_p + \nu, \tag{3}$$

where Z_p is a subset of the demographic variables captured in the experiment, ν is an iid error term, and λ is a vector of parameters to be estimated. The control function $CF(v;\phi) = \phi v$ is then included in the ordered probit model in order to remove the bias induced by also including the endogenous peer effects (Petrin and Train 2010). We chose the set of instruments conscious both of the need to maintain instrument validity (independence from the model error) and quality. While there are many sets of valid instruments in our data, the chosen set provides an F-statistic from the control function regression that is at least 10, thereby ensuring that our instruments are not "weak" in the sense of Staiger and Stock (1997). In the experimental data, there are more than 25 variables that are eligible to be valid instruments. As there are only a limited number of parameters in the model, the instruments at hand are sufficient to identify all peer effects.

Using this model, we compare the absolute values of the two social-network-effect parameters: The peer-effect parameter and the anonymous-reviewer effect parameter. If the peer-effects parameter is significantly larger than the anonymous-reviewer effect parameter, we conclude that the impact of trust among peers is stronger than the depth of knowledge expected from web-based anonymous sources.

4.2 Positive versus Negative Reviews

In the second model, we compare the effect of positive and negative reviews for both peer and anonymous Yelp reviews. We again use an ordered probit model because the dependent variable is the same ordered-rating variable as that used in the first model. Namely, the Likert-scale variable measuring the likelihood that the subject will return to the first-stage restaurant. Our hypothesis is that the observed rating-variable that describes each respondent's restaurant preferences is differentially affected by positive and negative reviews. Specifically, if the response of marginal utility with respect to positive and negative information is consistent with Prospect Theory (Kahneman and Tversky 1979), then we expect the utility function to be "steeper" in a negative relative to a positive direction. We define a "negative review" as one that is below 3-stars on the Yelp scale, and a "positive review" as one that is 3-stars or higher. We test our hypothesis using a variant on the ordered probit model described above.

For this model, the latent utility is written as:

$$\mathbf{u} = \beta_3\mathbf{P} + \beta_4\mathbf{N} + \gamma\mathbf{Z} + \boldsymbol{\varepsilon}, \tag{4}$$

where the vector **P** represents positive reviews and vector **N** represents negative reviews. Subjects are assumed to follow the same ordered-probit decision process as described above, so their choice reflects the option that reflects the highest level of utility. In terms of the estimated probability model, the estimating equation is written as:

$$\Pr(q_m = m) = \left\{ \begin{array}{ll} \Phi\left(\frac{\tau_1 - \beta_3\mathbf{P} - \beta_4\mathbf{N} - \gamma\mathbf{Z} - \boldsymbol{\varepsilon}}{\sigma}\right), & if\ m = 1, \\ \Phi\left(\frac{\tau_m - \beta_3\mathbf{P} - \beta_4\mathbf{N} - \gamma\mathbf{Z} - \boldsymbol{\varepsilon}}{\sigma}\right) - \Phi\left(\frac{\tau_{m-1} - \beta_1\mathbf{P} - \beta_4\mathbf{N} - \gamma\mathbf{Z} - \boldsymbol{\varepsilon}}{\sigma}\right), & if\ 1 < m < 5, \\ 1 - \Phi\left(\frac{\tau_4 - \beta_3\mathbf{P} - \beta_4\mathbf{N} - \gamma\mathbf{Z} - \boldsymbol{\varepsilon}}{\sigma}\right), & if\ m = 5, \end{array} \right\} \tag{5}$$

where τ_i is now the threshold level of utility that separates each rating level, and Φ is again the normal CDF with standard deviation σ. The elements of **P** and **N** are constructed from the star ratings on each Yelp review, where a star rating below average is deemed a "negative" review and an above average rating is defined as a "positive" review. The ε vector is an iid normal error term.

With this model, we test our second hypothesis that negative reviews have a stronger effect on restaurant choice relative to positive reviews. This "loss aversion" effect is a natural consequence of consumers behaving differently when experiencing gains than whey they experience losses (Kahneman and Tversky 1979). In this model, we again account for observed heterogeneity by adding demographic and behavioral variables through the **Z** matrix. In terms of the estimated coefficients, the Prospect Theory hypothesis is supported if the absolute value of β_4 is greater than the corresponding value of β_3. In other words, if the negative review coefficient is greater than the positive review coefficient, then we conclude that negative reviews can decrease the restaurant demand than positive reviews can increase it.

5 Results and Discussion

In this section, we present the results obtained from estimating both of the models described above. We establish the validity of each model by presenting a number of plausible alternatives – models that either do not consider the effect of unobserved heterogeneity, the spatial nature of network learning, or the potential endogeneity of peer effects. We then interpret the results from the model that provides the best fit to the data.

We examine the first empirical question, whether information obtained through anonymous networks have a greater or lesser effect on preferences than through peer networks, by pooling the data from stage one and stage two. Recall that, in the first stage, the treatment involves subjects viewing Yelp ratings, and in the second stage, subjects see peer reviews for a different restaurant. Model 1 in table 4 shows the estimates obtained from a model of non-spatial peer influence, that is, where the peer effect is measured as the simple arithmetic average of all peer reviews obtained from other members of the group. In this specification, only the peer effect is significantly different from zero, and the point estimate of the Yelp-review effect is, in fact, negative. In an ordered probit model, however, the structural parameter estimates are less important than the marginal effects of each explanatory variable on the probability of moving from one ordered value of the dependent variable to another. We report these values for each model in table 5.

From table 5, we see that the only statistically-significant marginal effects involve the peer effect in the middle three regimes. Recall that we coded each peer review into a Yelp-like five-star rating system for estimation purposes. Therefore, the marginal effects are interpreted as measuring the change in the probability of observing each "likelihood of returning" value for a one-star increment in the peer rating. For example, in table 5 the estimate of 0.1074 for the $\Pr(q = 5)$ category suggests that a one-star improvement in peer rating increases the probability of observing a value of $q = 5$ ("Very Likely" to revisit) by 10.74%. Because this estimate is positive, and statistically significant, it implies that a favorable peer review is able to increase the likelihood that a subject will return to the restaurant. The marginal effect of peer reviews is also relatively high for the $\Pr(q = 4)$ and $\Pr(q = 3)$ categories as a one-star improvement raises the probability of observing a "Likely" return response by 10.71% and "Undecided" response by 3.99%. On the other hand, a one-star improvement in Yelp review causes a reduction in the probability that the subject is either "Likely" or "Very Likely" to return to the restaurant, although each effect is only significant at a 10% level. At least in this simple model, these results imply that Yelp reviews may raise expectations – expectations that are not met by performance – so subjects' responses stand in contrast to the expected positive effect. If true, these results are consistent with the predictions of ACT as performance that varies widely from expectations, outside the latitude of acceptance, induces a contrasting response. Intuitively, people like to disagree with anonymous "experts" but not with their friends.

[tables 4 and 5 in here]

In Model 1, however, we do not account for the fact that peer effects are likely to vary by the strength of relationship between pairs of subjects, and how influential individual peers are. Model 2 accounts for differences in peer relationships by weighting each review by the inverse proximity of each member. That is, $\mathbf{W}$ is constructed such that an individual who is "closer" to the subject in a relationship-sense, should have greater influence on the subject relative to others. In table 4, we show that the "spatial" model provides a better fit to the data compared to the "non-spatial" alternative, based on a comparison of the log-likelihood function values, but the structural peer-effect estimate is smaller than the anonymous review effect. In fact, each of the marginal-effect parameters in table 5 are smaller in the spatial model relative to the non-spatial alternative. While finding that the spatial model provides a better fit to the data is expected, smaller marginal effects are perhaps counterintuitive. However, recall that the social proximity matrix is row-normalized so that the implied weights applied to each group member sum to 1 within each group. Because both the structural and marginal parameter estimates are averages over all members of all groups, it is entirely possible that the global average is lower in the spatial model. Even so, the estimates in Model 2 do not account for the potential endogeneity of peer effects.

In fact, neither of the first two models addresses the likely endogeneity of peer effects pointed out by Manski (1993). Before correcting for endogeneity, we first test for whether doing so is necessary as the logical potential for endogeneity does not ensure its existence. We use a Wu-Hausman test (Wu 1973, Hausman 1978) to test for the endogneity of peer effects. With the Wu-Hausman test, the null hypothesis is exogeneity, so rejecting the null implies that endogeneity is a feature of the data. The test involves comparing parameter estimates from an estimator that is efficient under the null hypothesis with those obtained with an estimator that is consistent under the alternative. Our consistent estimator is defined as the control-function model described above, where the instruments include a set of demographic variables (income, age, and education), behavioral measures taken from the survey instrument (would consider peer review, regards location, service, ambience, and taste important, has written an online review, and checks nutrient contents of packaged foods) and restaurant and round indicator variables. Using these instruments, the Wu-Hausman test statistic, which is chi-square distributed with degrees of freedom equal to the number of potentially endogenous variables, is 201.467, while the critical chi-square value is 3.84. Therefore, we reject the null hypothesis of exogeneity and conclude that the peer effects are endogenous. Table 6 presents the results from the first-stage, control-function regression. Although not all instruments are statistically significant on their own, the F-statistic of 14.436 is greater than 10, so our set of instruments is not "weak" according to the Staiger and Stock (1997) criteria.

[table 6 in here]

Models 1 and 2 also do not account for unobserved heterogeneity among experimental subjects. In Models 3 and 4, therefore, we estimate non-spatial and spatial models, respectively, that account for both endogeneity and unobserved heterogeneity. Not surprisingly, the results in table 4 show that Model 3 provides a better fit to the data than either of Models 1 or 2. However, the structural peer-effect parameter estimate is nearly 50% greater than in Model 1 and almost three times greater than in Model 2. Clearly, the bias induced by both endogeneity and unobserved heterogeneity is substantial, and negative in direction. Perhaps more important to the objectives of the paper, however, we find that the Yelp-review effect, which was not statistically significant in either of the first models, and negative in terms of the point-estimate, is now positive and statistically significant. Consistent with Anderson and Magruder (2012) and Luca (2011), we find a positive effect of Yelp-reviews, but the peer-effect is nearly four times as strong in Model 3, and three times in Model 4.

Comparing the non-spatial (Model 3) and spatial (Model 4) within the class of control-function models, we again find that the peer effect is substantially smaller in the spatial relative to the non-spatial model, although the spatial model provides a significantly better fit to the data. By disaggregating individual effects in the spatial model, we obtain a better estimate of the average influence of peers. While the non-spatial model assumes all others have the same level of influence, the spatial model allows for the fact that perhaps one or two individuals have an outsize influence on an individual's behavior, while the majority of the group may have no influence at all.

This pattern of results holds for the marginal effects estimated with Models 3 and 4. From table 5, we see that the marginal peer-effect in Model 3 is approximately four times the Yelp-review effect, while it is roughly three times the size of the Yelp-review effect in Model 4. Given that Model 4 is the preferred specification on the basis of fit, we interpret the marginal peer-effect as implying that a one-star improvement in peer rating will increase the probability that a subject will be very likely to return to the restaurant by nearly 16%, while a similar improvement in Yelp rating will increase the likelihood of returning by only 5%. The effect of a negative peer review is nearly symmetric in the opposite direction. The marginal effect of one-star reduction in peer rating leads to a nearly 25% greater chance of a "very unlikely" to return response, while the same change in Yelp-rating leads to only an 8% higher probability of not returning.

These marginal effects have important practical implications. While Anderson and Magruder (2012) find that there are considerable incentives for restaurant owners to manipulate the Yelp-review system in order to attract customers, we find that there are far greater incentives to – if not manipulate – at least manage peer reviews. If firms can devise ways to incentivize customers who have positive experiences in their restaurant to spread the word to their friends, the result is likely to be much more effective than paying others to log fraudulent online reviews. Indeed, Chevalier and Mayzlin (2006) argue that firm-generated WOM lacks credibility because consumers understand these incentives, so implicitly discount sources of information that they know are subject to manipulation. This is one reason why our peer effects may be so much stronger than the anonymous-reviewer effects. Peer effects are generally regarded as more genuine, and credible, than online reviews so are far more likely to be effective, even if they contain less real information than online alternatives. In short, depth and variety in anonymous reviews and presence of a large pool of information are dominated by the trust and familiarity in the peer networks. At least in the case of restaurant choice, peer networks are, in general, more trusted and hence likely overpower anonymous networks with unbiased information. Our estimated peer-effects also suggest a potential for WOM to generate substantial bandwagon effects, which may, in turn, help explain the boom-and-bust nature of restaurant patronage. Because information that "goes viral" in peer networks can spread deeply and quickly across all manner of social media, it is relatively easy for negative information to drive restaurant traffic below break-even levels. Moreover, because one of the attractions of popular restaurants is the mere fact of their popularity, and the attraction of crowds, positive WOM through peer-networks has a self-perpetuating aspect that no amount of marketing expenditure can replace.

Our second model compares the effect of positive and negative reviews on restaurant preference, and examines whether the marginal effect of negative reviews is larger (in absolute value) than positive reviews. Testing the symmetry of influence for online reviews amounts to a test of Prospect Theory (Kahneman and Tversky 1979) in a restaurant context as we would expect negative reviews to have a larger marginal effect on restaurant preference than positive reviews. The results obtained from estimating both a fixed and random coefficient version of an ordered probit model are shown in table 7. In the fixed-coefficient version of the model, the results in table 7 show that positive reviews have a positive effect on restaurant rating, and negative reviews have a negative effect, as expected. Notice also that positive reviews reduce the likelihood of a "very not likely" or "not likely" to revisit score, while negative ratings improve the probability that a customer does not return. This pattern is consistent over all response regimes. Further, the structural parameter estimates suggest that the absolute value of a negative effect is 18.5% larger than the positive effect in the fixed-coefficient model, and fully 37.3% larger in the random-coefficient model. In fact, controlling for unobserved heterogeneity in this way provides a better fit to the data (Likelihood Ratio statistic = 7.866, versus a critical chi-square value of 5.991). The random coefficient model implies a sharper difference between positive and negative reviews across the range of possible responses: Whereas there is an average gap in marginal effects of approximately 2% in the fixed-coefficient model, the difference grows to nearly 4% in the random-coefficient specification. Once we control for random unobserved subject attributes, the remaining asymmetry provides strong support for the implications of Prospect Theory in this experimental data.

[table 7 in here]

These findings also have important implications for both the theory of social learning, and management practice. In terms of the state of research on this issue, we document an important asymmetry that is suggested by theory, but has yet to be confirmed elsewhere in a social learning context. In other empirical models of social learning that involve reviews, positive and negative reviews should be separated for estimation purposes, or bias and inconsistency will ensue. From a practical standpoint, if our finding is true of restaurants, it is also likely to be true of any other category of good or service that is rated online. Although much of the discussion surrounds the manipulation of online reviews, which is typically manifest as owners paying third-party reviewers for posting positive reviews, our findings suggest that owners would be far better off paying online reviewers to not post negative reviews. These results also provide insight into the binary nature of restaurant success described in the introduction. If negative reviews are indeed as powerful as our results suggest, then it is not hard to imagine how news of either bad food or service can spread quickly through a community and doom a restaurant to failure.

6 Conclusion

Consumers have access to many sources of information that may help resolve uncertainty regarding untested experience goods. With the rise of online social media, researchers tend to focus on internet-based sources as an emerging, dominant influence on consumer behavior. However, social media also catalyze more traditional peer-based social network effects. In this study, we compare two categories of social networks (anonymous and peer networks) in terms of their effect on restaurant preference. We condition the effect of peer reviews by considering the proximity among individuals in their peer network, while comparing the effect of recommendations from one's peers to those obtained from anonymous reviewers through Yelp.com. We also compare negative and positive reviews and determine the relative effect of each on restaurant preference. We use a controlled, experimental approach in order to address the reflection problem (Manski 1993) that typically bedevils inference in empirical problems of social learning.

Our experiment consists of two stages using real online anonymous restaurant reviews from Yelp.com and peer reviews from multiple peer groups to compare anonymous and peer networks. In the first stage, subjects are provided anonymous Yelp reviews and asked to visit, and rate the likelihood of returning to a particular restaurant. In the second stage, they receive peer reviews for a different restaurant, and are again asked how likely they would be to return. By comparing preferences after receiving each type of review, we are able to compare the relative effectiveness of each.

Our empirical approach is unique in that we devise a spatial econometric method of testing for peer effects. That is, peer reviews are weighted by each subject's location in the peer network. These weights are constructed from adjacency matrices that reflect the proximity of each member to all others, where proximity is defined as how well each member knows the others. Controlling for both the endogeneity of peer effects, and unobserved heterogeneity, we find that both peer and anonymous reviews have a significant, positive impact on restaurant preference. However, we also find that peer reviews are approximately three times as influential as anonymous reviews in determining consumer preferences. We also find that both negative and positive reviews can influence preferences, but negative reviews have a larger adverse effect on restaurant preference than the demand-enhancing effect of positive reviews.

This research has many important implications for both future researchers as well as industry practitioners. From a methodological perspective, our experiment provides a new way of comparing peer and anonymous social network effects. Econometrically, we demonstrate how methods developed in spatial econometrics can be applied to the analysis of social relationships, and social learning. As a practical matter, it is well understood that a significant portion of advertising expenditure is lost because of poorly targeted advertising. By more accurately targeting these expenditures toward influential network members, much of this loss can be avoided. Our study focuses on restaurants, but the results likely extend to similar industries such as hotels, local contractors, bars and amusement parks. In each case, consumers face a high degree of prior uncertainty in consuming a multi-attribute good, and are likely to turn to social media – whether populated by peers or anonymous reviewers – to resolve this uncertainty. As user activity and the online user base increase in the future, social media marketing will become even more important and earn an even larger proportion of marketing expenditures from the traditional advertising media. Further, our findings suggest that small businesses with limited marketing budgets may be able to leverage peer-based WOM in creating a marketing program that may be more effective than programs with much greater expenditure on traditional marketing media.

There are also few limitations of this research. First, the necessarily-small size of peer groups can be a limiting factor for any controlled social-networking experiment. Larger experiments may be able to provide better data in a statistical sense, but using real networks of real people limits the scalability of any social network experiment. Future research in this area is required in comparing anonymous network effects with peer network effects in other high involvement categories such as durable home appliances, automobile, medical care, holiday packages, house purchases, and education investments. While we extend the existing research on anonymous network effect by including actual review content, future research may include attributes of the reviewers, review characteristics, and dynamic changes in reviews over time.

References

[1] Anderson, M., and Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. The Economic Journal, 122, 957-989.

[2] Banerjee, A. V. (1992). A simple model of herd behavior. The Quarterly Journal of Economics, 107, 797-817.

[3] Becker, G. S. (1991). A note on restaurant pricing and other examples of social influences on price. Journal of Political Economy, 99, 1109-1116.

[4] Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transportation Research Part B: Methodological, 35, 677-693.

[5] Bhatnagar, A., and Ghose, S. (2004). Online information search termination patterns across product categories and consumer demographics. Journal of Retailing, 80, 221-228.

[6] Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. Journal of econometrics, 150, 41-55.

[7] Brock, W. A., and Durlauf, S. N. (2002). A multinomial-choice model of neighborhood effects. The American Economic Review, 92, 298-303.

[8] Brock, W. A., and Durlauf, S. N. (2007). Identification of binary choice models with social interactions. Journal of Econometrics, 140, 52-75.

[9] Buda, R., and Zhang, Y. (2000). Consumer product evaluation: the interactive effect of message framing, presentation order, and source credibility. Journal of Product and Brand Management, 9, 229-242.

[10] Buttle, F. A. (1998). Word of mouth: understanding and managing referral marketing. Journal of Strategic Marketing, 6, 241-254.

[11] Cai, H., Chen, Y., and Fang, H. (2007). Observational learning: Evidence from a randomized natural field experiment. (No. w13516). National Bureau of Economic Research.

[12] Cheung, C. M., and Lee, M. K. (2012). What drives consumers to spread electronic word of mouth in online consumer-opinion platforms. Decision Support Systems, 53, 218-225.

[13] Chevalier, J. and D. Mayzlin. (2006). The effect of word of mouth on sales: online book reviews. Journal of Marketing Research, 43, 345-354.

[14] Dawes, J. G., (2008). Do data characteristics change according to the number of scale points used? An experiment using 5 point, 7 point and 10 point scales. International Journal of Market Research, 51, 61-77.

[15] De Bruyn, A., and Lilien, G. L. (2008). A multi-stage model of word-of-mouth influence through viral marketing. International Journal of Research in Marketing, 25, 151-163.

[16] Dierkes, T., Bichler, M., and Krishnan, R. (2011). Estimating the effect of word of mouth on churn and cross-buying in the mobile phone market with Markov logic networks. Decision Support Systems, 51, 361-371.

[17] Domingos, P. (2005). Mining social networks for viral marketing. IEEE Intelligent Systems, 20, 80-82.

[18] Duan, W., Gu, B., and Whinston, A. B. (2008). Do online reviews matter?—An empirical investigation of panel data. Decision Support Systems, 45, 1007-1016.

[19] Duflo, E., and Saez, E. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. Journal of Public Economics, 85, 121-148.

[20] Duflo, E., and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. The Quarterly Journal of Economics, 118, 815-842.

[21] Freeman, L. C. (1979). Centrality in social networks conceptual clarification. Social networks, 1, 215-239.

[22] Giorgi, G. D., Pellizzari, M., and Redaelli, S. (2007). Be as careful of the books you read as of the company you keep: Evidence on peer effects in educational choices, IZA Discussion papers no. 2833. Institute for the Study of Labor (IZA).

[23] Godes, D., and Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. Marketing Science, 23, 545-560.

[24] Godes, D., and Mayzlin, D. (2009). Firm-created word-of-mouth communication: Evidence from a field test. Marketing Science, 28, 721-739.

[25] Goldenberg, J., Han, S., Lehmann, D. R., and Hong, J. W. (2009). The role of hubs in the adoption process. Journal of Marketing, 73, 1-13.

[26] Hanneman, R., and Riddle, M. (2005). Introduction to social network methods; free introductory textbook on social network analysis. Department of Sociology, University of Northern Colorado.

[27] Hausman, J. A. (1978). Specification tests in econometrics. Econometrica: Journal of the Econometric Society, 46, 1251-1271.

[28] Hausman, J. A., Lo, A. W., and MacKinlay, A. C. (1992). An ordered probit analysis of transaction stock prices. Journal of Financial Economics, 31, 319-379.

[29] Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system, Econometrica, 46, 931–59.

[30] Herr, P. M., Kardes, F. R., and Kim, J. (1991). Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective. Journal of Consumer Research, 454-462.

[31] Hilligoss, B., and Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. Information Processing and Management, 44, 1467-1484.

[32] Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. Journal of Management, 21, 967-988.

[33] Hong, H., J. D. Kubik and J. C. Stein. (2004). Social Interaction and Stock-Market Participation. Journal of Finance, 59, 137-163.

[34] Hu, Y., and Li, X. (2011). Context-dependent product evaluations: an empirical analysis of internet book reviews. Journal of Interactive Marketing, 25, 123-133.

[35] Ibarra, H., and Andrews, S. B. (1993). Power, social influence, and sense making: Effects of network centrality and proximity on employee perceptions. Administrative Science Quarterly, 38, 277-303.

[36] Iyer, G., Soberman, D., and Villas-Boas, J. M. (2005). The targeting of advertising. Marketing Science, 24, 461-476.

[37] Jackson, M. O. 2008. *Social and Economic Networks* Princeton: Princeton University Press.

[38] Johns, N. (1999). What is this thing called service? European Journal of Marketing, 33, 958-974.

[39] Kahneman, D. and A. Tversky. (1979). Prospect theory: an analysis of decision under risk. Econometrica, 47, 263-291.

[40] Luca, M. (2011). Reviews, reputation, and revenue: The case of Yelp. com. Working paper no. 12-016, Harvard Business School, Harvard University.

[41] Mack R., R. Mueller, J. Crotts and A. Broderick. 2000. Perceptions, corrections and defections: implications for service recovery in the restaurant industry. Managing Service Quality, 10, 339-346.

[42] Mangold, W. G., Miller, F., and Brockway, G. R. (1999). Word-of-mouth communication in the service marketplace. Journal of Services Marketing,13, 73-89.

[43] Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. The Review of Economic Studies, 60, 531-542.

[44] Manski, C. F. (2000). Economic analysis of social interactions. (No. w7580) National Bureau of Economic Research.

[45] Maxham III, J. G. (2001). Service recovery's influence on consumer satisfaction, positive word-of-mouth, and purchase intentions. Journal of Business Research, 54, 11-24.

[46] Mayzlin, D. (2006). Promotional chat on the internet. Marketing Science, 25, 155-163.

[47] McCullough, R. D. (1998). The chemistry of conducting polythiophenes. Advanced materials, 10, 93-116.

[48] Mizerski, R. (1982). An attribution explanation of the disproportionate influence of unfavorable information. Journal of Consumer Research, 9, 301-310.

[49] Muller, C. C., and Woods, R. H. (1994). An expanded restaurant typology. The Cornell Hotel and Restaurant Administration Quarterly, 35, 27-37.

[50] Nam, S., Manchanda, P., and Chintagunta, P. K. (2010). The effect of signal quality and contiguous word of mouth on customer acquisition for a video-on-demand service. Marketing Science, 29, 690-700.

[51] Narayan, V., Rao, V. R., and Saunders, C. (2011). How peer influence affects attribute preferences: A Bayesian updating mechanism. Marketing Science, 30, 368-384.

[52] Nelson, K. (1974). Concept, word, and sentence: Interrelations in acquisition and development. Psychological review, 81, 267.

[53] Ong, B. S. (2012). The perceived influence of user reviews in the hospitality industry. Journal of Hospitality Marketing & Management, 21, 463-485.

[54] Opsahl, T., Agneessens, F., and Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks, 32, 245-251.

[55] Park, S., and Gupta, S. (2009). A simulated maximum likelihood estimator for the random coefficient logit model using aggregate data. Journal of Marketing Research, 46, 531-542.

[56] Parsa, H. G., J. T. Self, D. Njite, and T. King. (2005). Why restaurants fail. Cornell Hotel and Restaurant Administration Quarterly, 46, 304-322.

[57] Petrin, A., and Train, K. (2010). A control function approach to endogeneity in consumer choice models. Journal of Marketing Research, 47, 3-13.

[58] Richards, T. J., and Patterson, P. M. (1999). The economic value of public relations expenditures: Food safety and the strawberry case. Journal of Agricultural and Resource Economics, 440-462.

[59] Richards, T. J., W. Allender, and S. F. Hamilton. (2013). Social networks and new product introduction. American Journal of Agricultural Economics (forthcoming).

[60] Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. The Quarterly Journal of Economics, 116, 681-704.

[61] Staiger, D., and J. H. Stock. (1997). Instrumental variables regression with weak instruments. Econometrica, 65, 557-586.

[62] Train, K. (2003). Discrete choice methods with simulation. Cambridge, UK: Cambridge University Press.

[63] Trusov, M., Bodapati, A., and Bucklin, R. E. (2010). Determining influential users in internet social networks. Journal of Marketing Research, 47, 643–658

[64] Urberg, K. A., Luo, Q., Pilgrim, C., and Degirmencioglu, S. M. (2003). A two-stage model of peer influence in adolescent substance use: Individual and relationship-specific differences in susceptibility to influence. Addictive Behaviors, 28, 1243-1256.

[65] Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’networks. Nature, 393, 440-442.

[66] Weimann, G. (1983). The strength of weak conversational ties in the flow of information and influence. Social Networks, 5, 245-267.

[67] Wu, D. M. (1973). Alternative tests of independence between stochastic regressors and disturbances. Econometrica: Journal of the Econometric Society, 41, 733-750.

[68] Zenger, T. R., and Lawrence, B. S. (1989). Organizational demography: The differential effects of age and tenure distributions on technical communication. Academy of Management Journal, 32, 353-376.

Table 1: Social Networking Experiment Data Summary

Survey Question	Units	Mean	Std. Dev.	Min.	Max.
Eat out Frequently	# / wk	2.9779	0.8010	1	4
Number of Drinks	#	1.9338	0.8778	1	6
Food Quality Important	1=No, 5=Yes	3.3235	0.9786	1	5
Taste Important	1=No, 5=Yes	3.2978	1.0035	1	5
Service Important	1=No, 5=Yes	3.5000	1.0968	1	5
Location Important	1=No, 5=Yes	3.4816	0.9980	1	5
Price Important	1=No, 5=Yes	3.2500	0.8697	1	5
Ambience Important	1=No, 5=Yes	3.1471	0.9374	1	5
Variety Important	1=No, 5=Yes	3.2463	0.9020	1	5
Healthy Options Important	1=No, 5=Yes	2.3382	0.6567	1	5
Would you revisit?	1=No, 5=Yes	3.1471	1.2569	1	5
Age	Years	37.5338	12.8466	20	79
Gender	1=Male	0.4926	0.5009	0	1
Education	Years	13.3897	2.7360	10	20
Marital Status	1=Married	0.8529	0.5637	0	1
Dependents	#	2.9412	1.6424	0	7
Income	$,000	87.8677	59.2843	12.5	250
Online Time	Minutes / day	2.7059	1.0246	1	4
Social Networking Websites	#	2.4338	1.4048	1	6
Use Online Reviews	1=Yes	0.5993	0.3996	0	1

Table 2: Anonymous Reviews, Peer Reviews, and Restaurant Ratings

Overall	Anonymous Reviews			Peer Reviews		
Rating (1 - 5 Scale)	Control	Negative	Positive	Control	Negative	Positive
Restaurant	3.6111	3.5000	3.7200	3.6316	3.4583	3.4583
1	1.2433	1.1045	1.4000	1.1648	1.2151	1.3181
Restaurant	3.4737	3.7917	3.7917	3.7778	4.1538	3.7600
2	1.2188	1.4440	1.1413	1.1144	1.0466	1.0520

Note: Upper value is mean rating and lower value is standard deviation of rating.

Table 3: Rating Categories by Review Type

Overall	Stage-1 Yelp Reviews				Stage-2 Peer Reviews		
Rating	Negative	No Review	Positive	Total	No Reviews	Peer Reviews	Total
(1 - 5 Scale)	N=50	N=37	N=49	N=136	N=37	N=99	N=136
1	8.00%	5.40%	0.00%	4.40%	0.00%	0.00%	0.00%
2	6.00%	13.50%	22.40%	14.00%	16.20%	17.20%	16.90%
3	38.00%	32.40%	26.50%	32.40%	32.40%	35.40%	34.60%
4	10.00%	18.90%	4.10%	10.30%	16.20%	6.10%	8.80%
5	38.00%	29.70%	46.90%	39.00%	35.10%	41.40%	39.70%

Table 4: Ordered Probit Model Estimates

Model 1: Non-Spatial, Fixed Parameter			Model 2: Spatial, Fixed Parameter			Model 3: Non-Spatial, Random Parameter			Model 4: Spatial, Random-Parameter		
Variables	Estimates	t-ratios	Variables	Estimates	t-ratios	Variables	Estimates	t-ratios	Variables	Estimates	t-ratios
						Control	-0.2489*	-4.9367	Control	-0.7028*	-8.8142
						Random Parameter Means					
Peer Review	0.6464*	5.1225	W*Peer Reviews	0.3580*	6.1146	Peer Review	0.9232*	5.8913	W*Peer Review	0.7688*	10.4297
Yelp Review	-0.1363	-1.6070	Yelp Review	-0.1384	-1.6323	Yelp Review	0.2413*	6.6187	Yelp Review	0.2456*	6.624326
						Random Parameter Scales					
						Peer Review	0.0743	0.5643	W*Peer Review	0.0003	0.0058
						Yelp Review	0.1737*	5.1936	Yelp Review	0.1081*	3.3368
Threshold Parameters						Threshold Parameters					
α_1	0.2894*	7.8547	α_1	0.3037*	7.9181	α_1	0.3544*	7.1470	α_1	0.4275*	7.3607
α_2	0.7311*	13.5238	α_2	0.7568*	13.7013	α_2	0.8587*	16.1992	α_2	0.9842*	18.1534
α_3	1.4474*	17.7443	α_3	1.4816*	17.9597	α_3	1.6367*	18.6972	α_3	1.7996*	22.1598
LLF	-496.683			-491.365			-477.974			-457.946	
AIC/N	3.689			3.651			3.573			3.426	

Note: A single asterisk indicates significance at a 5% level. "Peer Review" is an arithmetic mean of group-level reviews, while "W*Peer Review" is peer reviews weighted by the spatial proximity of others in the group. "Control" is the control-function parameter estimate. LLF = log-likelihood function value, AIC = Akaike Information Criterion.

Table 5: Marginal Effects of Peer and Anonymous Reviews

		Model 1		Model 2		Model 3		Model 4	
		Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio
$\Pr(q=1)$	Peer Review	-0.2559*	-5.1985	-0.1403*	-6.3258	-0.3361*	-6.3778	-0.2474*	-12.8766
	Yelp Review	0.0540	1.6153	0.0543	1.6392	-0.0878*	-7.4573	-0.0790*	-7.7119
$\Pr(q=2)$	Peer Review	0.0015	1.0771	-0.0015	-0.9861	-0.0313*	-3.0765	-0.0515*	-5.7289
	Yelp Review	-0.0003	-0.8125	0.0006	0.9295	-0.0082*	-3.0833	-0.0164*	-4.3894
$\Pr(q=3)$	Peer Review	0.0399*	5.2939	0.0204*	6.3201	0.0316*	3.8101	0.0082	0.9357
	Yelp Review	-0.0084	-1.5828	-0.0079	-1.5943	0.0083*	3.8751	0.0026	0.9375
$\Pr(q=4)$	Peer Review	0.1071*	4.6002	0.0596*	5.3056	0.1583*	5.2075	0.1312*	7.8864
	Yelp Review	-0.0226	-1.5859	-0.0231	-1.6069	0.0414*	6.8390	0.0419*	7.2705
$\Pr(q=5)$	Peer Review	0.1074*	4.2220	0.0618*	4.6850	0.1774*	5.0105	0.1595*	7.4994
	Yelp Review	-0.0226	-1.5773	-0.0239	-1.6020	0.0464*	4.9141	0.0510*	5.0558

Note: A single asterisk indicates significance at a 5% level.

Table 6: Control Function Estimates

Variable	Estimate	t-ratio
Constant	-2.4711*	-3.3752
Income	-0.0002	-0.1889
Age	0.0290*	6.8373
Education	0.0097	0.5010
Consider Peer Review	-0.6439*	-2.3790
Location Important	0.0308	0.5004
Ambience Important	-0.0444	-0.6327
Service Important	0.0950	1.2089
Write Online Review	0.1482	1.3989
Nutritional Value Important	-0.0944	-0.8735
Taste Important	0.0988	1.1787
Restaurant 1	0.0034	0.0344
Round 1	1.0264*	10.4336
R^2	0.401	
F	14.436	

Note: A single asterisk indicates significance at a 5% level.

Table 7: Anonymous Review Symmetry Estimates

		Model 1		Model 2	
		Estimate	t-ratio	Estimate	t-ratio
Structural Estimates					
Positive		0.5823*	4.4099	0.6004*	2.1506
Negative		-0.6902*	-6.0288	-0.8245*	-6.1824
Threshold Parameters					
α_1		0.4285*	6.1990	0.4279*	2.9027
α_2		1.0273*	11.3038	1.0509*	10.3938
α_3		1.7220*	13.8958	1.8186*	9.3205
Standard Deviation of Random Parameters					
σ_{Pos}		N.A.		0.0666	0.5069
σ_{Neg}		N.A.		0.4944*	3.3596
Marginal Effects					
$\Pr(q=1)$	Positive	-0.2018*	-4.9888	-0.2004*	-2.5499
	Negative	0.2392*	7.1129	0.2752*	7.9428
$\Pr(q=2)$	Positive	-0.0293*	-2.4701	-0.0357	-1.3267
	Negative	0.0347*	2.8839	0.0490*	2.3552
$\Pr(q=3)$	Positive	0.0257*	2.8295	0.0200	1.8984
	Negative	-0.0305*	-2.7603	-0.0275*	-2.2976
$\Pr(q=4)$	Positive	0.0911*	3.5845	0.1025*	3.6390
	Negative	-0.1080*	-4.2643	-0.1407*	-3.7265
$\Pr(q=5)$	Positive	0.1143*	3.2385	0.1136	1.4693
	Negative	-0.1354*	-3.8883	-0.1560*	-2.8031
LLF		-233.672		-229.739	
AIC/N		3.510		3.481	

Note: A single asterisk indicates significance at a 5% level. LLF = log-likelihood function value; AIC = Akaike Information Criterion, N = number of observations.