Social Networks and New Product Choice

Timothy J. Richards, William Allender and Stephen F. Hamilton*

June 3, 2012

Abstract

Influential individuals in a social network environment are important in shaping preferences for new products. In this study, we adopt an incentive compatible choice-based conjoint analysis approach to generate data on the introduction of a new ice cream product. We use spatial econometric methods to determine how individuals are likely to change their preferences when exposed to the choices of other members in their social network. We find evidence that agents look to others for guidance in their preference for subjective or taste-specific parameters, but rely on own preferences for objectively measured attributes such as price. We also use spatial methods to determine which network-member is the most influential. We find that the most connected member is not necessarily the most influential, and that influence can be determined econometrically.

Keywords: choice-based conjoint, experimental economics, new product introduction, social network analysis, spatial econometrics

---

*Richards is the Morrison Professor of Agribusiness in the Morrison School of Agribusiness and Resource Management, College of Technology and Innovation, Arizona State University; Allender is Ph.D. candidate in the Morrison School and Hamilton is Professor, Orfalea College of Business, California Polytechnic State University, San Luis Obispo. Contact author: Richards. Address: 7231 E. Sonoran Arroyo Mall, Mesa, AZ. 85212. Ph. (480) 727-1488, Fax: (480) 727-1961, email: trichards@asu.edu.
Support from the Agriculture and Food Research Initiative - National Institute for Food and Agriculture, USDA is gratefully acknowledged. Copyright 2012. Please do not cite or quote without permission.

1 Introduction

Global consumer packaged good (CPG) manufacturers introduce approximately 20,000 new products each month (Mintel 2011). Developing and marketing innovative new products is essential for almost all firms, especially CPG manufacturer to keep up with constantly-changing consumer tastes, to evolve with the state of technology, or ensure that the product line doesn't become "stale" in the minds of consumers. Despite the high rate of new-product failure (approximately 80%, Sudhir and Rao 2006), the benefis to innovating firms is evident in the rate at which new products are brought to the market. Frequent failure, however, is also evidence of how difficult and research-intensive the commercialization process can be. In recent years, researchers have identified the importance of leveraging the value of social networking media as a relatively inexpensive means of advertising a new product. However, relatively little is known regarding how influential individuals within a social network affect the choices of others and how this impact can be used to improve the efficiency of the new product introduction process. In this study, we offer evidence from an incentive-compatible choice-based conjoint experiment implemented in a social network environment that shows how important individuals can affect the choices of others, and how these individuals can be identified. In this regard, our research considers two related problems: (1) how do different definitions of network membership impact a member's preference for attributes? and (2) how does network membership by one member affect the preferences of others?

While our objective is fundamentally empirical, answering these positive questions is important in testing theoretical mechanisms that drive social influence and provides managerially-relevant tools for new product development. The notion that one consumer's preferences could depend on another's choices has been well understood at least since Leibenstein (1950). Interdependent preferences, or the mechanism that underlies social contagion, may arise for any of a number of reasons. First, consumers may want to conform and simply be like others, or, conversely, strive to be different from others (Cowan, Cowan and Swann 1997). Second, if a consumer is not familiar with a product, she faces two alternatives: consume now and risk disappointment, or delay consumption and learn from others by word of mouth (McFadden and Train 1996; Godes and Mayzlin 2009). Third, a consumer's social network may consist of influential others that either provide useful information regarding attributes of uncertain value (Narayan, Rao and Saunders 2011; Iyengar, Van den Bulte and Valente 2011), or simply serve as aspirational models for goods that convey some notion of status (Yang and Allenby 2003). Few studies have investigated the nature of this effect, however, as influence and aspiration, in turn, have many dimensions. Specifically, the influence of a referent other in any social network is likely to depend on the strength of ties to the individual (Granovetter 1983; Borgatti and Everett 1999), the number of ties the individual has with others in the network (Goldenberg et al 2009), the level of respect network members have for his or her choices, the amount of trust others place in his or her ability to make a decision that could benefit themselves, or simply the perception of how influential the other is likely to be (Narayan, Rao and Saunders 2011). By framing our experiment and empirical analysis within a relatively simple product category (ice cream), we focus not on the question of *why* social network membership affects consumer choice, but rather *how* it does.

We recognize that there are two concepts of space that are relevant to the new-product development problem: social network space and attribute space. Individuals are arrayed at different locations in the social network. The strength of their social relationships is influenced not only by their location in the network, but the strength of ties to others in their vicinity. For example, if Jane and Sarah are both work associates of Mark and Tom who are close friends with eachother, but Sarah is also married to Tom, this puts Jane and Sarah and two fundamentally different locations in social network space. Empirical models of interdependent preferences adopt one of three approaches: (1) directly model interdependence, (2) use indirect or proxy measures of space, or (3) direct or structural modeling of the social network. Directly modeling interdependence requires no prior theory as to why choices of different individuals appear to be related, but simply captures the expectation that they do (Alessie and Kapteyn 1991; Brock and Durlauf 2001; Dugundji and Walker 2005; Brock and Durlauf 2007). Importantly, these models capture the essential feedback effect, or global interactions, that occur when one consumer's preferences depend on those of the group, but recognize that the individual is also part of the group. Marketers, however, are more interested in local interactions, when subsets of the group may have an outsize influence on individual decision making. Second, because consumers who happen to be closer to others geographically are more likely to share information, witness each other's choices or aspire to be like each other, geographic notions of relative location are a reasonable proxy for location in a social network (Case 1991; Yang and Allenby 2003; Grinblatt, Keloharju, and Ikaheimo 2008; Choi, Hui and Bell 2010). Third, a structural social network model is one that explicitly includes a variable that captures the nature of the linkage between a member and all others – typically a "social weight matrix" that measures the degree of proximity or connectedness between any two network members. Structural models of the social network capture not only the fact that preferences are interdependent, but attempt to explain the mechanism by which the local interaction occurs (Ballester, Calvo-Armengol and Zenou 2006; Calvo-Armengol, Patacchini and Zenou 2009; Narayan, Rao and Saunders 2011). Our approach falls into this third group as we use spatial econometric methods to account for the effect of relationship strength and centrality on the power of influential network members to alter others' preferences.

Researchers have long sought a way of modeling horizontal, or attribute, differentiation in a consistent way. We define horizontal differentiation as the distance – in Euclidean terms – from one product to all others in attribute space. Brownstone and Train (2000) demonstrate the value of using attribute-based models to study new product penetration through a mixed logit (or random parameters logit) model. Random parameter models allow attributes to enter the utility model explicitly, thus altering the competitive structure of the market depending on the location of the entrant in attribute space (Nevo 2003; Petrin 2002). While mixed logit models represented a significant advance in the development of attribute-based demand systems, they are still subject to the "crowding" problem noted by Akerberg and Rysman (2005). Namely, traditional discrete choice models tend to overestimate the benefit of new products because they do not allow the attribute space to become "more crowded" with entry (Trajtenberg 1989; Petrin 2002). Implicitly allowing the size of the attribute space spanned by existing products to grow with each new product has other serious implications. First, utility grows without bound with the introduction of new products (Berry and Pakes 2007). Second, equilibrium markups will not be driven to zero as crowding is not allowed to drive per-product demand to zero. Ackerberg and Rysman (2005), on the other hand, suggest that it is more intuitively reasonable that each new product should make the existing attribute space more crowded and, thus, compete for existing customers more aggressively than if there were no congestion effects. Berry and Pakes (2007) address this problem by developing a "pure characteristics" demand model in which consumers express a preference for attributes directly and not products as is usually the case. Our approach allows preferences to depend on attributes themselves by developing an explicitly spatial model of demand in which congestion arises naturally from our definition of utility.

Our research contributes to the emerging quantitative literature on social contagion and interdependent preferences in three ways. On a methodological level, we show how different definitions of social interaction can be used to estimate how preferences for one individual depend on others in his or her social network. While others find that usage rate, loyalty and the degree of friendship witnessed in any dyad may be inversely related to the degree of influence (Godes and Mayzlin 2009; Iyengar, Van den Bulte and Valenti 2011) we find support for the opposite, more conventional pathway, as social proximity is directly related to influence. Substantively, our empirical results help clarify the nature of the mechanism through which social contagion works by examining alternative ways of describing social-network location in a relatively simple, well-understood product category: Ice cream. By choosing a packaged-good category for the context of our study, we rule out the uncertainty effect modeled by Narayan, Rao and Saunders (2011) so that any remaining network effect likely derives from imitation and status (Yang and Allenby 2003) or new product learning (Godes and Mayzlin 2009). Because we consider both the primal (how do others influence an individual's choice?) and inverse (how does an individual influence others choices?) social network problems, our contribution is also on a practical level. Identifying inflential individuals in social networks has long been an objective of any firm using social media, but the methods to do so have proven somewhat elusive.

We find that individuals look to others for guidance in their preference for subjective or taste-specific attributes, but to their own preferences regarding objectively-measured attributes such as price. Mediating purely social network effects with measures of each product's location in attribute space tends to accentuate these findings: Individuals tend to rely on personal preferences when responding to price signals, but look for social cues as to more subjective attributes that affect taste (fat content, for example) or uniqueness (flavor, for example). Of the many measures of centrality and inter-linkage in the social network, we find that simple measures of social proximity are the most useful in explaining how subjects revise their preferences when exposed to the choices of others. We also find that influential individuals are not necessarily those that are most central or connected in the network but, similar to Goldenberg et al (2009), individuals that are regarded as leaders within the network. Our econometric methods can be used to identify such individuals.

In the next section, we describe the incentive-compatible choice-based conjoint experiment used to generate both willingness-to-pay and social network relationship data. In the second section, we use the emergent theory of interaction in social networks to develop a number of hypotheses regarding the mechanisms that cause individual network members to differ in influence. The third section contains our econometric model of social interaction, product choice and how an individual's location and power in the network can influence the choices of others. We present and interpret the data and econometric estimates in a fourth section, beginning with a summary of the experimental data and moving to parametric tests of network influence. A final section concludes and suggests avenues for future research in this area.

2 Social Network Experiment

In conventional models of product-choice, utility is assumed to refer only to one's own happiness and not the welfare of others. In a social network environment, decisions made by individuals spill over to affect other individuals through any one of the network effects described above: imitation, rejection, word-of-mouth, learning or a bandwagon contagion effect. When one individual tries a new product and informs others about the existence and nature of the product after consuming it, demand will differ depending upon the influential power of the individual disseminating the information and the strength of his connection to others. The key to incorporating social network effects into a marketing program is to design a mechanism by which the firm is able to internalize the positive network externalities generated by influential individuals. The basic idea is straightforward. In social networks, "word of mouth" advertising has been shown to be far more effective than other forms of promotion such as in-store sampling or trade promotions (Reingen and Brown 1987; Goldenberg et al. 2001; Domingos 2005) when designed to create an independent conversation, or one that wouldn't have occured in the absence of the network (Godes and Mayzlin 2009). In short, referrals from trusted network members are seen more credible than promotional activities from organizations that have an obvious vested interest in promoting a product or service (Reingen and Kernan 1986; Reingen and Brown 1987). In the terminology of viral marketing, targeting highly connected individuals in the social network is termed maximizing the "network value" of each customer (Domingos 2005). Maximizing network value requires the internalization of positive network externalities. The data gathered through the experiment described below is used to determine both the structure of the social network, and how network influence interacts with distance in attribute space in driving demand variation.

Our experimental design combines an exercise in product-attribute valuation and social network effects. The experiment consists of three stages, each of which is implemented through an online survey service (NetworkGenie.com – https://secure.networkgenie.com/). We conduct the experiment twice, once with a relatively small sample of students at a large Southwestern US university, and the second at a smaller university in California. The samples are both drawn from senior classes of business majors. Frequent group work and a common set of core classes mean that students are likely to at least know each other casually, and may have more intimate knowledge of each others' leadership abilities in many cases. The first sample consists of 34 students while the second consists of 73. Two features of a social networking sample are important to keep in mind. First, the subjects must have either interacted with each other, or been provided the opportunity to interact. Random, respresentative samples of large populations are not appropriate because of the low probability the subjects will know each other. Second, the sample size must be relatively small (Narayan, Rao and Saunders 2011) beause each network member is asked a series of questions regarding all others in the sample. Because each subject has to describe their relationship will all others, a larger sample size would be too burdensome for the subjects to complete in a reasonable time. Pilot surveys with more than 80 people resulted in a high drop-out rate, and much negative feedback.

In the first stage, we sent an email request that each subject fill out the survey through an online network data tool (NetworkGenie.com) that elicits profile information, including standard demographic and socioeconomic background (see instrument in the appendix) and, most importantly, data necessary to construct the social relationship matrix. There are five separate relationship matrices, each providing a different metric for the "strength of tie" between a subject and all others or measuring the likely degree of influence an individual may have. Multiple matrices are necessary because prior research shows that influence in a social network can derive from many possible sources: frequency of communication (Goldenberg et al 2009), degree of acquaintance (Godes and Mayzlin 2009), trust (Buskens 1998; Berrera 2007), respect or leadership (Mullen, Johnson and Salas 1991; Grippa and Gloor 2009). One individual in the class may know another very well from previous interactions, and yet not communicate frequently enough for the other to represent an influence on their decisions. Further, two individuals may be good friends, and communicate frequently, but one may not trust the other to provide valuable information with respect to product recommendations. For a similar reason, there is no reason to expect that friendship or trust should require one individual to respect the decisions made by the other. Finally, none of these attributes capture the influence an individual may have over group decisions if that individual is seen as a natural leader, trend-setter or early adopter. In each case, we ask subjects to respond on a Likert scale from very little communication, degree of acquaintance, trust, respect or leadership to very much. These questions were pre-tested on a sample of graduate students to ensure that the responses to each question produced some degree of independent information on strength of tie relative to the others.

In stage two, we administer an incentive-compatible (ie. non-hypothetical) choice-based conjoint (CBC, Louviere and Woodworth 1983; Louviere 1988; Louviere, Hensher and Swait 2000) experiment for ice cream products that vary in flavor, fat content, organic status and price. We select these particular attributes for a number of reasons. First, all are likely to be important determinants of consumer choice given that consumers are typically price conscious in their purchase of a discretionary item such as ice cream, many are likely to be conscious of the health-attributes of their choice, and flavor is always a key determinant of choice. Second, discrete or qualitative levels of these attributes are readily included in a conjoint analysis. Third, it seems reasonable to assume that the factors are not highly correlated. We use this experiment to provide data on each subject's willingness-to-pay (WTP) for ice-cream attributes within a particular brand: Ben&Jerrys. We choose Ben&Jerrys ice cream because they are known to be innovative, they change flavors often, they have many flavor-names that are not immediately associated with a particular flavor and are widely available in virtually any supermarket. Flavor has three levels (Vanilla, Dulce Delish and Steven Colbert's Americone Dream). These flavors represent one that is widely understood (and by far the most popular by volume share), one that is rising quickly in popularity according to firm sources, but likely to appeal to niche tastes, and one with a name that is impossible to associate with any ingredients, flavors or textures, but likely to have appeal based on popularity alone. Fat content has two levels (low fat and regular) that represent the most common method of differentiating ice cream by modifying nutrient content. Organic status has two levels (organic and non-organic). Organic ice-cream is becoming increasingly popular as the demand for organic foods as a whole rises over time, so consumers are likely to be willing to pay a significant premium for organic ice cream. Finally, price has four levels ($2.00 / pint; $2.50, $3.00 and $3.50) which spans the US average per-pint price for Ben&Jerry's ice cream at the time of writing.

As Kuhfeldt, Tobias and Garratt (1994) note, the sole aim of an experimental design should not be achieving orthogonality as there exist non-orthogonal designs that further minimize the variance of parameter estimates. They argue that maximizing efficiency, i.e., minimizing the variances and covariances of estimates, should take precedence over orthogonality considerations. Therefore, we implement a search algorithm to find the most D-efficient labeled, fractional factorial design. The SAS macro, %mktex, returned a 100% D- and A-efficient saturated, balanced, orthogonal fractional factorial design with 16 choice sets of three choice alternatives plus a "no buy" option. Furthermore, since the label (flavor) appeared in the same order in each set, we randomized the order of the choice alternatives within each set in order to minimize bias that may be introduced if flavor occurred in the same order for every set. The same CBC design is used for both the second-stage and third-stage choices, which we describe below.

The conjoint experiment is made incentive compatible by choosing one choice opportunity at random and making that choice binding. That is, the subjects will be given the opportunity to actually buy the ice cream that they are evaluating at a price to be determined through a Becker-DeGroot-Marshack (BDM 1964) mechanism. In this application, the BDM mechanism works as follows. We choose a price at random from a uniform distribution between the lowest and highest price on the cards ($2.00 and $3.50). If the random price is below the price on the chosen card, then the subject buys the ice cream at the random price. Subjects buy the ice cream by forfeiting an equal amount of their endowment. If, however, the random price is higher than the price on the preferred card, then the subject keeps his or her endowment and does not receive any ice cream.[1] This mechanism is well-understood to be incentive compatible under a wide variety of auction scenarios. The instructions handed out to the subjects carefully explain how the BDM mechanism works, why it is in the subject's best interest to bid exactly his or her willingness to pay (or as nearly as the price-attribute choice allows) and provides an example in a context similar to the ice cream experiment at hand.

In the third stage, we conduct a second CBC in which we send out a survey similar to that used in the second stage, but include recommendations from all other subjects as part of the initial information set. We follow Narayan, Rao and Saunders (2011) and offer the product selected by each subject in the first stage as their recommended choice.[2] Because each subject makes 16 independent choices, we construct a "recommended product" in the following way. After compiling all stage 1 data, we estimate the random parameter logit (RPL) choice model described below in which utility is a function of attributes of the choice, the chooser and an error term that represents the distribution of unobserved heterogeneity. We then use this choice model to calculate the implied utility of each choice. A subject's recommended product is the one that provides the highest level of utility. In this sense, we capture each respondent's "ideal product" as utility is directly a function of embedded attributes. Following the completion of the third-stage CBC, we then implement the BDM mechanism by drawing one choice opportunity at random, drawing a random price between $2.00 and $3.50 and determing which subjects indeed bought their ice cream. Coupons and cash are then disbursed online. In this way, we are able to assess not only the value of the ice cream attributes, but how these values are mediated by relationships with other, potentially significant, social network members.

[1] To avoid keeping ice cream on hand, we distributed coupons for the chosen ice cream at a local Safeway store. Students were told of this fact in the instructions beforehand.

[2] Selecting one product as the recommended product is both realistic and necessary. It is realistic in that product recommendations for relatively simple items are frequently made without reference to specific attributes – witness the "like" button used by Facebook members. It is necessary because modeling recommendations over multiple attributes over even a modest sample size would quickly become intractable.

3 Spatial Models of Social Interaction, Attribute Valuation and Influence

3.1 Overview

In this section, we develop two, complementary models of how social interaction influences product choice. In the first, we extend the model of social interaction developed by Narayan, Rao and Saunders (2011) to consider a number of different ways of measuring social network location and to examine how mediating interaction in social network space with an explicit consideration of attribute space affects how subjects value product attributes. In the second, we consider the inverse problem in which we investigate which individual in the network has the most influence on others' choices. While the former problem is interesting and important, the inverse problem is also of value to marketers interested in identifying influential individuals. These models, however, are built on fundamental spatial constructs for both social space and attribute space. We begin by explaining how we define each before they are integrated into the choice and influence models.

3.2 Social Network Space

Typically, empirical social network analysis research uses tools from graph theory to illustrate and analyze the relationships among members of a network (Hanneman and Riddle 2005). However, in this study we create a synthesis of social network analysis and spatial product differentiation, integrating the two with spatial econometric methods. Anselin (2002) describes the emerging field of social network analysis as one example of how spatial econometric methods can be applied to more general areas of research inquiry. The analogy is clear – members of a network are separated by perceived relationship distances just as products are separated by distance in attribute space or houses in geographic space. Thus, social network data can be analyzed using spatial econometric methods. An explicitly spatial approach, using parametric regression methods, provides valuable information that would not be available if analyzed using more typical graph theoretic software.

The first stage of the online experiment is designed to provide data on the nature of the relationships among our subjects. In formal terms, the responses to these five questions are represented in matrix format to describe the connections between each pair of network members. There are five such matrices, each corresponding to a separate question. Consider one of these matrices for purposes of illustration. Each cell of the matrix below the main diagonal contains the response to the question described above in (row, column) format for how the row individual feels about his or her relationship to the column individual. Above the main diagonal, the matrix element describes how the column member regards his or her relationship to the row member. This matrix does not have to be symmetric as each data point represents each members' own assessment of his or her relationship with the other member. Figure 1 provides an example of a network matrix for four individuals (A - D) using a five-point Likert scale to describe the nature of the relationship between each, where a response of 1 indicates the weakest communication, tie-strength, trust, respect or leadership assessment.

[figure 1 in here]

In mathematical terms, the matrix in figure 1 is referred to as an "adjacency matrix" because it contains information on the relationship between each member and all others. Although the Likert-scale responses are ordinal, consider a set of binary responses (1 = know; 2 = don't know) for illustration purposes. Defined this way, the adjacency matrix becomes an algebraic representation of a graph, or picture of the network. In our experiment, the adjacency matrix is fundamentally asymmetric, that is, just because one agent claims to know or be friends with another, the opposite is not necessarily true. The convention in social network analysis is to represent this asymmetry from the perspective of the row agent: each entry in the row represents how the row agent feels about his or her relationship to the column agent.[3] This form of the adjacency matrix is used to calculate five metrics regarding the degree of connectivity or potential influence of each network member: proximity, centrality, betweenness, farness and core membership (Freeman 1979; Hanneman and Riddle 2005; House et al., 2007). Identifying each individual's role in the social network is essential to the goals of this research as we evaluate each metric to determine the subject with the most evident importance in influencing others' attribute preferences.

First, proximity between two individuals in a network is calculated using a Euclidean distance metric. In the context of our survey, define a response r_{il} such that a "5" response to any of the social relationship questions indicates respondent i is relatively "close" to respondent l, and a response of "1" indicates the two are relatively far. Therefore, the valued adjacency matrix provides direct measures of proximity, which is row-normalized so that the updating mechanism in (7) produces weighted-average values of the initial parameter estimates of all other network members. We define the social distance matrix such that each element measures the relative proximity of i to l using the valued adjacency matrix given by:

$$\mathbf{WS}_p(i,l) = r_{il}, \tag{1}$$

Our proximity hypothesis is that the valuation of attribute k should change toward the valuation of respondents who are deemed to be relatively close, in a Euclidean sense. Proximity, however, does not account for how many other members of the network are known or trusted by any other member.

Second, centrality is measured using the simple count of "adjacencies" or connections for each member of the network. Let $\sum_l a_{il}$ be a simple count of the number of adjacencies between i and all other l members where $a_{il} = 1$ if i and l are adjacent, or connected, and zero if not. In the context of our experiment, we define two subjects to be adjacent if they know, trust or respect each other in a minimal way. That is, a reponse of either a "4" or "5" in any of the social relationship questions on the survey constitutes an adjacent

[3] Asymmetric adjacency matrices can also be represented as directed graphs, with arrows pointing from the source (row) to the target (column) agent (Hanneman and Riddle, 2005).

relationship. A measure of centrality normalizes this count by the maximum possible number of adjacencies in order to create a relative measure of centrality for member i: $\mathbf{WS}_c(i) = \sum_l a_{il}/(N-1)$,where there are N total network members. Intuitively, centrality in the class indicates the student with the largest circle of friends. In the context of new product marketing, Goldenberg et al. (2009) find that the centrality of an individual is more important than their innovativeness in contributing to the success of a new food product. Centrality, however, does not necessarily mean that a particular agent "brokers" many relationships between others. Such brokerage or intermediary services are necessary to maximize the flow of information through the network.

Third, betweenness is measured using the "partial betweenness" index described by Freeman (1979) and Jackson (2008). Partial betweenness builds on the concept of a "geodesic," which is the shortest path between two members of a network. For example, if student A knows student B who knows student C, but A does not know C, then the geodesic is a path two steps in length through student B. More generally, the betweenness of any student l, therefore, is measured by the probability that he or she lies on a geodesic between two others, say i and j. The total number of geodesics between student i and j is given by g_{ij} so, if they are indifferent between which path is taken, the probability of choosing each path is simply given by $1/g_{ij}$. Thus, the probability of a particular geodesic between i and j going through l is: $b_{ij}(l) = g_{ij}(l)/g_{ij}$, where $g_{ij}(l)$ is the total number of geodesics that run through l. Finally, the overall betweenness of l is measured by the sum of all partial betweenness indices for all other network members:

$$\mathbf{WS}_b(l) = \sum_{i=1}^{I} \sum_{j=1}^{J} b_{ij}(l), \ \forall i \neq j. \tag{2}$$

In the context of our experiment, betweenness among classmates measures the student with not necessarily the most friends, but the one with the most links to other classmates – the most interconnected person and, therefore, the most potential to exert influence over the product choices of others. If a person is the only connection between two others, then he or she has considerable power to influence the relationship between these two others. On the other hand, if there is another intermediary, the power of any of them is reduced proportionately.

The fourth measure is farness. Farness, or its inverse closeness-centrality (Hanneman and Riddle 2005), is a general concept that captures several different measures of network distance from one agent to all others. Freeman (1979) measures farness as the total number of interconnections required to get from one member to all others, whereas Hanneman and Riddle (2005) describe a measure that is the sum of the geodesic distances between and agent and all other agents. Intuitively, the farther the distance from one agent to all others, collectively, the less power that agent's product preferences will have over the choices of others.[4]

[4]Eigenvectors of the adjacency matrix can also be used as a measure of the centrality of each agent, but did not perform well in our data.

For our purposes, we use a farness measure in which we sum the total geodesic distance for each agent as a measure of their separation from the network. Farness, therefore, is calculated as:

$$\mathbf{WS}_f(l) = \sum_{i=1}^{I}\sum_{j=1}^{J} h_{ij},\ \forall i \neq j. \tag{3}$$

where h_{ij} is the length of each goedesic, g_{ij}. Farness can be defined in four different ways, depending on whether the researcher is interested in relationships others describe with the agent in question ("in-farness") or those the agent describes with others ("out-farness"), and whether farness or closeness is chosen as the relevant metric. For purposes of this study, and to maintain consistency with the perspective taken with respect to the proximity metric above, we define farness in terms of the perspective of out-closeness.

Finally, core membership is a measure of how correlated the interconnections of a given member are with all others. Information is expected to flow most quickly among members deemed to be in the core of the network (Borgatti and Everett, 1999). In terms of an agent's power to recommend products to others, core members are clearly expected to have more power than peripheral network members. We determine whether an agent belongs to the core or the periphery using the generalized algorithm of Borgatti and Everett (1999). Essentially, their approach uses a genetic algorithm to maximize the correlation between the observed data and an idealized core-periphery structure in which all members of the proposed core claim to be connected to each other, and all other non-core members, and none of those assigned to the periphery know any other members. In notation similar to that used for the other metrics, core membership is found by maximizing the correlation coefficient: $\rho = \sum_{i,l} a_{il}\delta_{il}$, where:

$$\delta_{il} = \left\{ \begin{array}{l} 1\ if\ a_i = CORE,\ a_l = CORE \\ 0\ else \end{array} \right\}, \tag{4}$$

and δ_{il} represents core membership in the ideal structure, while $a_{i,l}$ represents the observed assignment. Based on this assignment, the social weight matrix is defined as: $\mathbf{WS}_r(l) = CORE_i$ depending on the values of δ_{il} that maximize the probability of observing the experimental data. In our experiment, the core consists of 5 members in the smaller sample and 21 in the larger sample. However, because they do not all know each other, nor do all of the peripheral members know them, the maximal correlation coefficient is only 0.31 in the smaller and 0.43 in the larger sample. Despite these relatively low values, they are consistent with others reported by Borgatti and Everett (1999). Consequently, we use each of these five metrics to represent ways in which an agent may update their initial valuations of each product attribute, based on the estimates of all others. We calculate centrality, betweenness, farness and core membership for each of the five social-relationship questions in the survey and interact them with the measures of attribute proximity developed next.

3.3 Attribute Space

The degree of differentiation among the ice cream variants offered in our experiment depends on their location in attribute space. More precisely, each consumer forms a perception of the extent to which an ice cream is differentiated from others based on its distance from all others. Distance in attribute space is defined in the same Euclidean terms as that introduced above. Because utility depends not just on the levels of attributes, but on the distance each choice is from the others in attribute space (Feenstra and Levinsohn, 1995) our hypothesis is that attribute valuation will depend not only the recommendations of others, but where the product lies in attribute space.

Differentiation is defined in terms of the distance each product lies from the others. In attribute space, inverse Euclidean distance, or proximity, is a continuous measure defined as $\mathbf{WA}_{jl}$ between the attribute profile of item j and item l such that:

$$\mathbf{WA}_{jl} = \left(1 + 2\sqrt{\sum_k (n_{j,k} - n_{l,k})^2}\right)^{-1}, \tag{5}$$

where $n_{j,k}$ is the amount of nutrient k in item j and similarly for item l. For this application, we define the set of attributes to include fat content (grams per 1/2 cup serving), organic status (organic or non-organic), and flavor (Vanilla, Dulce Delish, Steven Colbert's Americone Dream). Because equation (5) is defined it terms of inverse-distance, it represents a measure of how close the items are in attribute space. We then define a $J \times J$ distance matrix, $\mathbf{WA}$, that includes all of the w_{jl} as elements and thus describes the distance between each pair of products in the sample. The main diagonal of $\mathbf{WA}$, which measures the proximity of each product to itself, is normalized to 0.0, and all entries are row-normalized (Yang and Allenby 2003) so that the matrix-products with the initial parameter vector results in a weighted-average of all other subjects' estimates. This point highlights the value of using a random parameter approach wherein each subject has his or her own attribute-valuation parameter vector. Clearly, $\mathbf{WA}$ is symmetric as the distance between j and l must be the same as the distance between l and j. By explicitly recognizing the extent of differentiation among the products in our experiment, we add an additional degree of flexibility to a standard random-parameter logit demand model in that substitution between products reflects not only preference-similarity among subjects, but the relative location of each variant as well (Slade 2004).

Modeling product differentiation in this way is not simply an empirical convenience, but is grounded in utility theory as the distance between products in attribute space is a primative of the model and must be a part of the choice model (Anselin 2002). As a primitive of utility, spatial differentiation must be included directly in the utility function as the demand for an ice cream product at one point in multi-dimensional attribute space depends on the demand for another product at a different point in space (Feenstra and Levinsohn 1995). This utility specification makes intuitive sense as it is likely that the probability an individual chooses an ice cream depends on the observed differences between all available alternatives. Consider the utility a consumer obtains from an assortment of ice cream that consists of only French Vanilla and Vanilla Bean relative to the potential utility from a choice set of full and low-fat Rocky Road, Vanilla Bean and Strawberry. Because the latter more nearly spans the potential space of attributes, the consumer is more likely to obtain greater utility from whatever choice is made. How accounting for attribute space interacts with preferences of others in the social network, however, is an empirical question. Based on previous social network research (Godes and Mayzlin, 2009; Narayan, Rao and Saunders, 2011), however, we expect that preference parameter values will be influenced in the direction of those with most power or influence in the network.

3.4 Social Interaction and Attribute Valuation

Our first model of demand investigates the interaction between an individual's location in a social network and a product's location in attribute space in determining discrete product choices. This model is intended to determine how different measues of social-network location affect an individual's own attribute valuation. We develop a model of the inverse problem in the next section: How an individual's location effects others' choices. We begin by describing a discrete choice model of demand in general notation, and then show how we modify this model to incorporate social and attribute space. In these extensions, we model the centrality of choosers, and the proximity of their choices, using constructs from spatial econometrics. We use a Bayesian updating mechanism to link these two concepts of space in a meaningful way.

In a discrete choice model, consumers choose the one product from all alternatives that provides the highest level of utility. Utility is assumed to be random, reflecting the underlying heterogeneity of consumer preferences. Attributes, defined narrowly as the set of nutrient attributes for this example, have both a direct effect on demand as arguments of the utility function (Berry, Levinsohn and Pakes 1995), and an indirect effect by defining a product's location relative to all others. Formally, mean utility for consumer i from purchasing item j in week t (the time subscript is suppressed below) is a function of a set of choice attributes ($\mathbf{x}_j$), chooser attributes ($\mathbf{z}_i$) and unobservable factors. There are $i = 1, 2, ..., N$ consumers choosing from among $j = 1, 2, ..., J$ products in our experiment. Utility for consumer i, therefore, is written as:

$$u_{ij} = \mathbf{x}_j^{\intercal} \boldsymbol{\beta}_i + \mathbf{z}_i^{\intercal} \boldsymbol{\gamma} + \varepsilon_{ij}, \tag{6}$$

where consumer heterogeneity is reflected in both the vector of random parameters, $\boldsymbol{\beta}_i$, and the random error term, ε_{ij}. As explained above, product attributes consist of price, fat content, organic status and flavor. Unobserved heterogeneity in the marginal utility of each attribute is assumed to reflect subject-specific factors that remain after controlling for observed heterogeneity so that: $\beta_{ik} = \beta_k + \sigma \nu_i,\ \nu_i \sim N(0, 1)$ where β_0 is the mean price-response over all subjects, σ is the dispersion of heterogeneity in attribute response, and ν_i

is a normally-distributed random **variable** [was factor]. The error term ε_{ij} is assumed to be extreme-value distributed, so the choice probabilities are of logit form. We estimate the resulting random-parameter logit (RPL) demand model using simulated maximum likelihood (Train 2003) as there is no closed-form expression for the choice probabilities. This demand forms the basis for our spatial model of demand incorporating social and attribute space.

Consider the location of subject i in social network space. In general notation allow $\mathbf{WS}_p(i,l)$ to reflect the proximity of subject i to subject l such that the opinion of l regarding the desirability of the product is valued more highly by i the greater is $\mathbf{WS}_p(i,l)$. For example, if the metric is a subjective measure of "friendship" then the higher the closer friends i and l are, the greater the effect of l's opinion on the choice of i. We model this effect by allowing each subject's marginal attribute valuation to depend on how close friends are with others in their network, and the attribute valuations of their friends. Intuitively, each subject has an initial or native response estimated using the first-stage experimental data, which is then updated given the attribute valuations of all others, and their relative location in the social network. Each subject's initial response is updated in a Bayesian manner as a distance-weighted function of the attribute preferences of all others in the social network. We model the second-stage or updated parameters using a Bayesian updating rule similar to that developed by Narayan, Rao and Saunders (2011) in the following way. Define the first-stage valuation of attribute k as β_{ik}^I to reflect the initial valuation of attribute k by subject i. The second-stage attribute valuation is updated according to:

$$\beta_{ik}^R = \rho_{ik}\beta_{ik}^I + (1-\rho_{ik})\left(\frac{\sum_{l=1}^{N}\mathbf{WS}(i,l)\beta_{lk}^I}{\max(\sum_{l=1}^{N}\mathbf{WS}(i,l)\boldsymbol{\eta},1)}\right), \tag{7}$$

where $0 < \rho_{ik} < 1$ reflects the weight given to others' responses in revising own attribute valuations and $\boldsymbol{\eta}$ is an $I \times 1$ unit vector. The term in parentheses represents the relative magnitude of all other individuals' attribute valuations, weighted by how "close" each individual is to subject i, while ρ_{ik} reflects the degree of importance assigned to *all* other valuations by this subject. Higher values of ρ_{ik} mean that individuals are less affected by the valuations of others in their social network, or at least information regarding others' valuations. This, in turn, may mean that subjects place little stock in others' choices, or that they are simply not close to anyone else in the network. Based on the experimental design described above, this updating mechanims relies on subjects' abilities to infer the individual attribute valuations from recommended products. If one subject, for example, recommends a high-fat, organic, Americone Dream flavor, then it is not unreasonable to expect another subject to infer that this individual places a high value on more indulgent properties, with unique-sounding names that are produced in an environmentally-sensitive way. It is the function of the social weight matrix to mathematically select those individuals' choices that are most important to the individual in question. In the extreme, if for subject A all other individuals but one have a social weight of 0, then the referent individual for subject A are irrelevant and the only choice that is evaluated is that of subject A. We can then reasonably expect this other individual to have an important effect on the second-stage choices of subject A. How we define social proximity in the context of this model is, therefore, of paramount importance. Next, we begin adding detail to this general model by describing several metrics of an individual's location in social network space, and how this location can affect demand, and then show how this model is generalized to include product location in attribute space.

3.5 A Model of Influence

In this section, we describe a model of social influence, or how influential individuals in a social network can be identified using econometric methods. Our model is intended to determine how an individual's position in the social network influences the choices of others. A well-connected member of the network need not be particularly influential. Rather, Goldenberg et al (2000) argue that influence derives from (1) respect for the individual not connected to any specific competence (or "who one is," (Weimann 1991)), (2) the member's posession of knowledge relevant to the product in question, (3) "market mavenism" or market knowledge that is associated with a high level of involvement with the product category, or (4) "social capital" or the member's cumulative investment in developing relationships with others in the network. While measuring these qualitative attributes is typically impossible in an empirically-meaningful way, they each describe individual-specific factors affecting influence that can be controlled for in estimating the relative importance of social network location.

In this model, we are less concerned with attribute valuation than we are with product choice. Intuitively, our approach is as follows. We first define how one's influence on others is to be measured using the two-stage choice experiment data described above. An individual is deemed to be influential if they are able to move other subject's choices closer to their own, where closer is defined in terms of attribute space. To implement this definition, we calculate the distance between each subject's choices and all others' choices from the first to the second stage data. We then calculate the difference between the first-stage distances and second-stage distances and calculate an average distance for each subject. Next, we need to determine how influence is related to each subject's position in the social network. Simply ranking subjects according to their influence is not sufficient because it does not show how influence is related to one's position in the social network. Therefore, we estimate a simple model in which influence is a function of an individual's network location, specifically how close they are to others. Allowing the marginal effect of social proximity on influence to vary by individual permits a ranking of network members' effect on others' decisions, controlling for their location in the network. Our hypothesis is that a relatively influential individual is able to move others' choices closer to their own between the two stages.

More formally, subjects in the experiment make a discrete choice among ice creams in both the first- and second-stages. Assume a simplified characteristic-demand framework (Berry and Pakes 2007) in which the utility obtained from choosing a product in the first stage depends on its price and attributes only. Because **we compare only 2 individuals at a time and each chooses** only one product, we use the same index for the subject and product. Assume one subject chooses product k with attribute vector (including price) $\mathbf{x}_k$, while another chooses product l with attribute vector $\mathbf{x}_l$. Utility for the first subject is, therefore, written as: $u_k = \mathbf{x}_k^{\intercal}\boldsymbol{\beta}$ while utility for the second is given by: $u_l = \mathbf{x}_l^{\intercal}\boldsymbol{\beta}$. With the characteristic-demand model, we project each choice into attribute space so the difference between the choices of two subjects is measured as a distance in attribute space. The distance between product k and l is expressed compactly in matrix notation by defining distance in Euclidean terms or, to follow the convention introduced in the spatial attribute model above, inverse Euclidean distance, or proximity. With this definition, the (k, l) element of the proximity matrix measures the closeness of the choices of two individuals and is defined as:

$$\mathbf{WA}_{kl}^1 = \left(1 + 2\sqrt{\sum_m (n_{k,m} - n_{l,m})^2}\right)^{-1}, \tag{8}$$

where $n_{k,m}$ is the amount of attribute m in choice k and similarly for choice l and the superscript refers to the first-stage choice. The proximity matrix used here is the same as that used in the previous model, but is used in a different way: We are not concerned about the impact of others' recommendations on individual attribute valuations, but rather how these recommendations are associated with different product choices. Matrix notation is convenient as we are not only interested in the proximity between two subjects, but between all pairs of subjects. The proximity between all pairs of individuals, therefore, is given by $\mathbf{WA}^1$. With this notation, the vector of proximities between the choices of each individual and the choices of *all* other individuals, therefore, is found by simply taking the average of the pairwise-proximities, or: $\overline{\mathbf{WA}}^1 = \mathbf{WA}^1\boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is an $N \times 1$ vector with elements: $[1/N, 1/N...1/N]$.

In the second stage, subjects again choose between products that differ in embedded attributes, but they are also informed of the choices of others. The average proximity between each subject's choice and all others is calculated in the same way as in the first stage and is written: $\overline{\mathbf{WA}}^2 = \mathbf{WA}^2\boldsymbol{\delta}$ in vector notation, which is again interpreted as the average of the pairwise proximity for each subject at the second stage.

Calculating attribute distance at the first and second stages provides a natural means of defining influence, and testing our theory of influence in social networks. Influence, in this framework, is defined as the ability of one subject to move another closer to his or her choice. Formally, define $\mathbf{INF} = \overline{\mathbf{WA}}^2 - \overline{\mathbf{WA}}^1$ as a vector of influence measures. $\mathbf{INF}$ measures the influence of each subject as large values suggest that others have moved closer to the subject's choices between stages 1 and 2 of the choice experiment. Large values of $\mathbf{INF}_k$, or the influence level of the k^{th} subject, may, however, simply reflect the interconnectedness of the subject in question. Consequently, our hypothesis regarding influence is that, controlling for network location, more influential individuals will be those who appear to be able to move others' choices toward their own.

Testing this hypothesis involves estimating a model in which we control for a subject's location in the network in order to estimate individual-specific influence parameters. Controlling for unobserved heterogeneity through a random-parameter estimation method, and a random individual-effect that controls for observed changes in choices that are not related to influence, we estimate subject-specific influence paramters that measure how each member's proximity in the network is associated with their level of influence. Because our spatial construct is similar to a difference-in-difference metric, any general trend toward picking a particular product in the second stage is removed through the differencing procedure.

More formally, influence is regressed on the mean of our preferred measure of social network location above: proximity, or $\overline{\mathbf{WS}_p(i,l)}$. By allowing the spatial coefficient in this model to vary by individual in a random-coefficients regression framework, we not only account for an important source of unobserved heterogeneity, but we obtain a vector of parameters that measure how proximity and influence are related. By ranking these individual estimates, we obtain a measure of who has the most power to influence others in the network. More formally, this model is written as:

$$\begin{aligned}\mathbf{INF}_i &= \lambda_{0i} + \lambda_{1i}\overline{\mathbf{WS}_p(i,l)} + \varepsilon_i, \\ \lambda_{0i} &= \lambda_0 + \sigma_0\nu_{0i},\ \lambda_{1i} = \lambda_1 + \sigma_1 v_{1i,}\ \nu_{ki} \sim N(0,1).\end{aligned} \tag{9}$$

where λ_{1i} is a spatial-correlation parameter (in social network space), here interpreted as measuring how much influence derives from location, while λ_{0i} is the autonomous level of influence that remains after controlling for network location. Among the other elements of the model, ε_i and ν_i are iid normal random variates. If the rankings according to λ_{0i} differ from the rankings according to proximity, we conclude that mere proximity is not a complete measure of importance within the social network, but influence is.

3.6 Estimation Methods

Each of the choice models described above is estimated using simulated maximum likelihood (Train, 2003). Although the logit assumption typically leads to closed-form expressions for the choice probability of each household, the random parameter approach means that no closed form expressions exist. Further, we estimate the models sequentially, beginning with the attribute-valuation model, and then adding the social network- and attribute-space models, in that order. While we realize that the choice of social network-space first and attribute-space second is arbitrary, we estimate each model in reverse and found the results to be qualitatively similar. We also estimate each model with variations on the question used to construct the adjacency matrix. Namely, we use a question that asked each subject to respond with "how well they know" each of the other subjects. This question is consistent with Narayan, Rao and Saunders (2011) and follows Godes and Mayzlin (2009) in construction. Other questions that asked subjects to assess the trusthworthiness of others in the network, their leadership abilities and the frequency of communication with other subjects. The results for each of these other adjacency matrices is similar to those presented below and are available from the authors.

We also estimate the random-coefficient influence model using simulated maximum likelihood. In this model, both the individual-specific and social network-proximity parameters are assumed to be distributed normal to allow for the possibility that both effects can be either positive or negative.

4 Results and Discussion

We begin this section by describing our experimental data in summary form, and then present and interpret the results of estimating attribute-valuation models without any spatial effects, by including social network space, and then by interaction social and attribute space. We then draw a number of implications regarding the practical importance of our findings.

4.1 Data Summary

Our social networks consist entirely of advanced undergraduates in business schools at a major research university in the US Southwest and at a smaller university in California. Therefore, we expect the demographic and socioeconomic profiles of these samples will reflect their local populations, and to not represent the more general population. Social networks, by definition, describe only a small part of the US population so will, in general, not be representative. In this section, we summarize the data from the larger sample. A similar summary for the smaller sample is available from the authors. Table 2a shows that our average subject is 20 years old, more than likely to be male, belong to a household consisting of 1.9 people and make nearly $46,000 in annual income. The typical subject is also white, a junior, single, intending to complete a master's degree and be a solid B student ($3.25 < \text{GPA} < 3.74$). A total of 73 subjects completed all parts of the experiment. While surveys and experiments generally use much larger samples, sample sizes for social network experiments are typically much smaller (see Narayan, Rao and Saunders 2011). Small sample sizes are necessary due to the response-burden placed on experiment subjects - each person has to evaluate how they know each of the other members of the sample. Response times and, therefore, inaccurate responses increase with the square of sample size. Based on the econometric results below, we find that $N = 73$ was sufficient to identify the parameters of primary interest.

Regardless of sample size, social network experiments are only valid if there is sufficient interlinkage among experiment subjects. Table 2b summarizes the nature of the social relationships among our sample subjects. Whether the existence of a "social relationship" influences product choice, however, may depend on how that relationship is defined. Therefore, we use five separate questions to assess how each subject describes his / her relationship with all others: frequency of communication, extent of acquaintance, trust, respect and leadership. Table 3 shows that a significant number of students do not know any of the others, do not interact frequently, but are able to provide opinions on their trustworthiness, earned respect and leadership abilities. We suspect that this seeming inconsistency arises from the extensive group and casework assigned to these students. While they may not be acquainted with others on a personal level, they form opinions based on relationships developed through purely class-related exercises. Ultimately, whether these social relationships are important are reflected in whether they influence product choices.

[tables 2 and 3 in here]

If social network membership has no effect on product choice, the first-stage and second-stage choices should be identical. Table 4 shows that they are not. Comparing the first- and second-stage choices, we find that subjects in the second stage were significantly more likely to choose the lowest-priced product (\$2.00), the Americone Dream or Dulce Delish flavor, both in organic form. We expect peer-influence to be an important factor in the evaluation of flavor and organic attributes because their effect on utility is inherently subjective. We included the Americone Dream flavor in order to provide a choice that subjects are not likely to be familiar with, and unable to form expectations as to whether they will like the flavor or not based on the name alone. Americone Dream suggests no particular flavor, nor any specific ingredients. Dulce Delish, on the other hand, is a variant on a popular flavor that subjects are more likely to be familiar with, but nonetheless not a common choice (B&J website). Further, products with credence attributes such as "organic" or "natural" are also likely to be strongly influenced by peer opinion because such choices are regarded as more socially-acceptable than conventional choices. Whether social influences are statistically significant while controlling for all other effects, however, requires a formal econometric model.

[table 4 in here]

4.2 Social Network Space Estimates

Results from both the choice and influence model estimates are presented for the larger-sample experiment. Again, estimates from the other sample are similar and are available from the authors. Estimates of the econometric models are shown in tables 5 and 6 below. In table 5, we show the initial marginal value of each attribute, estimated from the stage 1 choices with the RPL model, and how attribute values change after each subject is exposed to choices made by others in the social network. The same model is estimated five times with the stage 2 data, once for each alternative measure of how the subject relates to others in the network. For each model, we estimate random values of ρ_{ik} – for each subject i and attribute k – and then calculate the implied value of β_{ik}^{R} based on the intial β_{ik}^{I} value and the weighting-parameter estimate. Comparing goodness-of-fit statistics between these models provides a way of determining which metric yields the best information on how network relationships impact demand. All models in this table are based on the "degree of acquaintance" question from the survey, but others yield similar results. Before interpreting individual parameter estimates, note that the standard deviations for each of the random parameters are significantly different from zero. This suggests that the RPL model is preferred to a simple-logit alternative. Interpreting the mean values of the individual-specific attribute values, the results in table 5 show that price has a significant negative value, as expected, and flavor has a positive value.[5] Somewhat surprisingly, both regular fat and organic ice creams have lower marginal values than low-fat and non-organics, respectively, albeit at a lower level of significance than the other two attributes. These estimates, however, may differ depending on the choices of others in the network.

[table 5 in here]

Model 1 measures the location of each subject in the social network by his or her distance, as measured by the valued-responses to the second question in table 3 (degree of acquaintance, or proximity). In this table, a higher value of ρ_{ik} indicates that the revised attribute value β_{ik}^{R} is relatively dependent on the own initial estimate, and less so on the spatially-weighted average of others' marginal values. Because the weight placed on own initial estimates of the price parameter is relatively high (0.831), and choices throughout the network tended toward lower-price items, the mean revised price sensitivity is relatively close to the initial estimate, but moves toward greater price sensitivity. On the other hand, subjects placed a relatively low weight on initial flavor choices ($\rho_{ik} = 0.369$), but the strong general preference for the Americone Dream flavor (see table 4) led to only slight revision in flavor choices. In terms of fat valuation, subjects again placed a relatively low weight on own-choices ($\rho_{ik} = 0.131$), so there are apparently a sufficient number of subjects with a negative view of fat content to move the revised estimate below zero. This result is surprising because, on average, subjects chose regular fat ice creams more often in the second stage relative to the first (from table 4, regular fat is coded as 0 and low fat coded as 1). However, this estimate is not significant at conventional confidence levels. Because organic status is more of a credence attribute than the others, we expect a comparatively strong influence from the choices of others. This is indeed the case as the weight placed on the own initial marginal valuation is 0.307, causing the revised marginal valuation to fall from 0.227 to 0.177. Although the average respondent places a positive value on organic status both before and after seeing the choices of others near themselves in the network, we can infer from these results that referent individuals, or those near to a large number of subjects, appear to prefer non-organic ice cream. This result is consistent with the choice data summarized in table 4 as subjects chose organic ice cream less frequently in stage 2 relative to stage 1.

Model 2 captures the same dynamic as Model 1, but as a function of the betweenness of each subject rather than their proximity to others. Betweenness will be important if the agent moderates relationships between others, but is not necessarily "close" as measured by the proximity score in Model 1. With respect to the price parameter, we find that the estimated weight placed on initial estimates is slightly lower than in Model 1 ($\rho_{ik} = 0.761$), but the revised price-response estimate is significantly greater in absolute value than

[5] The three flavors were coded: Americone Dream = -1, Vanilla = 0, Dulce Delish = 1 in the choice-based conjoint model.

in Model 1. This finding suggests that some subjects with particularly high betweenness scores are unusually price sensitive. Among the other parameters, the weight placed on other members' choices is higher for flavor, fat and organic than in Model 1. With respect to flavor, the importance of betweenness is reflected in the observation that subjects' marginal flavor valuation moves more in the direction of Americone Dream than in Model 1 for only a slightly higher weight estimate. Similarly, subjects were more willing to revise their estimates for fat content and organic status after viewing the choices of subjects with high betweeness scores, but the ultimate effect on each estimate is relatively small.

Measuring network location by the subjects' centrality, we find a further strengthening of price response, and a lower reliance on others' choices relative to Model 2. Mathematically, this result occurs when the spatial weights in (7) are relatively low. Intuitively, subjects may rely on the choices of others to a lesser extent, or others may simply be less important by construction. Further in Model 3, the flavor results in table 5 show a similar network-weight to that found in Model 2, and the revised marginal valuation is nearly the same as well. With respect to fat content, the weight on initial estimates is relatively low ($\rho_{ik} = 0.313$), but the revised marginal valuation turns from positive to negative, as shown in Model 1. The results with respect to organic status are similar to those found in Model 1 as well. Although providing a goodness-of-fit similar to the other four models (ranked 4th out of 6 by log-likelihood function (LLF) value), the farness measure used in Model 4 provided only one significant ρ_{ik} value. Consequently, it is difficult to interpret any of the marginal valuation parameters with any degree of confidence, although the revised parameter estimates are broadly similar to those found with the centrality measure in Model 3. In the final model (Model 5, core membership), however, we find a number of significant results. First, this model provides the best fit to the data, measured by both the coefficient of determination and the LLF value. Second, if a subject is a core member, then others in the core will have a disproportionate effect on his or her valuation. Based on this measure, we find a relatively high weight on the own initial price-parameter estimate, so the revised price response moves only slightly toward others' in the network. Flavor and organic show moderate estimates of ρ_{ik}, with results similar to the proximity model: greater reliance on initial estimates is out-weighed by a combination of large structural weights on others' estimates and marginal valuations by referent others that differ sharply from the own-estimates.

Overall, we find that two models: core membership (Model 5) and the proximity model (Model 1) provide significantly better fit to the data relative to the other models, as measured by both the coefficient of determination and the LLF values. Simple measures of relative location such as proximity are likely to be most useful when the product is comparatively simple, and when consumers rely less on word-of-mouth for choices than individual tastes. Whether this is the case when we moderate the social network effect with measures of attribute differentiation, however, remains to be determined. Products that are more different, or more complex, may make social relationships more important.

4.3 Attribute Space Estimates

In this section, we examine whether the patterns that emerge from the results in table 5 are moderated or accentuated by combining social and attribute distance. These results are shown in table 6. Because of the volume of individual parameter estimates over all five models, we will summarize those that we regard as most important or salient to our findings. First, note that the general pattern of network influence for Model 1 (proximity) remains when we include attribute distance. That is, subjects tend to place greater weight on own choices with respect to price, but more weight on the example set by others with respect to attributes that reflect more subjective, experience or social influences. Particularly in the case of organic status, we find that accounting for distance in attribute space causes a strong reliance on others' preferences relative to own preferences. In Model 2 (betweenness), including attribute space leads to a significantly lower weight on initial estimates relative to the results in table 5, and a corresponding accentuation of the revision in marginal valuation due to social network influences. Again, this finding reflects the more general result that allowing for product differentiation leads to a greater role for the suggestions of others and reduces the importance of own preferences. In Model 3 (centrality), we find the most significant changes in ρ_{ik} again occur with respect to price and organic status. Unlike the betweenness case, however, we see a movement toward less reliance on the network and a marginal valuation revision closer to initial estimates. Centrality is more a measure of location in the network than connectivity, so this finding suggests that measures of location are likely to suggest a lesser reliance on others' choices when product attributes are introduced. In contrast to the social network model, the measure of farness (Model 4) produces significant ρ_{ik} values for all four attributes. For all attributes other than flavor, Model 4 produces ρ_{ik} estimates that suggest a greater reliance on others' choices than in the social network model without product attributes. Both price sensitivity and the revised flavor valuation show strong revision when subjects are exposed to others' choices. Among these changes, the most significant revision occurs with respect to the price parameter, which shows a strong movement toward greater price sensitivity. Clearly, the distance weights in the farness matrix ($\mathbf{WS}_f(l)$) tend to magnify the importance of others' choices when differentiation is taken into account. Finally, the Model 5 (core membership) results again show a significant change in ρ_{ik} to place more weight on other's choices, and movement away from each subjects' own initial estimates. Recall that choices in the core reflect the choices of others in the core more strongly than those of peripheral members. In this case, we again find that attribute distance increases the importance of word-of-mouth, both in terms of where subjects turn for guidance, and in how they value each attribute.

[table 6 in here]

Because there is no theory to guide the selection of one measure of network membership over another, social network analysis provides an embarassment of riches in terms of model alternatives. Therefore, goodness-of-fit statistics are useful in selecting which model provides the most useful measure for the purposes at hand.

Similar to the social network-only models in the previous section, the model that provides the best fit in the attribute-space case is Model 1 (proximity) according to both likelihood ratio and R^2 measures. Although the core model also performs well in both the social network and attribute-space applications, we regard the proximity model as providing the best information to guide our conclusions regarding the impact of social network membership on product choice, primarily because proximity provides a more concrete measure of the strength of social ties in the network. Despite evidence from some of the lesser-fitting models that subjects do indeed rely on others for attribute information, the preferred model suggests that social relationships are more relevant for subjective preferences regarding flavor, fat and organic status while home-grown or personal values are more important for more objective attributes such as price.

4.4 Identifying Important Individuals

Understanding the importance of social network relationships in product choice and attribute valuation begs the central question behind this research: how can we identify individuals that exert the greatest influence within the network? After reviewing each of the models above, we determined that the most useful metric of social network location is proximity. Proximity, however, is not the same as influence (Goldenberg et al. 2009). An individual may be "close" to others by having many ties, or be otherwise central in the network, and yet not exert much influence over other choices. Influence is a combination of location and personal attributes such as the ability to command respect, inspire admiration or create followers. Therefore, we use the results from estimating equation (9) above to determine each subject's influence on others' decisions.

Parameter estimates for (9) are in table 7. First, note that the model appears to fit the data well as the chi-square value, which compares the estimated log-likelihood function value to an alternative in which all parameters are restricted to zero, is 669.330. With two degrees of freedom, we easily reject the null hypothesis that all parameters are equal to zero. Second, note that all of the estimated parameters are strongly significant and of the expected sign. That is, there appears to be both a significant autonomous component to influence as well as an element that is positively related to social network proximity. The estimated mean autonomous effect, λ_0, implies that, on average, 16.5% of the influence exerted by one individual on others in the network, is not related to either social network location or unobservables (calculated as the parameter estimate divided by the mean influence value). In comparison, the social network effect ($\lambda_1 = 0.055$) means that 8.8% of influence is due to an individual's proximity to others in the network. Finding highly significant scale parameters for both random-coefficients means that there is considerable amount of unobserved heterogeneity in the exercise of influence among social network members. Nonetheless, our estimate provides a first explicit calculation of exactly "how important" influential members of a social network may be.

[table 7 in here]

We compare proximity and influence rankings to make this point more concrete. Rankings for the top

15 network members in terms of both proximity and influence are shown in table 8.[6] Recall that higher values of the proximity index are associated with members who know more of the other members well. By this measure, member #5 is clearly closest to more members than any other by a considerable margin. Members #70 and #1 also appear to be more closely-related to others, but are far below #5 according to the proximity metric. According to the influence rankings, however, we see that member #5, who was ranked first by proximity, is only the 22^{nd}-most influential, while the most influential member is member #50, who is only the 51^{st}-ranked subject according to proximity. Indeed, the correlation between influence and proximity is only 0.237, which is not statistically different from zero in our sample (t-ratio = 1.198). The influence rankings in table 8 show that many individuals who are not necessarily well-positioned in the network are nonetheless very influential. Many members who are not even in the top 15 ranking in proximity are among the top-ranked members in terms of their influence. Further, while the empirical measures of connectivity differ substantially from one member to the next, empirical estimates of influence exhibit greater clustering. This finding, which is important for practical application of our approach, suggests that influence is a more subtle construct than connectivity and the differences in influence among individuals are perhaps not as stark as believed *a priori.*

[table 8 in here]

In the choice model estimates shown above – both with respect to social network and social-attribute space combined – we demonstrate the practical importance of how social network space is defined by allowing the revised parameter estimates to vary across different definitions of connectivity. If the influence model is to have value as a practical tool for identifying influential individuals, it is important to examine the sensitivity of influence to how network interactions are defined as well. To that end, we conduct a series of counterfactual simulations using (9) under each alternative definition of social network space. These simulation results are shown in table 9. We generate simulated influence values by calculating fitted values for the influence model, including both the autonomous and network-dependent parts, for each social weight matrix. These estimates show both broad agreement across the different definitions, and some important differences. Note that subject #43 ranks first in "total influence" for the proximity and betweenness definitions and ranks sixth in both the farness and core definitions, but does not appear in the top 15 with respect to centrality. However, subject #37 is ranked in the top five using all five definitions of social space. Similarly, #29 is highly-ranked under all alternatives. In fact, the pairwise-correlations between each set of rankings is over 99%, suggesting that there are far more similarities than there are differences in the entire set of rankings. Managers who may want to use our influence model, therefore, should recognize that there are important differences in identifying influential individuals with different definitions of interdependence, but the likely

[6] The proximity index is calculated by expressing the total Euclidean distance from each member to all others as a ratio to the mean distance of the group. Member #5, therefore, is interpreted as being 30% closer to all other members compared to the mean. The influence index is the estimated parameter from the influence model.

error in rankings is small.

[table 9 in here]

The implications of our findings are important for both academic and practical reasons. Combining data from a social network experiment with explicitly spatial econometric methods provides a way of identifying influential individuals in a network. While others in the literature answer the important problem of how an individual's location in the network affects their response sensitivity (Narayan, Rao and Saunders 2011), our approach goes a step further to investigate the inverse problem – how an individual's location impacts others' choices. On a practical level, applying our method to large-scale social networks would provide firms the ability to more efficiently and effectively design marketing messages, or price promotions. Our method of identifying important individuals also has the benefit of simplicity. Given data on proximity such as that collected here, it is a simple matter to estimate equation (9) and rank consumers according to their level of revealed influence. Instead perhaps give a bit more detailed response on how managers can use the methods outlined to identify those consumers to target who will shift other's choices. For example, Marketers are often concerned with those consumers that have the ability to shift other consumer's preferences. These consumer's are often termed 'influential' but as we have shown the specific definition of 'influence' is of the utmost importance. Based on the results of our study managers should be targeting those consumers that are central to the network, even if those consumers do not necessarily rank high using standard 'influence' metrics as these consumers are more likely capable of shifting other's preferences. What's more, knowing how exactly to find those consumers to target, mangers and markerers alike can avoid the often expensive and time consuming social network-purchase habit study and instead just find those consumers central to the network using secondary data from Facebook or Twitter perhaps.>>

5 Conclusions and Implications

In this study, we consider the interaction between social network membership and location in attribute space on preferences for differentiated food products. We consider how network membership impacts a member's preference for attributes and how network membership by one member affects the preferences of others. By addressing these two, complementary issues, we provide a practical means of quantifying interpersonal effects in a social network, and determining who are influential individuals.

We design and implement a two-stage choice-based conjoint study in order to gather data on both the structure of the social network and how network relationships affect product choice and attribute valuation. The subject matter of our experiment consists of ice cream products that vary in price, flavor, fat content and organic status. In the first stage, we gather demographic information and data on how well each sample subject knows each of the other network members. We also ask respondents to make a number of choices from among 16 choice sets of different ice cream-attribute combinations. In the second stage, we reveal the choices of all network members to all others and ask them to make a second set of product choices. In answering problem (1), our econometric approach also proceeds in two stages. We estimate marginal attribute values implied by the first-stage product choices, in the absence of any social network influences, and then re-estimate the choice model using the second stage data on the assumption that subjects' marginal attribute valuations will reflect their exposure to others' choices. By constructing five different social relationship (adjacency) matrices reflecting proximity, betweenness, centrality, farness and core membership, and estimating separate second-stage models for each we are able to determine both the most useful measure of network structure, and how this structure influences preferences. The empirical model for problem (2) involves constructing a metric of how all others' choices change between the first and second-stage experiments and then using subject-specific estimates of autonomous and network-related influence to explain observed changes in others' behavior. The most influential individual is the one who is not necessarily the most connected, but the one who is associated with changes in others' choices.

We find that simple measures of network proximity provide the best fit to our sample data. Using the results obtained from this model, our results show that social network effects can have significant impact on consumer choices and attribute preferences, both in an economic and statistical sense. Social network effects are most important, however, in attributes that are more subjectively evaluated, such as flavor and organic status, while personal preferences remain important for attributes that are objectively measured such as price. These effects are accentuated when attribute space is interacted with choice data from social network members. We also find that the most connected network member is not necessarily the most influential. Rather, influence derives from unobserved consumer attributes that may include others' respect, trust or admiration for the member's choices. In other words, marketers can determine those individuals that are most likely to influence other's choices simply by finding the consumer who has the most network connections.

There are many implications of our research, both for future research in this area and for practitioners. For others interested in the economic analysis of social network data, our study shows that it is important to include both the location of the chooser and the choice. We have long understood that location in attribute space is an important determinant of choice, so social network analysis simply allows us to layer another determinant of choice on our existing models.

One weakness of our study is that we cannot guarantee that our results generalize beyond ice cream. While we chose ice cream purposefully to consider how imporant social network relationships are for a packaged good category, future studies may consider products with more attribute complexity, or a greater range of products. Food marketers are perhaps the most active users of social network marketing methods so it is important to identify how they can better use the tools at their disposal, but greater attribute complexity would likely make social network relationships more important. By choosing an inherently simple product for the subject matter of our research, we purposefully avoid testing hypotheses regarding why network members influence others, but future research may want to combine our methods with more complex products to examine the specific mechanism at work. Finally, our approach provides a way to both estimate the social dimension of individual-level attribute valuation parameters, as well as a means of estimating network influence directly. Future research is necessary to explore the second of these problems in a more complete way, given the number of ways in which network connectivity and influence can be defined.

References

[1] Ackerberg, D. and M. Rysman. 2005. "Unobserved Differentiation in Discrete Choice Models: Estimating Price Elasticities and Welfare Effects." *RAND Journal of Economics* 36: 771-788.

[2] Allessie, R. and A. Kapteyn. 1991. "Habit Formation, Interdependent Preferences and Demographic Effects in the Almost Ideal Demand System." *Economic Journal* 101: 404:419.

[3] Anderson, S., de Palma, A. and J. -F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation* Cambridge, MA: MIT Press.

[4] Anselin, L. 1988. *Spatial Econometrics: Methods and Models* Dordrecht, NL: Kluwer.

[5] Anselin, L. 2002. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models," *Agricultural Economics* 27: 247-267.

[6] Armstrong, M. 2006. "Competition in Two-Sided Markets." *RAND Journal of Economics* 37: 668-681.

[7] Ballester, C., A. Calvo-Armengol, and Y. Zenou. 2006. "Who's Who in Networks. Wanted: The Key Player." *Econometrica* 74: 1403–1417.

[8] Becker, G., S., M. H. DeGroot, and J. Marschak. 1964. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science* 9: 226-232.

[9] Berrera, D. 2007. "The Impact of Negotiated Exchange on Trust and Trustworthiness." *Social Networks* 29: 508-526.

[10] Berry, S. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25: 242-262.

[11] Berry, S., J. Levinsohn, and A. Pakes. 1995. "Automobile Prices in Market Equilibrium," *Econometrica* 63: 841-890.

[12] Berry, S. and A. Pakes. 2007. "The Pure Characteristics Demand Model." *International Economic Review* 48: 1193-1225.

[13] Besanko, D., S. Gupta and D. Jain. 1998. "Logit Demand Estimation Under Competitive Pricing Behavior: An Equilibrium Framework." *Management Science* 44: 1533-1547.

[14] Bhattacharya, H. and R. Innes. 2007. "Does Market Concentration Promote or Reduce New Product Introductions? Evidence from US Food Industry." Paper presented at American Agricultural Economics Association meetings, Portland, OR.

[15] Borgatti, S. P., and M. G. Everett. 1999. “Models of Core / Periphery Structures.” *Social Networks* 21: 375-395.

[16] Brock, W. A. and S. N. Durlauf. 2007. "Identification of Binary Choice Models with Social Interactions." *Journal of Econometrics* 140: 52-75.

[17] Brock, W. A. and S. N. Durlauf. 2001. "Discrete Choice with Social Interactions." *Review of Economic Studies* 68: 235-260.

[18] Brownstone, D. and K. Train. 2000. "Forecasting New Product Penetration with Flexible Substitution Patterns." *Journal of Econometrics* 89: 109-129.

[19] Buskens, V. 1998. "The Social Structure of Trust." *Social Networks* 20: 265-289.

[20] Calvo-Armengol, A., E. Patacchini and Y. Zenou. 2009. "Peer Effects and Social Networks in Education." *Review of Economic Studies* 76: 1239-1267.

[21] Cardell, N. S. 1997. “Variance Components Structures for the Extreme Value and Logistic Distributions.” *Econometric Theory* 13: 185-213.

[22] Caplin, A. and B. Nalebuff. 1991. “Aggregation and Imperfect Competition: On the Existence of Equilibrium.” *Econometrica* 59: 26-61.

[23] Case, A. C. 1991. "Spatial Patterns in Household Demand." *Econometrica* 59: 953-965.

[24] Choi, J., S. K. Hui, and D. R. Bell. 2010. "Spatiotemporal Analysis of Imitation Behavior Across New Buyers at an Online Grocery Retailer." *Journal of Marketing Research* 47: 75-89.

[25] Cowan, R., W. Cowan and P. Swann. 1997. “A Model of Demand with Interactions Among Consumers.” *International Journal of Industrial Organization* 15: 711-732.

[26] dePalma, A., V. Ginsburgh, Y. Y. Papageorgiou and J. F. Thisse. 1985. "The Principle of Minimum Differentiation Holds Under Sufficient Heterogeneity." *Econometrica* 47: 1045-1050.

[27] Domingos, P. 2005. "Mining Social Networks for Viral Marketing." *IEEE Intelligent Systems* 20: 80-

[28] Driskill, R. and S. McCafferty. 2001. “Monopoly and Oligopoly Provision of Addictive Goods,” *International Economic Review* 42: 43-72.

[29] Dugundji, E. R. and J. L. Walker. 2005. “Network Interdependencies: An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects.” *Transportation Research Record: Journal of the Transportation Research Board* 1921: 70-78.

[30] Farrell, J. and G. Saloner. 1986. “Standardization and Variety,” *Economics Letters* 20: 71-74.

[31] Feenstra, R. C. and J. A. Levinsohn. 1995. “Estimating Markups and Market Conduct with Multidimensional Product Attributes.” *Review of Economic Studies* 62: 19-52.

[32] Freeman, L. C. 1979. “Centrality in Social Networks Conceptual Clarification.” *Social Networks* 1: 215-239.

[33] Godes, D. and D. Mayzlin. 2009. "Firm-Created Word-of-Mouth Communication: Evidence from a Field Test." *Marketing Science* 28: 721-739.

[34] Goldenberg, J., S. Han, D. R. Lehmann and J. W. Hong. 2009. “The Role of Hubs in the Adoption Process.” *Journal of Marketing* 73: 1-13.

[35] Granovetter, M. 1983. “The Strength of Weak Ties: A Network Theory Revisited,” *Sociological Theory* 1: 201-233.

[36] Grinblatt, M., M. Keloharju, and S. Ikaheimo. 2008. "Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors." *Review of Economics and Statistics* 90: 735-753.

[37] Grippa, F. and P. A. Gloor. 2009. "You are Who Remembers You. Detecting Leadership Through Accuracy of Recall." *Social Networks* 31: 255-261.

[38] Hanneman, R. A. and M. Riddle. 2005. “Introduction to Social Network Methods,” Working paper, University of California at Riverside, Riverside, CA.

[39] Horsky, D. and P. Nelson. 1992. “New Brand Positioning and Pricing in an Oligopolistic Market.” *Marketing Science* 11: 133-153.

[40] House, L., M. House, and J. Mullady. 2007. “Do Recommendations Matter? Social Networks, Trust and Product Adoption,” paper presented at European Association of Agricultural Economics Forum on Innovation and System Dynamics in Food Networks, Innsbruck-Igls, Austria. February 15 - 17.

[41] Iyengar, R., C. Van den Bulte and T. W. Valente. 2011. "Opinion Leadership and Social Contagion in New Product Diffusion." Marketing Science 30: 195-212.

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

[43] Kaiser, J. and U. Wright. 2006. “Price Structure in Two-Sided Markets: Evidence from the Magazine Industry,” *International Journal of Industrial Organization* 24: 1-28.

[44] Kuhfeld, W. F., R. D. Tobias and M. Garratt. 1994. "Efficient Experimental Design with Marketing Research Applications." *Journal of Marketing Research* 31: 545-557

[45] Liebenstein, H. 1950. "Bandwagon, Snob, and Veblen Effects in the Theory of Consumer's Demand." *Quarterly Journal of Economics* 64: 183-207.

[46] LeSage, J. P. 1998. Spatial Econometrics Monograph, Department of Economics, University of Toledo, Toledo, OH. December.

[47] Louviere, J. J. and G. Woodworth. 1983. "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data." Journal of Marketing Research 20: 350-367.

[48] Louviere, J. J. 1988. "Analyzing Decision Making: Metric Conjoint Analysis." Sage University Papers Series in Quintitative Applications in the Social Sciences No. 67, Newbury Park.

[49] Louviere, J. J., D. A. Hensher and J. Swait. 2000. *Stated Choice Methods: Analysis and Applications* Cambridge, U.K.: Cambridge University Press.

[50] McFadden, D. and K. Train. 1996. "Consumers' Evaluation of New Products: Learning from Self and Others." *Journal of Political Economy* 104: 683-703.

[51] Mintel. Global New Products Database. (http://www.gnpd.com/sinatra/gnpd/frontpage). September 19, 2011.

[52] Mullen, B., C. Johnson, and E. Salas. 1991. "Effects of Communication Network Structure: Components of Positional Centrality." *Social Networks* 13: 169-185.

[53] Narayan, V., V. Rao and C. Saunders. 2011. "How Peer Influence Affects Attribute Preferences: A Bayesian Updating Mechanism." 30: 368-384.

[54] Nevo, A. 2003. "New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems" *Review of Economics and Statistics* 85: 266–275.

[55] Petrin, A. 2002. "Quantifying the Benefits of New Products: the Case of the Minivan." *Journal of Political Economy* 110: 705-729.

[56] Pinkse, J., M. E. Slade and C. Brett. 2002. "Spatial Price Competition: A Semiparametric Approach." *Econometrica* 70: 1111-1153.

[57] Pinkse, J., and M. E. Slade. 2004. "Mergers, Brand Competition and the Price of a Pint." *European Economic Review* 48: 617-643.

[58] Reingen, P. H. and J. B. Kernan. 1986. "Analysis of Referral Networks in Marketing: Methods and Illustration." *Journal of Marketing Research* 23: 370-378.

[59] Reingen, P. H. and J. J. Brown. 1987. “Social Ties and Word-of-Mouth Referral Behavior.” *Journal of Consumer Research* 15: 350-362.

[60] Slade, M. E. 2004. “Market Power and Joint Dominance in U.K. Brewing.” *Journal of Industrial Economics* 52: 133-163.

[61] Suarez, F. 2005. “Network Effects Revisited: The Role of Strong Ties in Technology Selection,” *Academy of Management Journal* 48: 710-720.

[62] Sudhir, K. and V. R. Rao. 2006. "Do Slotting Allowances Enhance Efficiency or Hinder Competition." Journal of Marketing Research 43: 137-155.

[63] Thaler, R. 1985. “Mental Accounting and Consumer Choice,” *Marketing Science* 4: 199-214.

[64] Train, K. 2003. *Discrete Choice Methods with Simulation.* Cambridge, U.K.: Cambridge University Press.

[65] Trajtenberg, M. 1989. "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners." *Journal of Political Economy* 97: 444-479.

[66] Yang, S. and G. M. Allenby. 2003. “Modeling Interdependent Consumer Preferences.” *Journal of Marketing Research* 40: 282-294.

Table 1: Social Network Adjacency Matrix

	Network Member			
Network Member	A	B	C	D
A	-	2	3	4
B	2	-	1	2
C	3	1	-	0
D	4	1	0	-

Table 2: Summary of Demographic and Socioeconomic Data

Variable	Measure	Mean	Std. Dev.	N
Age	Yrs	20.1	1.4	73
Gender	% Male	50.1	47.8	73
Household Size	#	1.9	1.5	73
Income	$,000	45.8	53.1	73
Race	% White	60.3	48.9	73
	% African-American	1.4	11.6	73
	% Asian	15.1	35.8	73
	% Hispanic	6.9	25.3	73
	% Mixed	10.9	31.2	73
	% Other	5.5	22.8	73
Class	% Sophomore	6.8	25.3	73
	% Junior	53.4	25.9	73
	% Senior	39.7	38.9	73
Degree Expected	% Bachelor	32.9	36.9	73
	% Masters	60.3	48.9	73
	% Ph.D.	0.0	0.0	73
Marital Status	% Single	61.8	48.6	73
	% Married	38.2	48.6	73
	% Divorced	0.0	0.0	73
GPA	% 3.75-4.00	10.9	31.2	73
	% 3.25-3.74	46.6	49.9	73
	% 2.75-3.24	37.1	48.3	73
	% 2.25-2.74	5.5	22.8	73

Table 3: Summary of Social Relationship Data

Q1. How often do you communicate with __ on matters of mutual interest on a weekly basis?	
1 = none	78.2%
2 = once	7.9%
3 = 2-5 times	10.9%
4 = more than 5	0.0%
Q2. How well do you know __?	
1 = have never met	78.2%
2 = met once	10.9%
3 = somewhat acquainted	5.0%
4 = well acquainted	2.9%
5 = know them well	0.0%
Q3. How well do you respect the decision making ability of ___?	
1 = no respect	20.6%
2 = a little respect	2.9%
3 = indifferent	64.7%
4 = some	2.9%
5 = very much	5.9%
Q4. How well do you trust the decision making ability of ___?	
1 = no trust	26.5%
2 = some trust	5.9%
3 = indifferent	61.8%
4 = some	2.9%
5 = very much	0.0%
Q5. How would you rate the leadership abilities of ___?	
1 = no leadership ability	8.8%
2 = some leadership ability	2.9%
3 = indifferent or don't know	82.4%
4 = good leadership ability	0.0%
5 = excellent leadership ability	2.9%

Table 4: Summary of Choice Data

Attribute	Value	% Choice Stage 1	% Choice Stage 2
Price	$2.00	39.7%	43.1%
	$2.50	29.4%	30.9%
	$3.00	19.9%	22.5%
	$3.50	14.0%	13.5%
Flavor	Americone Dream	27.9%	31.6%
	Vanilla	25.4%	21.9%
	Dulche Delish	23.9%	29.0%
Fat	Regular Fat	27.3%	29.9%
	Low Fat	24.1%	25.1%
Organic	Conventional	26.8%	27.6%
	Organic	24.6%	27.5%

Note: Values are defined as the % of observations in which the chosen product has the associated attribute value.

Table 5: Choice Model Estimates: Social Network Effects

			Model 1: Proximity			Model 2: Betweenness			Model 3: Centrality			Model 4: Farness			Model 5: Core		
	β_i^I	t-ratio	ρ_{ik}	t-ratio	β_i^R	ρ_{ik}	t-ratio	β_i^R	ρ_{ik}	t-ratio	β_i^R	ρ_{ik}	t-ratio	β_i^R	ρ_{ik}	t-ratio	β_i^R
Price	-2.298	-15.117	0.831	6.423	-2.357	0.761	14.68	-3.359	0.498	31.616	-2.768	0.726	5.42	-2.539	0.925	15.813	-2.745
Flavor	0.624	6.213	0.369	4.314	0.888	0.486	5.872	1.174	0.445	18.359	1.014	0.301	1.032	1.206	0.417	4.514	0.753
Fat	0.061	0.391	0.131	0.742	-0.150	0.658	5.023	0.069	0.313	8.425	-0.353	0.924	3.781	-0.003	0.481	3.425	-0.218
Organic	0.227	1.616	0.307	5.571	0.177	0.792	3.234	0.208	0.249	3.845	0.245	0.889	1.252	0.235	0.584	5.933	0.116
Non-Random Parameters:																	
Age	-0.046	-6.577															
Income	-0.426	-2.346															
HH size	1.529	5.986															
Gender	0.036	0.109															
Standard Deviation of Random Parameters:																	
Price	1.087	13.443	2.111	9.963		0.783	12.402		0.069	33.627		1.435	13.166		1.218	13.566	
Flavor	1.559	10.952	1.029	7.847		0.673	9.091		0.012	5.862		1.022	9.692		2.277	7.964	
Fat	1.401	8.846	1.106	4.672		0.221	2.988		0.001	0.063		0.406	2.903		1.084	6.8	
Organic	1.008	5.615	1.075	5.039		0.986	4.444		0.009	2.134		1.178	5.062		1.997	5.945	
LLF	-1033.1		-962.3			-1011.9			-1203.8			-1026.1			-935.4		
χ^2	1172.6		1314.3			1214.4			830.8			1185.9			1367.5		
R^2	0.362		0.406			0.375			0.257			0.366			0.422		

Table 6: Choice Model Estimates: Social Network and Attribute Space Interactions

			Model 1: Proximity			Model 2: Betweenness			Model 3: Centrality			Model 4: Farness			Model 5: Core		
	β_i^I	t-ratio	ρ_{ik}	t-ratio	β_i^R	ρ_{ik}	t-ratio	β_i^R	ρ_{ik}	t-ratio	β_i^R	ρ_{ik}	t-ratio	β_i^R	ρ_{ik}	t-ratio	β_i^R
Price	-2.299	-15.117	0.883	5.261	-2.585	1.041	15.233	-3.658	0.786	15.171	-2.088	0.431	6.599	-3.686	0.054	16.064	-1.057
Flavor	0.624	6.213	0.336	7.916	0.886	0.818	9.120	0.282	0.605	11.792	0.457	0.405	5.118	1.185	0.031	7.779	0.691
Fat	0.061	0.391	0.345	1.687	-0.280	0.053	-0.373	-0.057	0.284	2.792	0.349	0.491	3.72	-0.232	0.186	1.444	-0.194
Organic	0.227	1.616	0.172	4.357	0.166	0.192	6.744	0.167	0.569	6.077	0.059	0.225	3.527	0.201	0.581	3.077	0.093
Non-Random Parameters:																	
Age	-0.047	-6.577															
Income	-0.426	-2.346															
HH size	1.529	5.986															
Gender	0.036	0.109															
Standard Deviation of Random Parameters:																	
Price	1.087	13.443	0.818	7.714		0.712	11.119		0.471	12.101		0.542	11.299		0.918	13.736	
Flavor	1.559	10.952	1.763	9.589		1.568	9.527		0.132	11.326		0.997	10.18		1.143	10.811	
Fat	1.401	8.846	1.191	3.116		1.314	9.817		0.288	11.286		0.626	6.719		1.269	7.085	
Organic	1.008	5.615	1.604	4.685		0.977	6.422		1.963	6.588		0.756	8.456		2.832	8.811	
LLF	-1033.1		-943.9			-972.9			-1070.4			-1018.5			-945.4		
χ^2	1172.6		1350.5			1292.5			1097.1			1201.3			1347.6		
R^2	0.362		0.417			0.399			0.339			0.371			0.416		

Table 7: Estimates of Influence Model Parameters

Parameter	Variable	Estimate	t-ratio
λ_0	Individual Effect	0.103*	9.910
σ_0	Scale of Ind. Effect	0.055*	44.178
λ_1	Social Network Location	0.056*	52.566
σ_1	Scale of Social Net. Location	1.109*	563.042
σ^2	Variance	0.142*	157.065
LLF	Likelihood Function Value	334.665	
χ^2	Chi-Square Value	669.330	

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

Table 8: Ranking of Subject Influence

Subject ID	Proximity Index	Subject ID	Influence Index
5	1.305	50	0.229
70	1.256	24	0.210
1	1.184	60	0.199
34	1.184	8	0.197
38	1.172	11	0.190
29	1.159	29	0.179
57	1.159	6	0.178
33	1.148	13	0.176
48	1.136	48	0.175
22	1.124	65	0.173
23	1.124	68	0.171
35	1.122	22	0.170
72	1.122	15	0.162
51	1.111	51	0.158
56	1.111	69	0.151

Note: The proximity index is calculated as the total Euclidean distance from each member to all others, expressed as a ratio to the mean. The influence index is the subject-specific estimate of λ_{0i}.

Table 9: Influence and Social Network Definition

ID	Proximity	ID	Betweenness	ID	Centrality	ID	Farness	ID	Core
43	2.097	43	0.656	69	0.652	37	0.646	28	0.683
27	1.907	37	0.656	25	0.648	29	0.642	29	0.676
29	1.890	28	0.640	37	0.647	28	0.641	27	0.675
28	1.854	25	0.637	28	0.636	69	0.640	37	0.672
37	1.758	68	0.634	29	0.634	25	0.630	25	0.671
67	1.736	29	0.630	67	0.632	43	0.630	43	0.666
25	1.726	69	0.627	68	0.623	67	0.627	69	0.655
69	1.723	67	0.621	27	0.620	68	0.624	67	0.644
68	1.586	27	0.619	73	0.607	27	0.620	68	0.641
73	1.473	73	0.606	51	0.551	73	0.598	73	0.617
48	1.020	51	0.555	4	0.534	4	0.549	51	0.556
15	0.984	4	0.540	23	0.532	51	0.547	71	0.537
4	0.899	23	0.537	32	0.531	32	0.541	42	0.536
22	0.849	71	0.534	71	0.531	63	0.539	24	0.531
34	0.846	32	0.533	56	0.528	17	0.537	4	0.531

Note: Values in table represent fitted influence values from the influence model developed above.