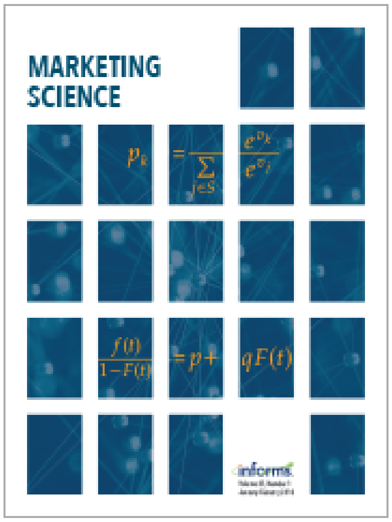


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Modeling Choice Interdependence in a Social Network

Jing Wang, Anocha Aribarg, Yves F. Atchadé

MARKETING SCIENCE
Vol. 32, No. 6, November–December 2013, pp. 977–997
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2013.0811

Modeling Choice Interdependence in a Social Network

Jing Wang
McKinsey & Company, Shanghai 200020, China, jingwang16@gmail.com

Anocha Aribarg
Stephen M. Ross School of Business, University of Michigan, Ann Arbor, Michigan 48109, anocha@umich.edu

Yves F. Atchadé
Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109, yvesa@umich.edu

This paper investigates how individuals' product choices are influenced by the product choices of their connected others and how the influence mechanism may differ for fashion- versus technology-related products. We conduct a novel field experiment to induce and observe choice interdependence in a closed social network. In our experiment, we conceptualize individuals' choices to be driven by multiattribute utilities, and we measure their initial attribute preferences prior to observing their choice interdependence and collecting network information. These design elements help alleviate concerns in identifying social interaction effects from other confounds. Given that we have complete information on choices and their sequence, we use a discrete-time Markov chain model. Nonetheless, we also use a Markov random field (MRF) model as an alternative when the information on choice sequence is missing. We find significant social interaction effects. Our findings show that whereas experts exert asymmetrically greater influence on a technology-related product, popular individuals exert greater influence on a fashion-related product. In addition, we find choices made by early decision makers to be more influential than choices made later for the technology-related product. Finally, using the MRF with snapshot data can also provide good out-of-sample predictions for a technology-related product.

Key words: social interactions; social network; social influence; homophily; conjoint experiment; Markov random field

History: Received: October 31, 2010; accepted: July 27, 2013; Preyas Desai served as the editor-in-chief and Gary Russell served as associate editor for this article. Published online in *Articles in Advance* October 8, 2013.

Introduction

Researchers have long recognized that social interactions among individuals in a social network have an impact on their preference evolution and product adoptions. The rapid growth of social networking websites has facilitated social interactions to an unprecedented level. In addition to encouraging users to maintain current connections and build new ones on its site, Facebook automatically passes along information about users' site activities to their friends. For instance, when users become a fan of a brand page or respond to an interactive ad on the site, their friends receive a notification. These notifications are sent with the expectation that they may influence other users. Social interaction effects are also undoubtedly prevalent in off-line contexts. Individuals' decisions to purchase a hybrid vehicle, for example, may be influenced by recommendations from their neighbors. Teenagers' decisions to purchase certain brands of jeans can be influenced by observing their peers' choices. In this paper, we examine social interactions in a multiattribute choice context where we conceptualize *choice interdependence* in a social network as being driven by individuals' preference shifts induced by social interactions after controlling for their initial preferences.

Although extant research has attempted to measure social interaction effects (e.g., Nair et al. 2010, Nam et al. 2010, Tucker 2008), it does not recognize that such effects may arise from different types of social influence. Prior behavioral research (Bearden and Etzel 1982, Burnkrant and Cousineau 1975, Deutsch and Gerard 1955) distinguishes a value-expressive influence from an informational social influence. Whereas an informational influence is driven by an individual's goal to gain more knowledge, a value-expressive influence is driven by an individual's goal to maintain and enhance his or her self-perception or identity in relation to his or her peers. Given different goal orientations, consumers rely more on the credibility of the source of influence when it comes to an informational influence and more on desirable characteristics of the source when it comes to a value-expressive influence (McGuire 1969). Given that we expect an informational influence to play a more important role for a technology-related product and a value-expressive influence for a fashion-related product, we examine how experts (i.e., with more credibility) versus popular individuals (i.e., with a more desirable social characteristic) are asymmetrically more influential for technology-related versus fashion-related products. In addition, previous research suggests that two distinct groups of influentials and imitators may play different roles in the diffusion of innovations (Van den Bulte and Joshi 2007). That is, influentials tend to be opinion leaders (Katz and Lazarsfeld 1955) and early adopters who influence, but are not or to a lesser degree influenced by, imitators. Provided with information about the sequence of choices, we also assess whether individuals are more influenced by the choices made by early decision makers than they are by choices made later.

Prior research has attempted to identify the causal effects of social interactions (Hartmann et al. 2008; Manski 1993, 2000; Moffitt 2001). With observational data, social interactions are likely confounded with homophily or endogenous group formation, correlated unobservables, and simultaneity. We propose a novel conjoint experiment for studying social interactions that helps alleviate these endogeneity concerns. Specifically, we examine choice interdependence in a social network and conceptualize consumer choices to be driven by multiattribute utilities. We conduct a two-stage choice experiment with a preinfluence stage and a postinfluence stage. The experiment also involves two product categories: one that is fashion-related and another that is technology-related. In the preinfluence stage, we use a conjoint choice experiment to measure each participant's initial preferences for different product attributes that govern the utilities of the product choices. We also collect information about each participant's social connections within a closed social network. In the postinfluence stage, we conduct an incentive-aligned choice experiment to induce and observe choice interdependence within the network. This experimental design helps us identify preference shifts as a result of social interactions separately from initial preferences.

To study choice interdependence and identify the social interaction effects, researchers should be able to observe both individual choices and the sequence of choices (complete data). We fit a discrete-time Markov chain (DMC) model to the complete data. However, researchers often do not keep records of the sequence because of a capacity constraint, or in off-line contexts, it may not even be feasible to observe the choice sequence. In these situations, researchers have to work with cross-sectional data, where realized choices are recorded only at a particular time point (snapshot data). We demonstrate how researchers can use a Markov random field (MRF) model as an alternative to a DMC model when only snapshot data are available. Given their theoretical connection, both DMC and MRF models are appropriate for modeling choice interdependence in a social network because both models properly account for the dependence structure in the data (i.e., the joint probability of individual choices).

Controlling for individuals' heterogeneous initial preferences, we find significant social interaction effects based on both the DMC and MRF models. Our findings suggest that, on the one hand, popular individuals (i.e., those with more connections) exert greater influence for a fashion-related product where value-expressive influence is more prominent. On the other hand, experts exert greater influence for a technology-related product where informational influence dominates. We also find that individuals who make choices earlier are more influential in the technology-related product category. This result alludes to the opinion leadership of early decision makers for this product category. Finally, our model comparison shows good out-of-sample predictive performance of the MRF model for the technology-related product.

The contribution of our paper is threefold. First, we assess how different types of influences, informational versus value-expressive, lead experts versus popular individuals to asymmetrically exert greater influence, respectively, for technology-related versus fashion-related products. We also examine the potential disproportionate influence of early decision makers. Second, coupled with network information, we demonstrate how researchers can use a two-stage conjoint choice experiment, which we conceptualize based on the multiattribute utility framework (Keeney and Raiffa 1993), to identify individual preference shifts induced by social interactions after controlling for their initial preferences. Our approach helps mitigate the endogeneity concern. Finally, given that we use the DMC to model choice interdependence when the information on both choices and their sequence is available, we examine how the theoretically connected MRF model can be used as an alternative when the sequence information is missing—a common occurrence in most online and off-line situations.

The rest of this paper is structured as follows. We begin with the literature review in relation to our conceptual framework and then present our modeling framework. This is followed by a description of our experimental design and data collection procedure. Our empirical analysis section includes (i) descriptive statistics of the participants and their social network, (ii) the analysis of individual-specific initial preferences, (iii) the detailed model specification and estimation procedure, (iv) model comparison, and (v) estimation results. The discussion section concludes our paper.

Literature Review and Conceptual Framework

There is a sizable body of literature on social interaction effects (also referred to as peer, neighborhood, and social contagion effects) in the fields of economics, sociology, and marketing. In economics, the primary focus is on identifying an "active" social interaction that gives rise to a social multiplier using observational data (Brock and Durlauf 2001; Manski 1993, 2000). Hartmann et al. (2008) distinguish active from passive social interactions; the former involves two individuals' actions being simultaneously affected by the actions of the other and resolved as an equilibrium outcome (e.g., Hartmann 2010), whereas the latter treats the impact of an individual's action on the other's action as being sequential in nature. Although the concept of active social interactions originates in sociology (Granovetter 1978, Schelling 1971), the majority of work in this discipline does not hinge on an equilibrium assumption (i.e., focus on passive social interactions), and the emphasis is placed on how the structure of individuals' social network affects how social interactions propagate information diffusion in the network (Granovetter 1973, Katz and Lazarsfeld 1955, Watts and Dodds 2007). Compared with sociology, not much attention has been paid to the structure of a social network in the economic literature. Early research in economics and sociology mainly proposes theoretical frameworks to characterize the mechanism of social interactions without corroborating empirical evidence.

Despite no explicit attempt to model social interactions, the marketing literature has long recognized their role (i.e., influentials versus imitators) in the diffusion of innovations (Bass 1969, Van den Bulte and Stremersch 2004, Van den Bulte and Wuyts 2007). With the availability of observational data especially in online contexts, research in marketing explicitly links social interactions to firm performance and product adoptions (Godes et al. 2005, Hartmann et al. 2008). On the one hand, some researchers have examined the effects aggregate word of mouth (Chen et al. 2011, Chevalier and Mayzlin 2006, Godes and Mayzlin 2004, Zhu and Zhang 2010), online viral marketing activities (De Bruyn and Lilien 2008, Toubia et al. 2011), and online community participation (Algesheimer et al. 2010, Manchanda et al. 2013, Stephen and Toubia 2010) have on aggregate sales. On the other hand, others attempt to establish the causal effects of social interactions on adoption decisions at the individual level (Nair et al. 2010, Nam et al. 2010, Tucker 2008, Zhang 2010). A few researchers examine social interactions in a multiattribute utility framework[1] (Narayan et al. 2011, Yang and Allenby 2003) or explicitly collect network information (Nair et al. 2010, Narayan et al. 2011, Reingen et al. 1984). In studying choice interdependence, we conceptualize individual choices to be driven by multiattribute utilities and focus on the passive form of social interactions. We incorporate explicit information on a network of individuals that is not endogenously determined by choice behavior.

None of the aforementioned research considers that social interaction effects may arise from different types of social influence. Behavioral research (Bearden and Etzel 1982, Bearden et al. 1989, Burnkrant and Cousineau 1975, Deutsch and Gerard 1955) distinguishes a value-expressive influence from an informational social influence. According to Kelman (1958, 1961), whereas an informational influence operates through the internalization process, a value-expressive influence operates through the identification process. An informational influence is driven by an individual's goal to gain more knowledge to make informed decisions. Faced with uncertainty, individuals seek information from sources that are viewed as credible (DeBono and Harnish 1988, Kelman 1958, McGuire 1969). In contrast, a value-expressive influence[2] is driven by someone's goal to maintain and enhance their self-perception or identity by associating with a reference group (Belk 1988). These individuals would associate (disassociate) themselves with a reference group they perceive as being positive (negative) (Miller et al. 1993). As a result, they rely on desirable characteristics of the information source, such as attractiveness (DeBono and Harnish 1988, Kelman 1958), as criteria to judge the value of the information.

Park and Lessig (1977) suggest that the extent to which different types of social influence affect an individual's brand selection varies by product category. Informational influence plays a more important role for a technology-related product; therefore,

[1] Some research relies on the random utility framework (Brock and Durlauf 2001, Hartmann 2010) but not the multiattribute utility framework.

[2] Previous research distinguishes two types of normative influence: (i) a value-expressive influence, which is motivated by self-maintenance or enrichment, and (ii) a utilitarian influence, which is motivated by external rewards. We focus only on a subtype of normative influence, the value-expressive influence, because it simplifies the connection between the type of social influence and technology-related versus fashion-related products. Note, however, despite their conceptual differences, researchers have had difficulty in empirically distinguishing value-expressive and utilitarian influences (Bearden et al. 1989, Burnkrant and Cousineau 1975).

we expect that individuals pay more attention to the source credibility, and experts exert an asymmetrically larger influence in this product category. In contrast, value-expressive influence dominates in a fashion-related product; therefore, we expect that individuals focus more on the desirable characteristics of the source, and popular individuals exert an asymmetrically larger influence in this product category. The asymmetric effects of social interactions have been documented in the prescription drug (Nair et al. 2010) and video messaging (Tucker 2008) categories. However, the study contexts primarily involve informational influences. In addition, Nair et al. (2010) examine asymmetric social interaction effects for prescription decisions—decisions for others, not for one's self. Tucker (2008) studies social interactions in the presence of network externalities—a larger number of adopters enhance product performance and benefits. We study choice interdependence in the absence of network externalities.

In the context of diffusion of innovations, influentials may have a greater influence than imitators (Van den Bulte and Joshi 2007, Van den Bulte and Wuyts 2007). Influentials are early adopters of innovations, who are likely to be opinion leaders (Katz and Lazarsfeld 1955, Myers and Robertson 1972, Rogers and Cartano 1962). Czepiel (1974) reports early adopters to exhibit great opinion leadership in the diffusion of technological innovation among competing firms. Nair et al. (2010) find that research-active specialists who are supposed to be early decision makers exert asymmetrically higher influence. Although not in the context of innovation adoptions, Trusov et al. (2010) also document that users who have been a member of a social networking site for a longer period of time (measured by the months a user has been a member) are more influential than others in stimulating activities on the site. Given the information about the choice sequence, we also assess whether early decision makers (i.e., influentials) have a greater influence than later decision makers (i.e., imitators). In studying choice interdependence, we examine how experts versus popular individuals exert asymmetrically greater influence for technology-related versus fashion-related products, respectively, and how the choices made by early decision makers may be disproportionately more influential.

Modeling Framework

To examine choice interdependence in a social network, it is desirable to have complete longitudinal choice data where both choices and their sequence are observed. With the choice sequence information, we can examine how observed choices in a particular order may be more influential than others. It also allows us to model not only which choice is made but also who is likely to make a choice at a particular time. Snapshot data, on the other hand, are cross-sectional choice data collected at a fixed time point. This type of data is easier and less expensive to obtain. However, a lack of information on the choice sequence makes it difficult to understand the dynamic process because different processes can give rise to the same snapshot data. No matter what type of data is available, we need to model all consumers' choices jointly rather than independently to properly preserve the dependence structure in the data. In modeling choice interdependence, we treat the social network as being exogenous and fixed instead of modeling its formation or properties (Ansari et al. 2011, Iacobucci and Hopkins 1992, Reingen and Kernan 1986, Watts and Dodds 2007).

Discrete-Time Markov Chain for Complete Data

Besides sociology, social networks and their influences on individuals' behavior have also sparked interests of researchers in other fields, such as statistics and computer science. Some researchers (Hoff et al. 2002, Hunter 2007, Robins et al. 2007, Wasserman and Pattison 1996) primarily focus on modeling the structure (i.e., connections among nodes) and properties of a social network, whereas others (Hill et al. 2006), like us, attempt to model the behavior of the network nodes, e.g., choice decisions such as in our study. For instance, Koskinen and Snijders (2007) and Snijders et al. (2007) model the dynamics of network formation and behavioral evolution simultaneously by collecting the snapshot data of network structure and behavioral outcomes at more than one time point (e.g., every three months). However, the authors collect data at a few time points. As such, between two consecutive time points there are multiple processes that can lead to the same observed behavioral outcomes. They hence use a continuous-time Markov process to fit the data and treat the possible processes between consecutive time points as latent variables to be augmented. With complete data, we do not need to rely on augmented variables and adopt a simpler DMC model (Ross 1996).

Following Koskinen and Snijders (2007), we assume that at most one individual may make a choice decision at a given time (i.e., no multiple decisions are allowed simultaneously). This assumption is reasonable in the context of sequential choice such as ours because it is unlikely that two individuals will make choice decisions at the exact same time. The assumption simplifies our model specification because we do not need to handle the possible dependence of the transitions made at the same time. The state of the Markov chain hence has two parts: (i) who is likely to make a choice and (ii) the choices made by all n individuals in a network at a particular time. Again, we emphasize that the state of the Markov chain involves all individuals' choices rather than a single person's as a result of the interdependence. We denote "who is likely to make a choice" at time t by c^t, which takes values from 1 to n, and individual i's choice at time t by Y_i^t, which takes nominal values $0, \ldots, K$. The state at time point t is $S^t = (c^t, Y_1^t, \ldots, Y_n^t)$. The first-order Markovian property dictates that

$$p(s^{t+1} \mid s^1, \ldots, s^t, \theta) = p(s^{t+1} \mid s^t, \theta). \quad (1)$$

Under the assumption that only one transition is allowed, only one person's choice $Y_{c^{t+1}}^{t+1}$ can be different from $Y_{c^{t+1}}^t$, while all other choices remain the same, $Y_j^{t+1} = Y_j^t$ for $j \neq c^{t+1}$. Then we have

$$\begin{aligned} p(s^{t+1} \mid s^t, \theta) &= p(c^{t+1} = i, y_i^{t+1} \mid c^t, y^t, \theta) I(y_{-i}^{t+1} = y_{-i}^t) \\ &= p(c^{t+1} = i \mid c^t, y^t, \theta) p(y_i^{t+1} \mid c^{t+1} = i, c^t, y^t, \theta) \\ &\quad \cdot I(y_{-i}^{t+1} = y_{-i}^t), \end{aligned} \quad (2)$$

where $y_{-i} = (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$. The likelihood is $p(s^1, \ldots, s^T \mid s^0, \theta) = \prod_{t=0}^{T-1} p(s^{t+1} \mid s^t, \theta)$, where y^0 and c^0 are to be specified later. Researchers can postulate any probability form of interest for $P(c^{t+1} = i \mid c^t, y^t, \theta)$. Conditional on c^t, there is great flexibility in specifying the multinomial choice probability $p(y_i^{t+1} \mid c^{t+1} = i, c^t, y^t, \theta)$ (e.g., a multinomial logit). We include $I(y_{-i}^{t+1} = y_{-i}^t)$ for completion to ensure that only one choice can be made at a time. Note that a transition is determined by an individual attempting to make his or her choice. Given that DMC models everyone's choices together, it preserves the interdependence within the system.

Markov Random Field for Snapshot Data

Recognizing that many times researchers may not have information on the choice sequence, we examine the use of an MRF model with snapshot data. Despite the differences in the formulation and type of data they handle, DMC and MRF are closely connected. We provide a proof in Online Appendix A (available at http://dx.doi.org/10.1287/mksc.2013.0811) that under certain conditions, the DMC converges to the MRF, or, in other words, the MRF is an equilibrium model for the DMC.

The MRF is the general generating function for a multivariate logistic (MVL) choice model. (A detailed description of the MRF structure is provided in Online Appendix B.) MVL models have been used in marketing for studying consumer market baskets (Russell and Peterson 2000), multiple-category price competition (Song and Chintagunta 2006), multiple-category incidence and quantity (Niraj et al. 2008), product recommendations (Moon and Russell 2008), and coordinated choice decisions (Yang et al. 2010). An MVL that involves binary choices is referred to as an autologistic model and one that involves multinomial choices as an auto-model.

An auto-model (Besag 1974, 1975) is essentially a special case of the MRF. We denote individual i's choice by y_i; $i = 1, \ldots, n$. In Yang et al. (2010), each choice is dummy coded as a vector of 0 and 1; e.g., $y_i = (0, 1, 0)$ indicates that among the three choice options, individual i chooses option 2. The joint distribution of the auto-model takes the form of $p(y_1, \ldots, y_n) = \exp(\sum_{1 \leq i \leq n} y_i U(y_i) + \sum_{1 \leq i < j \leq n} \beta_{ij} y_i' y_j)/z(\beta)$, where $U(y_i)$ is a function capturing individual i's own preferences, $\beta_{ij} y_i' y_j$ captures choice interdependence, and $z(\beta)$ is the normalizing constant that sums over all possible values of $y = (y_1, \ldots, y_n)$. The parameter $\beta_{ij} = 0$ when choice decisions of individuals i and j are independent. Given the restrictive specification of $\beta_{ij} y_i' y_j$, the auto-model can only accommodate symmetric social influence effects $\beta_{ij} = \beta_{ji}$. The MRF allows for any functional form to capture choice interdependence and thus can accommodate asymmetric effects of social interactions. (A technical discussion appears in Online Appendix B.)

To incorporate influence asymmetry, some necessary conditions need to be met for both DMC and MRF. The first condition involves a clearly defined identity of each node at either the individual or role level (e.g., some people are experts on high-tech products whereas others are novices). The second condition is that statistics corresponding to the parameters for asymmetry must have different values, which most of the time are satisfied in a directed network. To identify asymmetric influence at the individual level, researchers also need to observe more than one choice for each individual. If node identities are defined at the role level, one decision per individual will be sufficient as long as aggregate statistics associated with each role differ.

As the group size grows, the MRF, as well as the auto model, suffers from the intractable normalizing constant problem. Specifically, because $z(\beta)$ involves summing over all possible values of y, its computation can become intractable. Given that our paper involves a large social network, we demonstrate the use of an efficient approximate sampling algorithm (Murray et al. 2006, Wang and Atchadé 2014; see also Online Appendix C) to circumvent the intractable normalizing constant problem in estimating an MRF model.

Experimental Design and Data Collection

Experimental Design

Overview. We conduct a two-stage conjoint choice experiment: a preinfluence stage and a postinfluence stage. The experiment also involves two products: a bundle of university sports paraphernalia and a Bluetooth® headset. Our multistage conjoint design is in line with previous research examining reference group influence[3] (Aribarg et al. 2002, Arora and Allenby 1999, Narayan et al. 2011) using conjoint experiments. In the preinfluence stage, for each product we use a standard conjoint choice experiment (Haaijer and Wedel 2003) to measure each participant's initial preferences (i.e., part-worths) for different product attributes that govern the utilities of the product choices he or she makes. We then collect information about the participants' connections within a relatively well-connected and closed social network. Our network involves undergraduate students who took an introduction to marketing class in the fall of 2009 at a Midwest university. The majority of participants knew each other. We expect this network to be denser than that of a randomly selected group of students.

In the postinfluence stage, we conduct a field conjoint experiment where we ask participants to enter two raffles as a token of appreciation—in addition to class credits—for their participation in the preinfluence stage. Each participant may opt out of the raffles. Our design is consistent with the incentive-aligned design discussed in Ding (2007). We notify the participants that they will receive an email with a Web page link from which they can make a prize choice (from four possible options) should they win one of the raffles. Each of these prize options is described by product attributes for which we measure individual-specific part-worths of the participants in the preinfluence stage. Participants make a series of choices in the preinfluence stage but only make one prize choice for each product in the postinfluence stage. The choice set in the postinfluence stage also does not overlap with any choice set in the preinfluence stage. To induce choice interdependence, we show the participants on the Web page the choices made by their friends (including "no choice"). We let the choice interdependence process evolve organically and collect longitudinal data of choices sequentially made. All choices are recorded in real time (complete data). As a result, we also obtain cross-sectional data that are the realization of the complete data at the end of the study (snapshot data).

Fashion-Related vs. Technology-Related Products. We regard university sports paraphernalia as a fashion-related product and a Bluetooth headset as a technology-related product. Choice decisions for both products are likely affected by reference group influences because they can be easily "seen or identified by others" (Bearden and Etzel 1982, Bourne 1957). The products are also of interest to the student population. Park and Lessig (1977) develop a series of Likert scale questions to measure the extent to which informational and value-expressive influences are relevant to a consumer's brand selection for a product.[4] We use these questions to validate our product selection.

We recruit 78 participants from the undergraduate student population at the same university. Each participant rates both products on this scale. We randomize the order of products being rated across participants.[5] The findings show that for sports paraphernalia, the mean value-expressive influence (4.139, SD = 0.549) is significantly higher than the mean informational influence (2.994, SD = 0.980; $t = 9.177$), whereas for the Bluetooth headset, the mean informational influence (4.034, SD = 0.655) is significantly higher than the mean value-expressive influence (3.709, SD = 0.791; $t = 2.884$). The mean informational influence is also significantly higher for the Bluetooth headset than for sports paraphernalia (4.034 versus 2.994; $t = 7.745$). In contrast, the mean value-expressive influence is significantly higher for sports paraphernalia than for the Bluetooth headset (4.139 versus 3.709; $t = 4.509$). These results validate the use of university sports paraphernalia as a product that involves a value-expressive influence and a Bluetooth headset as a product that involves an informational influence.

Identification of Social Interaction Effects. Observed correlation in the behavior of individuals who belong to the same reference group may not indicate that someone's action has a causal effect on the actions of the others in the group. Previous research has extensively discussed the identification (i.e., the causal inference) of social interaction effects from potential confounds, including homophily or

[3] We study individual choice decisions, not joint (Aribarg et al. 2002, 2010; Arora and Allenby 1999) or coordinated choice (Hartmann 2010, Yang et al. 2010) decisions.

[4] We include all 14 items that are supposed to measure three types of influence, informational, utilitarian, and value-expressive, although we only report the results for the two types of influence that are relevant to our research. The complete analysis is available from the authors upon request.

[5] A factor analysis shows that for the value-expressive influence, all five items form a single factor. However, for the informational influence in the Bluetooth category, two out of five items appear to capture a different underlying factor from the other three that capture the majority of the explained variance. Upon further investigation, we decide to use only these three statements because they are also more relevant to the product categories in our study. We drop two statements: (i) the brand that the individual selects is influenced by observing a seal of approval of an independent testing agency and (ii) the individual's observation of what experts do influences his choice of a brand (such as observing the type of car the police drive or the brand of television TV repairmen buy).

endogenous group formation, correlated unobservables, and simultaneity (Manski 1993, 2000; Moffitt 2001). Homophily describes the reverse causality effect of social interactions where the correlated behavior of individuals within a reference group results from the individuals' tendencies to form social connections with others who share similar tastes and preferences, and not a causal effect of someone's behavior on another.

In economics, the collection of exogenous social network information is rare. Social connections are sometimes inferred from the observed behavior with a sensitivity to social interactions that researchers aim to study (Tucker 2008) or the geographic and/or demographic proximities of individuals (Bell and Song 2007, Nam et al. 2010, Yang and Allenby 2003). Such procedures have the potential to interject a homophily problem into a study. Different methods are proposed as remedies. With the availability of panel data, one can specify a model with individual fixed (Nair et al. 2010) or random effects (Hartmann 2010) to control for individual preferences for a behavior or an action. Alternatively, researchers can simultaneously model the group formation process, which requires exclusion restrictions (Steglich et al. 2010, Tucker 2008), or use other exogenous variables to help separate social interaction effects from homophily (Nam et al. 2010). Aral et al. (2009) propose a dynamic matched sample estimation framework to tease out correlated product adoption that can be explained by observable individuals' characteristics and behavior.

Bramoullé et al. (2009), Hartmann et al. (2008), and Manski (1993, 2000) advocate more research to collect network information that is exogenous to the focal behavior to circumvent homophily. The social network in our study is determined by the frequency with which participants interact, not their choice behavior. However, one may expect individuals who interact frequently to have similar preferences for certain product attributes, such as a particular brand (Reingen et al. 1984). We control for potential homophily by accounting for each participant's utility (i.e., overall preference) for each prize option presented to him or her in the postinfluence stage. Each option's utility is derived from each participant's partworths—capturing his or her initial preferences for different product attributes—estimated from the preinfluence stage. In addition to network information, Manski (2000) encourages more research to explicitly elicit preferences using an experimental approach. Our design employs both elements.

Shalizi and Thomas (2011) distinguish latent homophily from homophily. The authors argue that even when researchers tease out homophily by conditioning on previous behavior or action (e.g., jumping off a bridge) and observable individual traits including preference (e.g., fondness of jumping of a bridge), the presence of unobservable individual traits that drive the preference for the action (e.g., a thrill-seeking propensity trait) and engender the social tie (e.g., a bond formed by a common thrill-seeking propensity) still confounds the social interaction effects with latent homophily.[6] Most research studying social interactions cannot easily measure the latent preference driving the action under investigation (e.g., technology adoption, school performance) and controls for preference only at the action level (e.g., adding fixed effects to a model intercept). In contrast, conjoint experiments allow us to precisely measure individuals' intrinsic preferences at the characteristic of the action level (i.e., preferences for attributes of a choice). In a way, we measure the latent attribute preferences (e.g., the element of thrill-seeking in jumping off a bridge) that drive the utility (i.e., overall preference) of each choice option. Our multiple choice tasks in the preinfluence stage are created based on an optimal design, which aims to maximize the efficiency of part-worth estimates. As a result, each choice option, including those in the postinfluence stage, is hypothetically created and entirely determined by the attributes that characterize the option. As such, the possibility that a prize choice is influenced by latent factors other than a participant's latent initial attribute preferences and induced social interactions is minimal.

Correlated unobservables can be explained by individuals behaving similarly because of their likely exposure to the same external stimuli (e.g., individuals in the same zip code receive the same promotion from a car dealership). The use of geographic and demographic proximities to infer social relations can further exacerbate this type of endogeneity. The inclusion of individual fixed effects helps mitigate both homophily and correlated unobservables concerns. However, with observational data, researchers may need to also include time fixed effects and additional variables that help control for time-varying changes that are specific to each individual (Nair et al. 2010). In our study, all participants make their prize choices on an identical Web page within a short time frame. There are no promotions on the sports paraphernalia items we use during the raffle period, and the Bluetooth headset options are hypothetical. In fact, outside marketing activities should not concern us; our focal behavior is prize choice, which is costless to the participants, and everyone has an equal probability of winning. The only difference across the participants is the information about prize choices made by

[6] The authors provide the empirical support for their arguments via simulation studies.

their friends.[7] We thus expect the impact of correlated unobservables to be minimal.

Finally, the simultaneity issue may arise if individuals make choices at the same time. Given our emphasis on studying passive social interactions in a sequential choice context, our experimental design separates preinfluence from postinfluence choices and exploits the choice sequence information. The imposition of temporal ordering makes simultaneity not relevant in our research (Hartmann et al. 2008, Manski 2000, Narayan et al. 2011).

More recent research advocates the use of randomized experiments to identify social interaction effects. Several experimental procedures have been proposed: (i) a random assignment of group membership (Moffitt 2001); (ii) with mutually exclusive reference groups, a random assignment of members in each group to a treatment (e.g., a policy intervention) or a control condition (Moffitt 2001); (iii) a random assignment of individuals either to one of the treatment conditions where different types of social cues are present or to a control condition where the social cue is absent (Aral and Walker 2011, Bakshy et al. 2012); and (iv) for each adopter, a random assignment in which some of his or her peers receive information about the adoption and the rest receive no information (Aral and Walker 2012). Despite the merits of these approaches (Rubin 1974), a few caveats exist. The random assignment of group membership can change the structural nature of social interactions; hence, the less obtrusive nature of the last three approaches makes them more appealing. Nonetheless, the last three approaches are more appropriate for a social network wherein reference groups do not overlap. Overlapping connections across members can create leakage (Aral and Walker 2011) or incompatibility across conditions. In our social network, friendships are intertwined across participants. Therefore, it is not plausible to create nonoverlapping treatment and control conditions that allow us to conduct simple mean comparison analyses. Obtaining good data from a randomization approach also requires having an adequate sample size in each experimental condition. As a result, our research uses a statistical modeling approach where we control for initial preferences directly as an alternative to using a randomized controlled experiment.

Data Collection Procedure

As mentioned earlier, our data collection procedure consists of pre- and postinfluence stages. The preinfluence stage involves measuring initial preferences for different attributes of a university sports paraphernalia bundle and a Bluetooth headset and collecting their network information. For sports paraphernalia, we created product bundles based on four attributes: jacket style (four levels: blue style 1, grey style 1, blue style 2, and grey style 2), T-shirt color (two levels: blue and white), hat style (two levels: 1 and 2), and price (four levels: $39, $49, $59, and $69). We included six attributes for the Bluetooth headset: brand (two levels: Motorola and Plantronics), color (two levels: black and silver), weight (two levels: 8 and 13 grams), talk time per battery charge (two levels: five and eight hours), noise cancellation (two levels: yes and no), and price (four levels: $49, $59, $69, and $79). For each product, a blocked design involving 10 sets of 12 quadruples was created using an optimal design generated by SAS OPTEX (Kuhfeld et al. 1994). Participants were randomly assigned to 1 of the 10 sets and complete 12 choice tasks (each with four options). We also included a fifth baseline option of "an earplugged wired headset at $29" for the Bluetooth headset category.

To collect social network information, we first asked the participants to nominate 10 students in the class with whom they had interacted most frequently.[8] To ensure privacy,[9] we later provided the participants with an option to indicate to which nominees they were unwilling to disclose their choice information. We then invited the participants to enter the two raffles. We finally collected demographic information including gender, time spent on email daily, confidence in apparel taste, knowledge about cell phone headsets, and interest in winning the products.

In the postinfluence stage, we first launched the online raffle for sports paraphernalia. All participants received an email that included the URL that directed them to a Web page. On the page, they could select one of four sports paraphernalia bundles they wished to receive should they win the raffle. The four prize options were described by the same attributes as those used in the preinfluence stage (see Figure 1).[10] The four options were relatively close to one another with regard to their overall utilities (computed from the part-worth estimates obtained from the preinfluence stage) and were pretested with a similar pool of participants to ensure similar choice shares.

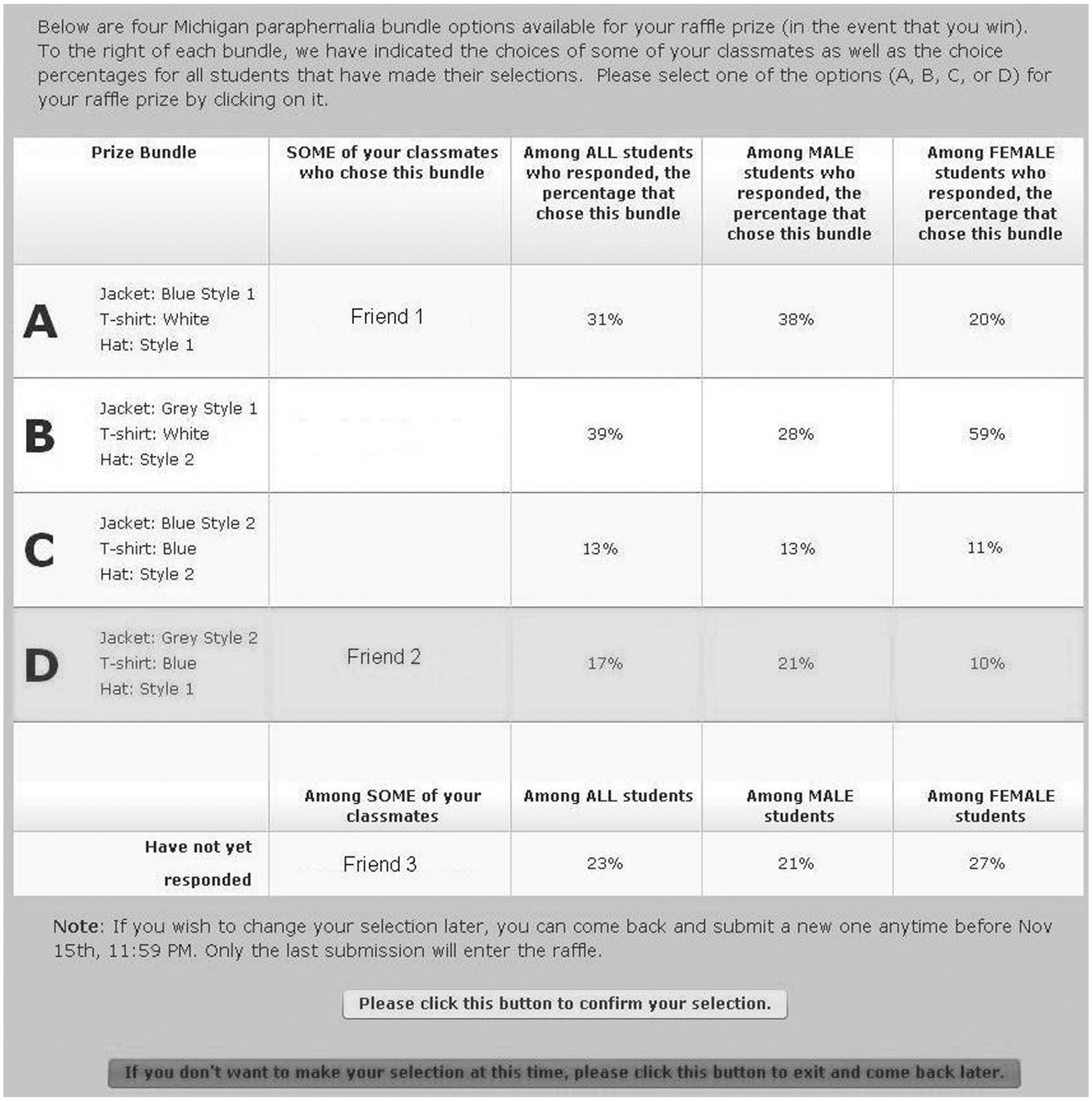
[7] The participants also see the names of their friends who have not made their choices by the time they log on to the site.

[8] To facilitate the nomination process, our Java program has a built-in search tool. Participants can easily locate the nominees' names by typing part of their first or last names. The search tool helps us avoid typos and solves the problem that participants may not necessarily remember the full name of all the nominees.

[9] This procedure was included on the request from the Institutional Review Boards of Health Sciences and Behavioral Sciences.

[10] We omitted price because including it may have prompted students to choose the most expensive bundle.

Figure 1 Web Page for Sports Paraphernalia Raffle



Below are four Michigan paraphernalia bundle options available for your raffle prize (in the event that you win). To the right of each bundle, we have indicated the choices of some of your classmates as well as the choice percentages for all students that have made their selections. Please select one of the options (A, B, C, or D) for your raffle prize by clicking on it.

Prize Bundle	SOME of your classmates who chose this bundle	Among ALL students who responded, the percentage that chose this bundle	Among MALE students who responded, the percentage that chose this bundle	Among FEMALE students who responded, the percentage that chose this bundle
A Jacket: Blue Style 1 T-shirt: White Hat: Style 1	Friend 1	31%	38%	20%
B Jacket: Grey Style 1 T-shirt: White Hat: Style 2		39%	28%	59%
C Jacket: Blue Style 2 T-shirt: Blue Hat: Style 2		13%	13%	11%
D Jacket: Grey Style 2 T-shirt: Blue Hat: Style 1	Friend 2	17%	21%	10%
	Among SOME of your classmates	**Among ALL students**	**Among MALE students**	**Among FEMALE students**
Have not yet responded	Friend 3	23%	21%	27%

Note: If you wish to change your selection later, you can come back and submit a new one anytime before Nov 15th, 11:59 PM. Only the last submission will enter the raffle.

Please click this button to confirm your selection.

If you don't want to make your selection at this time, please click this button to exit and come back later.

To facilitate choice interdependence, we showed the participants the choices that were previously made by their "friends."[11] One's friends are other participants who specified him or her as someone they had interacted with frequently and agreed to share choice information with in the preinfluence stage. For sports paraphernalia, we expected that the participants' product decisions may also be influenced by a larger-scale social interaction based on gender (i.e., gender norm) (Brock and Durlauf 2001, Miller and Prentice 1996). To examine whether the influence of gender norm can coexist with the social influence at the reference group level (i.e., global versus local social interactions), we showed each participant the aggregate proportions of other participants choosing each option. The aggregate proportions were computed based on (1) all participants who had already chosen each option, (2) only female participants who had already chosen each option, and (3) only male participants who had already chosen each option at a particular time point. The participants could change their prize choice as often as they wished before the end of the raffle. We informed them that if they did not make a choice, one of the four bundles would be drawn randomly as their prize should they win the raffle. Our database was updated in real time to ensure that the Web page offered up-to-date

[11] To further induce social interactions, the experiment featured the real-time delivery of friends' choices during the raffle period. The procedure works as follows: The first time a participant makes a choice, or every time he or she changes a choice, emails are sent out to his or her friends. The email informs each participant's friends about his or her most recent activity and urges the friends to visit the raffle page. The Web page URL is included in the email.

information about friends' activities. To increase the response rate, we sent an email reminder in the middle of the raffle.

A week later we launched the second raffle for the Bluetooth headset. The setup was similar to the paraphernalia raffle except that participants did not see the aggregate proportions of male and females who chose each Bluetooth option. We did not expect a gender norm to be relevant for this product category. In addition, the default option for the participants who did not make a prize choice was a wired headset priced at $29 (the outside option we used in the conjoint experiment in the preinfluence stage). The raffle was closed after a week. Two email reminders were sent out during this period. A winner was drawn for each raffle at the conclusion of the study.

Empirical Analysis

Descriptive Statistics of Participants and Their Social Network

Participant Profile. There were 292 participants in the preinfluence stage, 215 of which agreed to participate in the raffles in the postinfluence stage. The general demographic profiles of the participants who agreed to enter the raffles and those who did not were not significantly different, except that those in the raffles were more interested in the products they could win (see Table 1). All the analysis discussed below is based on these 215 participants.

Among the 215 raffle participants, 21.4% rated their confidence in apparel taste greater than 6 on a 1 to 7 scale, with 1 being least confident, and 18.6% rated their level of knowledge about cellphone headsets greater than 5 on a scale from 1 to 7, with 1 being not knowledgeable. More than half of the participants were "very interested" in winning a sports paraphernalia bundle, giving this a rating of 7 on a 1 to 7 scale, with 1 being "not interested all at"; the mean rating was 5.99. For the Bluetooth headset, interests were more evenly distributed, although one-third of the ratings were 7 (on a 1–7 scale), and the mean rating was 4.51.

Table 1 **Descriptive Statistics**

Variable	Participants' mean (SD)	Nonparticipants' mean (SD)	Test statistic[a] (df)
Gender	60.9% male	63.6% male	0.08 (1)
Apparel confidence	5.52 (1.18)	5.51 (1.30)	0.09 (123)
Headset knowledge	3.59 (1.75)	3.27 (1.70)	1.38 (138)
Time spent on email	38.6%[b]	28.6%	5.43 (3)
Interest in sports paraphernalia	5.99 (1.34)	3.22 (2.06)	10.05 (78)**
Interest in Bluetooth headset	4.51 (2.12)	2.35 (1.77)	8.11 (119)**

[a]A chi-square test is used for gender and time spent on email; an independent-sample t-test is used for the remaining variables.

[b]Percentage of participants spending more than an hour on email every day.

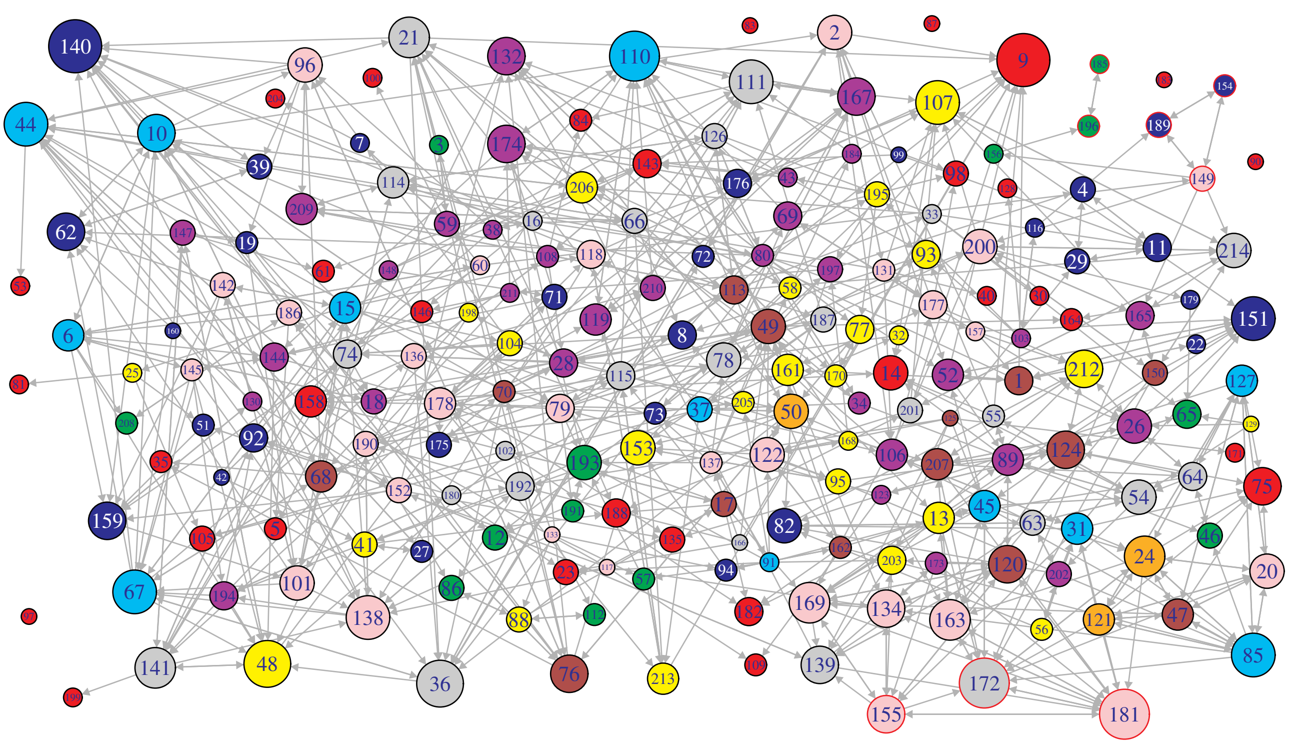
$^{**}p < 0.05$ (two-tailed).

Social Network Profile. Figure 2 depicts the directed social network among the raffle participants. Each node represents a participant. Recall that in our experiment one's friends are defined as participants who nominate him or her as one of those with whom they have frequently interacted and agree to share choice information. In other words, a participant cannot choose his or her friends. It is other participants who nominate themselves to be his or her friends. An arrow indicates the direction of the nomination. For instance, an arrow from node 110 to 2 means participant 110 nominates participant 2 and becomes one of participant 2's friend. A double-headed arrow between 185 and 196 means both participants nominate each other. Among the 628 pairs of connected participants (the same two individuals are only counted as one pair), 26.6% of them have double-headed arrows in Figure 2 (i.e., have reciprocal connections).

It is not surprising that a small group of participants in the network are substantially more popular (i.e., many other students nominate each of them as one of their friends) than the rest. We use the size of each node to represent the number of friends each participant has. The bigger the node is, the larger the number of friends he or she has (i.e., the more arrows pointing inward in Figure 2). Participants 9, 110, and 140 are examples of the popular participants. On the flip side, we use color to denote the number of others each participant nominates. Each of the three orange nodes, 24, 50, and 121, nominates nine students (nine arrows pointing outward) although they themselves do not get as many nominations from others (i.e., fewer arrows pointing inward). From this plot, we do not see a strong correlation between popularity (being nominated) and the number of nominations. Nodes of the same size can thus have a variety of colors.

The network solicitation procedure leads to participants having 3.7 friends on average. The most popular participants have 12 friends, whereas 14 participants have no friends. Table 2 provides some network structure statistics. As expected, the density, reciprocity, and transitivity are low because we worked with a relatively large network. Two characteristics of our experimental setup also contribute to the network's sparse nature. First, we asked each participant to select up to 10 friends with whom they interacted most frequently for a practical reason. Narayan et al. (2011) allow participants to nominate up to $N-1$ (the N sample size $= 70$) others as their

Figure 2 **Social Network Plot**



influencers. This would have been a timely prohibited task for us with 292 participants—and we propose a design that is applicable for a large network. Other researchers take a similar approach in restricting the number of nominations per participant (Leider et al. 2009, Nair et al. 2010). Second, because participants' choices were shared along with their names, to protect privacy, participants were given the right to opt out from sharing choice information with the friends they nominated, which makes our network even sparser. Unlike in Centola (2010, 2011) where people do not know each other before the experiment and their information is shared with the screen names they choose—which eliminates the privacy issue—our experiment hinges on actual friendship and cannot get around the privacy issue. Privacy is also not a concern in a field experiment, where focal behavior is naturally expected to be observed among peers (Aral and Walker 2011, 2012).

As noted in Table 2, 4.2% of the participants receive no nomination from other students, although they nominate at least one other student as their friend (i.e., arrow pointing outward only). We observe that 85.1% of the participants share choices with at least one friend, and 93.5% of the participants can observe at least one other participant's choice. At the same time, 14.9% (the percentage with arrows pointing inward only + the percentage that are isolated) are not willing to nominate anybody with whom they will eventually share their choice information. The major difference between them and those sharing choice information is that they are more socially peripheral; i.e., they have fewer friends sharing choice information with them (on average, 2.4 versus 3.9

Table 2 **Network Structure Statistics**

Effects	Network statistics	Values
Mean in-/out-degree	$\sum_{1\le i,j\le n} e_{ij}/n$	3.7
Variance of in-degree (prestige)	$\mathrm{var}\left(\sum_{1\le j\le n} e_{ji}\right)$	7.0
Variance of out-degree	$\mathrm{var}\left(\sum_{1\le j\le n} e_{ij}\right)$	6.0
% isolated (no arrows pointing in- or outward)	$\#\left(\sum_{1\le j\le n} e_{ji} = \sum_{1\le j\le n} e_{ij} = 0\right) \Big/ n$	2.3%
% arrows pointing outward only	$\#\left(\sum_{1\le j\le n} e_{ji} = 0,\ \sum_{1\le j\le n} e_{ij} > 0\right) \Big/ n$	4.2%
% arrows pointing inward only	$\#\left(\sum_{1\le j\le n} e_{ji} > 0,\ \sum_{1\le j\le n} e_{ij} = 0\right) \Big/ n$	12.6%
Density	$\sum_{1\le i,j\le n} e_{ij}/n(n-1)$	1.7%
% reciprocal connections	$\sum_{1\le i<j\le n} 2e_{ij}e_{ji}/n(n-1)$	0.73%
% transitive triads	$\sum_{1\le i,j,i'\le n,\, i\ne j\ne i'} e_{ij}e_{ji'}e_{ii'}/n(n-1)(n-2)$	0.007%

Note. $e_{ij} = 1$ if i nominates j, 0 otherwise.

friends). There is no significant difference in other demographic profiles except that participants not sharing are on average less knowledgeable about a cell phone headset (3.1 versus 3.7 on a 1–7 point scale).

Raffle Choice Summary. For sports paraphernalia, 167 participants visited the Web page, and 62.3% of them were male, which is similar to the proportion of males among the 215 raffle participants. Among the 25 participants who visited the page multiple times, 17 did not change their choices in all the visits, whereas the remaining 8 first did not choose any option but then picked one later. The percentages of people choosing options 1 to 4 and not choosing any of the options are 30.5%, 38.9%, 9.0%, 16.8%, and 4.8%, respectively.

For the Bluetooth headset, 97 participants visited the Web page, and 25.8% claimed to be knowledgeable (rating their knowledge greater than 5 on a 1–7 scale), which is higher than the percentage among the 215 enrolled (18.6%). The main reason could be that the participants who consider themselves knowledgeable are more interested in winning a Bluetooth headset than those who are less knowledgeable, as the mean winning interests are 5.9 and 4.2 for the more knowledgeable and less knowledgeable groups, respectively. Among the 97 participants, 92 also logged on to the paraphernalia raffle page. Thirty-four participants visited the page multiple times, and 28 of them maintained the same choices in all the visits. The remaining six first did not choose an option but later chose one. The percentages of people choosing options 1 to 4 and not choosing any of the options are 14.4%, 3.1%, 74.2%, 2.1%, and 6.2%, respectively.[12]

Analysis of Participants' Initial Preferences

We use the hierarchical Bayes logit model (Allenby and Rossi 1999) to fit the conjoint choice data from the preinfluence stage. Participant i's utility of option k is $\sum_{p=1}^{P}\beta_{ip}^{I}x_{ipq,k}^{I}+\varepsilon_{ik}^{I}$, where β_{ip}^{I} captures individual i's initial preference for the pth product attribute, $x_{ipq,k}^{I}$ indicates the pth attribute's value for option k in task q, ε_{ik}^{I} is an error term following Type I extreme value distribution, and the superscript I denotes initial preference. All the attributes are dummy-coded, including price. The model specifications for the sports paraphernalia and the Bluetooth headset are the same except that for the headset, the participants also have an option to choose an earplug version at $29. Hence, we estimate the preference (β_{i0}^{I}) for this outside option for the headset. Let n be the number of participants; let $\beta_i^I=(\beta_{i1}^I,\dots,\beta_{iP}^I)$; let $\beta^I=(\beta_1^I,\dots,\beta_n^I)$; let $y_i^I=(y_{i1}^I,\dots,y_{i12}^I)$ be a vector of i's choices in 12 conjoint choice tasks, taking values of 1 to 4 (as well as 0 for a no-choice option for the headset); let the corresponding attributes for option k in qth task be $x_{iq,k}^I=(x_{i1q,k}^I,\dots,x_{iPq,k}^I)'$, $y^I=(y_1^I,\dots,y_n^I)$, $x^I=(x_{11,1}^I,\dots,x_{n12,4}^I)$; and let P be the number of attributes. Then the likelihood is given by

$$p(y^I \mid x^I,\beta^I)=\prod_{i=1}^{n}\prod_{q=1}^{12}\frac{\exp\{\beta_i^I x_{iq,y_{iq}^I}^I\}}{\sum_{k=1}^{4}\exp\{\beta_i^I x_{iq,k}^I\}}. \tag{3}$$

We follow the conventional prior setting of a hierarchical Bayes model: $b_0\sim \mathrm{N}(0,100I_{P\times P})$, $V_0\sim \mathrm{IW}(P+2, I_{P\times P})$, and $\beta_i^I\sim \mathrm{N}(b,V_\beta)$, where $I_{P\times P}$ is a $P\times P$ identity matrix.

We estimate the model using the standard Metropolis-within-Gibbs algorithm with 10,000 iterations. According to the trace plots and autocorrelation plots, the Markov chain Monte Carlo (MCMC) chain converges and has good mixing. We keep every 10th draw from the last 8,000 of the MCMC chain for model estimation and inference. We obtain the posterior means of the hyperparameters for the part-worths. To perform a hypothesis test, we calculate the probability of a parameter being greater or less than 0 based on the kept draws. The posterior means of almost all of the part-worths are significant, indicating that the participants have preferences for actual attributes of these products, not focusing only on price when making their choices. We calculate the posterior means of the individual-specific part-worths to represent the initial preferences for each participant, and we examine the extent to which they share their preferences with their friends (i.e., evidence for homophily).[13]

We first compute the correlation between the initial preferences of each participant and his or her friend for each attribute (795 pairs[14]). Then, to maintain the same sample size, for each participant, we randomly select the same number of other participants who are not identified as one of his or her friends as the number of friends he or she has. Then we compute the correlation between the initial preferences of each participant and his or her randomly selected "nonfriend" for each attribute. In the presence of homophily, we

[12] Option 1: Motorola, black, 13 grams, 8-hour talk time, no noise cancellation; Option 2: Motorola, silver, 8 grams, 5-hour talk time, no noise cancellation; Option 3: Plantronics, silver, 13 grams, 8-hour talk time, noise cancellation; and Option 4: Plantronics, black, 8 grams, 8-hour talk time, no noise cancellation.

[13] To be more rigorous, one is expected to conduct the analysis using the actual draws from the MCMC chain to account for the uncertainty in the parameter estimates. We use the posterior means of individuals' draws to simplify our analysis.

[14] We decided to use double-counted pairs, 167 of 628 pairs, for the participant–friend pairs so as to make a comparison with the randomly selected participant pairs.

Table 3 Patterns of Participants' Initial Preferences

		Correlation in initial preferences	
Attribute	Posterior mean	Preference between participant and each friend	Preference between participant and a randomly selected nonfriend
Sports paraphernalia			
Style 1 blue jacket[a]	1.033**	0.118	0.037
Style 1 grey jacket	1.549**	0.070	0.059
Style 2 blue jacket	−0.230	0.019	0.017
Blue T-shirt[b]	0.317**	−0.015	−0.028
Style 1 hat[c]	0.241**	0.026	−0.015
$39[d]	3.466**	−0.004	−0.045
$49	2.677**	0.004	0.034
$59	1.341**	−0.037	0.007
Bluetooth headset			
Earplug headset (outside good)	−3.663**	0.103	−0.155
Motorola[e]	1.021**	−0.037	0.004
Black[f]	0.052	−0.009	0.049
8 grams[g]	1.837**	0.033	−0.035
8 hours[h]	1.895**	−0.106	0.041
Noise cancellation	2.033**	0.001	0.026
$49[i]	5.347**	0.036	−0.095
$59	3.479**	0.053	−0.066
$69	1.829**	0.055	−0.038

Notes. The baseline levels for the paraphernalia bundle are as follows: (a) style 2 grey jacket, (b) white T-shirt, (c) style 2 hat, and (d) $69. The baseline levels for the headset are as follows: (e) Plantronics, (f) silver, (g) 13 grams, (h) 5 hours, and (i) $79.

**Probability of parameter estimates being greater or less than 0 is 0.95.

should expect the correlations of the preferences from the former calculation to be large and much higher than those in the latter. Our findings are reported in Table 3, and we do not observe high correlations for either case.

Finally, given that prior research (Aral et al. 2009) suggests that homophily may be more prevalent among early adopters, we also compute the correlations of initial preferences between early adopters, defined as the first 16% of participants who made choices in each product category,[15] and their friends. The results show relatively higher correlations for a few attributes including hat style (0.107) for sports paraphernalia and the outside option (0.241) for the Bluetooth headset. Nonetheless, we do not consider these correlations to be alarming high. To summarize, we do not find strong evidence that our social network is endogenously determined by the participants' preferences, which drive the observed choices.[16]

Model Specification and Estimation

For each product, we fit two models: (i) the baseline model (Equation (3)) to the preinfluence stage conjoint data and a DMC model to the postinfluence stage raffle choice data and (ii) the same as (i) except that we fit an MRF model to the raffle data instead of the DMC model. In both cases, we estimate the baseline model and the model for the raffle data simultaneously in the MCMC chain.

Discrete-Time Markov Chain. We fit a DMC model to the complete raffle data. We only model the choices of participants who at least logged on to the raffle website and treat the rest as if they declined participating in the raffle in the preinfluence stage. However, the information from participants who did not choose any option still enters the model in the calculation of social interaction statistics for the participants who logged on. Let m be the number of participants in each raffle. To model who is likely to make a choice at time point $t+1$, c^{t+1}, we use a logit model with the deterministic part of the utility for participant i, $u_i^{t+1} = \alpha' v_i$, where v_i is a vector that includes interest in winning the raffle (rated on a scale from 1 to 7, with 1 being the least interested) and proportions of friends making any prize choice before participant i does, and α is the corresponding parameter vector. The intuition is that the participants who are more interested in winning the raffle are more likely to make their choice, and they may be prompted to make their choice if they observe that many of their friends have already made choices. The corresponding likelihood is given by

$$p(c^{t+1} = i \mid c^t, y^t, v, \alpha) = \frac{\exp\{u_{c^{t+1}}^{t+1}\}}{\sum_{i=1}^{m} \exp\{u_i^{t+1}\}}. \tag{4}$$

Conditional on $c^{t+1} = i$, we next specify the component of the model that captures which choice a participant makes. We specify the participants' choices to be influenced by their initial preferences as well as the preference shifts induced by social interactions. The initial preferences are identified by observed choices from both stages, whereas the shifts are identified only by the prize choice from the postinfluence stage. We further decompose the impact of social interactions into five parts: (i) influence from their connected friends (i.e., the reference group effect),

[15] We classify 27 (16% of 167) and 16 (16% of 97) participants who made prize choices earlier as early decision makers for sports paraphernalia and the Bluetooth headset, respectively. According to the adoption of innovation curve, 16% of consumers who first adopt a new innovation are regarded as innovators (2.5) or early adopters (13.5%).

[16] We also conducted a similar analysis with the utilities of the prize options in the raffles. That is, we used individual-specific partworths to compute these utilities before performing the correlation analysis. The results exhibit similar patterns and are available from the authors upon request.

(ii) incremental influence from friends who are more popular, (iii) incremental influence from friends with more expertise, (iv) incremental influence from early decision makers, and (v) gender norm influence (only for sports paraphernalia).

To capture the reference group effect for each participant, we use the summation of observed choices made by his or her friends normalized by the number of friends (i.e., the proportion of friends choosing each prize option). This normalization ensures that we control for different numbers of friends across different participants. This statistic is consistent with a measure called normalized degree centrality (Wasserman and Faust 1994). The statistics to capture asymmetry based on popularity and expertise are computed in an analogous manner, except that they are based only on friends who are considered popular or experts relative to participant i. Specifically, we take a relative approach instead of an absolute approach in identifying experts and popular individuals associated with participant i to avoid subjectivity from imposing certain cutoffs. Based on the self-reported ratings on confidence in apparel taste and knowledge about cell phone headsets, we define an expert as a friend who provides a higher rating on each of these two questions than a target participant i. For popularity, we define a popular friend as someone who has more friends than a target participant i. To capture the disproportionate influence of early decision makers, we include a term to allow for the incremental impact of the first choice each participant i observes. For sports paraphernalia, we also include the proportions of choices made by those with the same gender as participant i to capture the gender norm effect. Finally, we include an intercept to account for the case where the participants log on but do not choose either product. For sports paraphernalia, we inform the participants that if they do not make a prize choice, they will receive a randomly chosen option if they win. We hence include the average utility of the four prize options in the model.

Let m be the number of participants making choices in each raffle and let n be the number of participants agreeing to participate in the raffles (215); let y_i^{t+1} be participant i's choice at time $t+1$; let $u_{iy_i^{t+1}}^{t+1}$ be the utility associated with participant i choosing option y_i^{t+1}; let $x=(x_1,\ldots,x_4)$, where x_k is the vector of attributes of option k in the raffle; let $\bar{x}=\sum_{k=1}^4 x_k/4$ be the vector of mean attribute values across options, let γ be a vector of parameters associated with the social interaction effects, let $W=(w_{ji})$ be the matrix of a binary directed network with $w_{ji}=1$ if participant j is a friend of i and 0 otherwise; and let g_i be i's gender in the paraphernalia raffle, $E=(e_{ji})$, where $e_{ji}=1$ if j is a relative expert compared with i. We define H in a way similar to E, but it is associated with popularity. Participant i's opinion leader (the friend who first made a non-zero choice before i first logged on) is denoted as O_i, and t_{O_i} is the first time O_i makes a choice. The probability of a choice made at time $t+1$, $p(y_i^{t+1}\mid c^{t+1}=i, c^t, y^t, x, \beta^I, \gamma, g, W, H, E)$, has a logit form, with

$$\begin{aligned}
u_{iy_i^{t+1}}^{t+1} &= I(y_i^{t+1}=0)(\gamma_0+\gamma_1\beta_i^I\bar{x}) + I(y_i^{t+1}>0)\gamma_1\beta_i^I x_{y_i}\\
&\quad+\gamma_2\frac{\sum_{j=1}^m w_{ji} I(y_i^{t+1}=y_j^t)}{\sum_{j'=1}^n w_{j'i}}\\
&\quad+\gamma_3 I(y_i^{t+1}>0)\frac{\sum_{j=1}^m h_{ji}w_{ji} I(y_i^{t+1}=y_j^t)}{\sum_{j'=1}^m h_{j'i}w_{j'i}}\\
&\quad+\gamma_4 I(y_i^{t+1}>0)\frac{\sum_{j=1}^m e_{ji}w_{ji} I(y_i^{t+1}=y_j^t)}{\sum_{j'=1}^m e_{j'i}w_{j'i}}\\
&\quad+\gamma_5 I(y_i^{t+1}>0) I(y_i^{t+1}=y_{O_i}^{t_{O_i}})\\
&\quad+\gamma_6 I(y_i^{t+1}>0)\frac{\sum_{j=1,\neq i}^m I(y_i^{t+1}=y_j^t, g_i=g_j)}{\sum_{j=1,\neq i}^m I(y_j^t>0, g_i=g_j)}, \qquad (5)
\end{aligned}$$

where γ_0 is the intercept for the no-choice option, γ_1 is a scaling factor that adjusts the magnitude of the initial preferences, γ_2 is the baseline reference group influence, γ_3 is the asymmetric popularity influence, γ_4 is the asymmetric expert influence, γ_5 is the early decision maker influence, and γ_6 is the gender norm influence. We also define $\frac{0}{0}=0$ for all the ratios. In the Bluetooth headset raffle, because we have the information about the initial preference $-\beta_{i0}$ for a wired headset from the preinfluence data, we use $-\gamma_0\beta_{i0}$ instead of γ_0. The term $\gamma_1\beta_i^I\bar{x}$ is also not relevant.

Observing the choice sequence and assuming that only one choice can be made at a particular time point allows us to use a simple logit specification to model the choice made at each time point. The full likelihood function is thus specified as

$$\begin{aligned}
&p(s^1,\ldots,s^T\mid s^0, x, \beta^I, \gamma, g, W, H, E)\\
&\quad=\prod_{t=0}^{T-1}\frac{\exp\{u_{c^{t+1}}^{t+1}\}}{\sum_{i=1}^m \exp\{u_i^{t+1}\}}\frac{\exp\left\{u_{c^{t+1}y_{c^{t+1}}^{t+1}}^{t+1}\right\}}{\sum_{k=I(y_{c^{t+1}}^t=0)}^4 \exp\{u_{c^{t+1}k}^{t+1}\}}\\
&\qquad\cdot I(y_{-c^{t+1}}^{t+1}=y_{-c^{t+1}}^t). \qquad (6)
\end{aligned}$$

We specify $c^0=0$, meaning no one is likely to make a choice before the raffle starts, and y^0 is a vector of 0, indicating that everyone starts without any selection. We include data from all 215 participants who agreed to participant in the raffles to get more efficient estimates of individual-specific initial preferences (i.e., better shrinkage), although the number of participants in each raffle is smaller.

The full likelihood contains both preinfluence and postinfluence raffle choices, $p(s^1,\ldots,s^T\mid s^0, x, \beta^I, \gamma, g, W, H, E)p(y^I\mid x^I, \beta^I)$, where $p(y^I\mid x^I,\beta)$ is defined in Equation (4). The prior for the social interaction parameters is $\gamma^0 \sim \mathrm{N}(0, 100 I_{4\times 4})$. Although we estimate initial preferences at the individual level, we can only estimate the social interaction effects at the aggregate level, because we only observe one final choice from the raffle (participants do not change their choices after the first selection) for each participant and each product.

Markov Random Field. We fit the MRF to the snapshot data. The model specification for the MRF is similar to that of the DMC. (The detailed specification of the MRF model is in Online Appendix B.) However, the statistics corresponding to the social interaction effects are quite different. Take the baseline reference group influence γ_2 as an example. In the DMC model, when participant c^{t+1} makes a choice decision at time $t+1$, we compute the statistic as the proportion of friends who made the same choice as c^{t+1} up to time t, $(\sum_{j=1}^{m} w_{ji} I(y_i^{t+1} = y_j^t))/(\sum_{j'=1}^{n} w_{j'i})$ (recall that is a vector of everyone's choices at time t). The computation of this statistic takes the time stamp of participant c^{t+1}'s decision into account. In contrast, for MRF for each individual i, we calculate the proportion of friends who make the same choice as i observed at the end of the raffle (i.e., time T), $\sum_{j=1}^{m} w_{ji}(I(y_j^T = y_i^T))/(\sum_{i'=1}^{n} w_{j'i})$. This means that the MRF statistics associated with individual i also include individual i's friends who make choices after individual i. The statistics for each individual i can therefore take different values in the DMC and MRF models. Also, we may calculate the statistic for individual i multiple times if he or she makes decisions more than once, and the statistic each time might be different depending on his or her friends' choices up to that time point. All other statistics to capture asymmetry and gender norm are computed analogously.

Similar to the hierarchical Bayes logit model, we carry out the Metropolis-within-Gibbs algorithm for 10,000 iterations for both models (see Online Appendix D for the estimation procedure), and we use every 10th draw from the last 8,000 for parameter estimation and inference. For the MRF model, we use the approximate sampling to circumvent the intractable normalizing constant problem. Online Appendix C describes our approximate sampling procedure.

Model Comparison

In this section, we compare the DMC and MRF models using in-sample and out-of-sample hit rates. To quantify the impact of social interactions on preference shifts, we also fit a simple logit choice model instead of the DMC and MRF model to the raffle data as a baseline model. Superior fits of the DMC and MRF models, which account for social interaction effects, to those of the baseline model indicate the significant role of the social interactions in the prize choice decisions made in the social network. An in-sample hit rate is computed as the mean percentage of correctly predicted choices for all participants and an out-of-sample hit rate as the mean percentage of correctly predicted choices for out-of-sample participants. We opt to use a hit rate instead of traditional fit measures such as log marginal density and deviance information criterion because these traditional measures require the calculation of the normalizing constant for the MRF model, which is computationally expensive. To make the results more comparable across models, in the DMC model the two hit rates are computed only for the final choices, and in prediction we condition on who is likely to make a choice. To compute out-of-sample hit rates, we first fit the models to the participants who make choices earlier to predict choices of the last 20 holdout participants. The model prediction procedure is further described in Online Appendix E.

Based on both in-sample and out-of-sample hit rates, the MRF and DMC models perform better than the baseline model, especially for the Bluetooth headset category (see Table 4). These findings suggest the significance of social interaction effects particularly for a technology-related product. The participants appear to rely less on their friends' choices in making their prize selections for sports paraphernalia compared with a technology-related product, such as Bluetooth headset. For sports paraphernalia, although the in-sample hit rate of the MRF is higher than that of the DMC, the superior out-of-sample hit rate suggests better predictive performance of the DMC in the fashion-related category. Surprisingly, although we observe that the DMC and MRF provide similar in-sample hit rates, the out-of-sample hit rate of the MRF is higher than that of the DMC in the Bluetooth headset category. Given that the DMC takes advantage of additional information on the choice sequence,

Table 4 **Model Comparison**

	Measure	Baseline (%)	DMC (%)	MRF (%)
Paraphernalia bundle	In-sample hit rate	40.0 (0.11)[a]	40.2 (0.21)	43.8 (0.47)**
	Out-of-sample hit rate	38.3 (0.17)	43.3 (0.25)**	40.6 (0.32)**
Bluetooth headset	In-sample hit rate	28.4 (0.27)	41.1 (0.58)**	41.1 (0.25)**
	Out-of-sample hit rate	22.5 (0.18)	43.5 (0.33)**	46.0 (0.47)**

Notes. Within parentheses is the standard error of the hit rate, computed from the kept draws of the posterior predictive distribution. We use two independent sample t-tests (assuming that the central limit theorem holds) to determine whether one model gives a significantly better hit rate than the other.

**Significantly greater than hit rate of baseline model at the significance level of 0.05.

we expect the DMC to provide better predictive performance than the MRF.

One possible explanation could lie in the relationship between the DMC and the MRF. That is, the MRF can be regarded as an equilibrium model for the DMC, as also shown by our proof in Online Appendix A. Yang et al. (2010) propose an automodel, a special case of the MRF, to model coordinated choice. They refer to their model as an equivalent conditional approach to Hartmann's (2010) unconditional or equilibrium approach to modeling coordinated choice. It is possible that for a technology-related product, individuals may make choices based on how they have anticipated choices to be realized at the equilibrium (e.g., an option with the most sophisticated technology will be the one their relatively more high-tech friends adopt). Therefore, we may not lose much from the prediction standpoint using the MRF with snapshot data as opposed to the DMC with complete data. Given these results, for the Bluetooth headset we draw conclusions from the DMC model with some caution. Despite its inferior out-of-sample prediction, the exploitation of the choice sequence information using the DMC model alleviates the simultaneity concern. The findings from the DMC model are also more insightful. That is, we can explore the potential disproportionate influence of early decision makers in the social interaction process. We can also examine who is likely to make a choice. This information can be useful for firms to target influential consumers. Neither of these insights can be drawn from the MRF model.

Estimation Results

Tables 5 and 6 report the posterior means of the model parameters for sports paraphernalia and the Bluetooth headset, respectively. We omit the estimates for the initial preferences based on the preinfluence and postinfluence data in all the models, because they lead to the same inferences as to the estimates based only on the preinfluence data (see Table 3). We present results from the DMC model with different specifications (Models 1–4) and the MRF model (Model 5). Model 1 is our proposed model (Equation (5)). In Model 2, the numbers, instead of the proportion, of friends are used to as statistics to capture social interactions.

Using the proportion of friends (i.e., average friends' behavior) to capture social interactions is consistent with the majority of prior research (e.g., Brock and Durlauf 2001, Nair et al. 2010, Tucker 2008). However, some research that allows participants to discuss preferences between the preinfluence and postinfluence stages uses the average initial preference, instead of behavior, of a reference group (Aribarg et al. 2002, 2010; Arora and Allenby 1999) as a statistic. The decision to use the average behavior is driven by our primary mechanism to induce choice interdependence—that is, the observation of friends' choices on a Web page. However, the participants may have discussed their preferences off-line. We add the average preference of a participant's friends who have made prize choices before him or her in Model 3. We assume that the participants only rely on information from their friends who have made choice commitments. Finally, we further examine the potential impact of homophily on our results by estimating Model 4, where we do not control for the participants' initial preferences in the estimation of social interaction effects.

For sports paraphernalia, the results from the MRF (Model 5) and DMC (Model 1) are quite similar.

Table 5 Parameter Estimates: Sports Paraphernalia

	DMC				
Parameter	Proposed model (Model 1)	Statistics based on no. of friends (Model 2)	Controlling for avg. friends' preference (Model 3)	Not controlling for initial preference (Model 4)	MRF (Model 5)
Social interaction effects					
Outside good	0.524**	0.528**	0.510**	n/a	0.301**
Scale parameter for initial preference (γ_1)	0.517**	0.516**	0.501**	n/a	0.418**
Overall reference group influence ($\gamma 2$)	−0.009	−0.020	−0.184	−0.174	0.354
Asymmetric popularity influence ($\gamma 3$)	0.766**	0.492**	0.839**	0.416	0.604
Asymmetric expertise influence ($\gamma 4$)	−0.103	0.024	−0.229	0.242	−0.526
Early decision maker influence (γ_5)	0.053	0.056	0.008	0.268	n/a
Gender norm (γ_6)	1.892**	1.899**	1.730**	2.211**	1.510**
Intrinsic preference of the reference group (γ_7)	n/a	n/a	0.194**	n/a	n/a
Factors determining who is likely to make a choice decision					
% friends made a choice	0.377*	0.029	0.349*	0.363*	n/a
Interest in winning the raffle	0.010	0.004	0.010	0.012	n/a

Note. Baseline levels are defined in Table 3.
*Probability of parameter estimates being greater or less than 0 is 0.9; **probability of parameter estimates being greater or less than 0 is 0.95.

Table 6 Parameter Estimates: Bluetooth Headset

	DMC				
Parameter	Proposed model (Model 1)	Statistics based on no. of friends (Model 2)	Controlling for avg. friends' preference (Model 3)	Not controlling for initial preference (Model 4)	MRF (Model 5)
Social interaction effects					
Outside good	0.075*	0.090*	0.106*	n/a	−0.038
Scale parameter for initial preference (γ_1)	0.394**	0.430**	0.373**	n/a	0.316**
Overall reference group influence (γ2)	0.576	0.285**	0.664	−0.429	0.127
Asymmetric popularity influence (γ3)	0.390	−0.248	0.069	0.149	1.598**
Asymmetric expertise influence (γ4)	1.034**	0.255	0.822**	1.071**	1.235**
Early decision maker influence (γ_5)	0.805**	1.270**	1.216**	1.646**	n/a
Initial preference of the reference group (γ_7)	n/a	n/a	−0.165	n/a	n/a
Factors determining who is likely to make a choice decision					
% friends made a choice	0.011	0.036	0.035	0.034	n/a
Interest in winning the raffle	0.055*	0.056*	0.057**	0.054	n/a

Note. Baseline levels are defined in Table 3.
*Probability of parameter estimates being greater or less that 0 is 0.9; **probability of parameter estimates being greater or less that 0 is 0.95.

Given its superior out-of-sample fit, the DMC model becomes our focus. The participants' initial preference (γ_1) is found to be significant in determining participants' prize choices, suggesting that they make prize choices based on their own evaluations of options' utilities. Interestingly, the baseline reference group effect (γ_2) is not significant. Nonetheless, as expected for a fashion-related product where value-expressive influence is more prominent, we find the asymmetric popularity effect (γ_3) to be significant and the asymmetric expertise effect to be insignificant (γ_4). These results suggest that participants are affected by only friends in their reference group who are more popular than them (i.e., have more friends) in this product category. We also do not find early decision makers (γ_6) to be disproportionately more influential but did find gender norm (γ_5) to also play a significant role. Finally, the results from the who is likely to make a choice part indicates that observing their friends making prize choices prompts participants to also make their prize choices.

The results from the model based on the number of friends (Model 2) and that controlling for the initial preference of the reference group (Model 3) lead to similar conclusions as Model 1. The significant effect of the initial preference of the reference group suggests possible off-line communication among the participants. Interestingly, the MRF model (Model 5), which is fit to the incomplete snapshot data, cannot detect the significant effect of asymmetric popularity.

Looking at the proposed model (Model 1) for the Bluetooth headset, similar to sports paraphernalia, the initial preference plays a significant role (γ_1); the baseline reference group effect does not (γ_2). As expected for a technology-related product where informational influence dominates, the asymmetric expertise effect is significant (γ_4), whereas the asymmetric popularity effect is not (γ_3). Nonetheless, the MRF model (Model 5), which provides a better out-of-sample fit, leads to a slightly different conclusion. That is, the MRF suggests significance of both the asymmetric expertise and popularity effects (the parameters related to the two types of asymmetry are also not significantly different from each other). Despite the better out-of-sample fit, the results based on the DMC better represent the actual social interaction process because it exploits the complete data. To be conservative, we urge future research to reinvestigate the presence of the asymmetric popularity effect in the technology-related category.

Unlike sports paraphernalia, we find the disproportionate influence of early decision makers (γ_6) for the Bluetooth headset. That is, the participants are more significantly influenced by the first choice they observe compared with choices observed later. We also find that the participants who are more interested in the raffle are also more likely to make the prize choice early. Taken together, these results suggest that for a technology-related product, individuals who are more interested in a technology product are likely to adopt the product earlier than others. These early adopters also exert a disproportionate influence on later adopters (i.e., taking an opinion leadership role). These conclusions are consistent with previous theoretical and meta-analysis work (Van den Bulte and Joshi 2007) that touts the important role of opinion leaders in the area of diffusion of innovations. Our results complement prior research, as we provide empirical evidence for these effects at the individual level.

Compared with other specifications of the DMC model, we do not find a significant effect of the initial preference of the reference group in this category (Model 3). Using the number of friends (Model 2) also leads to a different conclusion that there exists only the baseline reference group effect, but not the asymmetric popularity or expertise effect. We draw our conclusions based on our proposed DMC model (Model 1) instead of Model 3, because the statistics in Model 1 are more consistent with the normalized degree centrality measure widely used in the sociology literature (Wasserman and Faust 1994) to capture social network effects.

For both products, we find controlling for the participants' initial preferences is more important for sports paraphernalia than for Bluetooth headset. Despite the difference in the magnitude of parameter estimates, the estimates from the model that does not control for the initial preference (Model 4) for Bluetooth headset suggest a similar pattern of social interaction effects as that from Model 1. For sports paraphernalia, however, Model 4 does not detect the asymmetric popularity effect. These findings indicate that even with low correlations in the initial preferences within a reference group (see Table 3), researchers can benefit from controlling for heterogeneous initial preferences across individuals in the estimation of social interaction effects. It is noteworthy that the observed low correlations and absence of discernible differences in the correlations between the participants and their friends' initial preferences and those between the participants and randomly selected nonfriends' initial preferences may also be masked by the uncertainty in the part-worth estimates at the individual level.

Discussion

Our paper delineates an experimental and modeling framework to examine social interaction effects. Relying on the multiattribute utility framework and two-stage conjoint experiment where we measure individuals' initial attribute preferences and collect exogenous social network information, we can separately identify choice interdependence—shifts in choice preferences as a result of social interactions—in a social network. Our conjoint experiment coupled with statistical modeling is an alternative approach to a controlled randomized experiment, particularly when reference groups largely overlap in a social network. Our findings prompt the importance of controlling for individual initial preferences when examining social interaction effects, even with the incorporation of exogenous network information.

Previous behavioral research has recognized the differences between informational and value-expressive influences (Burnkrant and Cousineau 1975, Deutsch and Gerard 1955, Kelman 1958) and extensively examined different factors that moderate the impact of these two types of influence. These factors can be categorized as those related to the characteristics of the influenced individuals, such as the extent of self-monitoring and susceptibility to social influence (Batra et al. 2001, Bearden and Rose 1990, Bearden et al. 1989, DeBono and Harnish 1988); the characteristics of the influencers, such as their credibility and attractiveness (Kelman 1958); and the characteristics of information or decision contexts, such as the transparency of decisions to others, the uniformity of opinions, and product necessity and conspicuousness (Bearden and Etzel 1982, Burnkrant and Cousineau 1975, Childers and Rao 1992, Cohen and Golden 1972). Extending previous research, we explore how the interaction between the type of product and the characteristics of influencers moderates the two types of social influence. Specifically, we examine differential asymmetric influences based on popularity and expertise for fashion-related versus technology-related products. We find popular individuals exert greater influence for a fashion-related product, but experts exert greater influence for a technology-related product. Nonetheless, we urge future research to reexamine whether the asymmetric popularity effect can also exist in the technology-related product category given our results from the MRF model, which provides good predictive performance.

Given the choice sequence information, our findings show that choices made by early decision makers are more influential than choices made by later decision makers, and early decision makers are those who are more interested in owning the product in the technology-related category. Previous research views a combination of expertise or knowledge, high category involvement, and influence as characterizing opinion leaders (Jacoby and Hoyer 1981, Midgley 1976, Robertson et al. 1984). Our result is in line with the previous research and suggests the important role of early decision makers as highly involved and knowledgeable opinion leaders for technology-related products. Finally, in the absence of the choice sequence information, we demonstrate that the MRF model can be used as an alternative to the DMC model. In fact, the MRF model provides good predictive performance for the technology-related product.

As avenues for future research, our design can be further benefited by a technique to solicit social network information that forces individuals to be more accurate in their friend nominations (Leider et al. 2009). Our research also does not account for potential nonlinear social interaction dynamics, where the marginal effect of observing behavior of one more friend may not be constant across different numbers of friends (Aral and Walker 2011, Aral et al. 2009, Bakshy et al. 2012). In the context of nonoverlapping reference groups, future research may combine a statistical modeling approach with a two-stage conjoint experiment such as ours with a randomized experimental approach. For example, for each person who makes a choice, researchers can randomly send the choice information to only a subset of his or her friends. One benefit of a modeling approach is its ability to facilitate researchers in performing predictive and counterfactual exercises. Finally, in this paper we propose a DMC to model the sequences of choices (i.e., who is likely to make a choice). However, additional information on the time between two consecutive choices will allow us to also model how long it takes for someone to make a choice. It is generally difficult to track the time between two consecutive choices. Even with social media technology, it can become cumbersome to record large amounts of data. However, if such data are available, a continuous-time, instead of a discrete-time, Markov chain model can be applied to study the factors that influence the time it takes for someone to make a choice. Modeling this type of data is beyond the scope of our paper, but it could be a fruitful avenue for future research.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mksc.2013.0811.

Acknowledgments

The paper is based on the first author's dissertation. The authors thank Stefan Witwicki for his assistance in Java and Flash programming to facilitate their data collection process. They thank Natasha Zhang Foutz for great comments on this paper. Finally, they thank the review team for providing them with valuable comments and directions for the paper.

References

Algesheimer R, Borle S, Dholakia UM, Singh SS (2010) The impact of customer community participation on customer behaviors: An empirical investigation. *Marketing Sci.* 29(4):756–769.

Allenby GM, Rossi PE (1999) Marketing models of consumer heterogeneity. *J. Econom.* 89(1–2):57–78.

Ansari A, Koenigsberg A, Stahl F (2011) Modeling multiple relationships in social networks. *J. Marketing Res.* 48(4):713–728.

Aral S, Walker D (2011) Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Sci.* 57(9):1623–1639.

Aral S, Walker D (2012) Identifying influential and susceptible members of social networks. *Science* 337(6092):337–341.

Aral S, Muchnik L, Sundararajan A (2009) Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc. Natl. Acad. Sci. USA* 106(51):21544–21549.

Aribarg A, Arora N, Bodur HO (2002) Understanding the role of preference revision and concession in group decisions. *J. Marketing Res.* 39(3):336–349.

Aribarg A, Arora N, Kang MY (2010) Predicting joint choice using individual data. *Marketing Sci.* 29(1):139–157.

Arora N, Allenby GM (1999) Measuring the influence of individual preference structure in group decision making. *J. Marketing Res.* 36(4):476–478.

Bakshy E, Eckles D, Yan R, Rosenn I (2012) Social influence in social advertising: Evidence from field experiments. *Proc. 13th ACM Conf. Electronic Commerce* (ACM, New York), 146–161.

Bass F (1969) A new product growth model for consumer durables. *Management Sci.* 15(5):215–227.

Batra R, Homer PM, Kahle LR (2001) Values, susceptibility to normative influence, and attribute importance weights: A nomological analysis. *J. Consumer Psych.* 11(2):115–128.

Bearden WO, Etzel MJ (1982) Reference group influence on product and brand purchase decisions. *J. Consumer Res.* 9(2):183–194.

Bearden WO, Rose RL (1990) Attention to social comparison information: An individual difference factor affecting consumer conformity. *J. Consumer Res.* 16(4):461–471.

Bearden WO, Netemeyer RG, Teel JE (1989) Measurement of consumer susceptibility to interpersonal influence. *J. Consumer Res.* 15(4):473–481.

Belk RW (1988) Possession and the extended self. *J. Consumer Res.* 15(2):139–168.

Bell DR, Song S (2007) Neighborhood effects and trail on the Internet: Evidence from online grocery retailing. *Quant. Marketing Econom.* 5(4):361–400.

Besag J (1974) Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* 36(2):192–236.

Besag J (1975) Statistical analysis of non-lattice data. *J. Roy. Statist. Soc. Ser. D* 24(3):179–195.

Bourne FS (1957) Group influence in marketing and public relations. Likert R, Hayes SP, eds. *Some Applications of Behavioral Research* (UNESCO, Basel, Switzerland), 207–255.

Bramoullé Y, Djebbari H, Fortin B (2009) Identification of peer effects through social network. *J. Econometrics* 150(1):41–55.

Brock WA, Durlauf SN (2001) Discrete choice with social interactions. *Rev. Econom. Stud.* 68:235–260.

Burnkrant RE, Cousineau A (1975) Informational and normative social influence in buyer behavior. *J. Marketing Res.* 2(3): 206–215.

Centola D (2010) The spread of behavior in an online social network experiment. *Science* 329(5996):1194–1197.

Centola D (2011) An experimental study of homophily in the adoption of health behavior. *Science* 334(6060):1269–1272.

Chen Y, Wang Q, Xie J (2011) Online social interactions: A natural experiment on word of mouth versus observational learning. *J. Marketing Res.* 48(2):238–254.

Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3):345–354.

Childers TL, Rao AR (1992) The influence of familial and peer-based reference groups on consumer decisions. *J. Consumer Res.* 19(2):198–211.

Cohen JB, Golden E (1972) Informational social influence and product evaluation. *J. Appl. Psych.* 56(1):54–59.

Czepiel JA (1974) Word-of-mouth processes in the diffusion of a major technological innovation. *J. Marketing Res.* 11(2):172–180.

De Bruyn A, Lilien GL (2008) A multi-stage model of word-of-mouth influence through viral marketing. *Internat. J. Res. Marketing* 25(3):151–163.

DeBono KG, Harnish RJ (1988) Source expertise, source attractiveness, and the processing of persuasive information: A functional approach. *J. Personality Soc. Psych.* 55(4):541–546.

Deutsch M, Gerard HB (1955) A study of normative and informational social influences upon individual judgment. *J. Abnormal Soc. Psych.* 51(3):626–636.

Ding M (2007) An incentive-aligned mechanism for conjoint analysis. *J. Marketing Res.* 44(May):214–223.

Godes D, Mayzlin D (2004) Using online conversations to study word-of-mouth communication. *Marketing Sci.* 23(4):545–560.

Godes D, Mayzlin D, Chen Y, Das S, Dellarocas C, Pfeiffer B, Libai B, Sen S, Shi M, Verlegh P (2005) The firm's management of social interactions. *Marketing Lett.* 16(3–4):415–428.

Granovetter M (1973) The strength of weak ties. *Amer. J. Sociol.* 78(6):1360–1380.

Granovetter M (1978) Threshold models of collective behavior. *Amer. J. Sociol.* 83(6):1420–1443.

Haaijer R, Wedel M (2003) Conjoint choice experiments: General characteristics and alternative model specifications. Gustafsson A, Herrmann A, Huber F, eds. *Conjoint Measurement Methods and Applications*, 3rd ed. (Springer, New York), 371–412.

Hartmann WR, Nair H, Manchanda P, Bothner M, Dodds P, Godes D, Hosanagar K, Tucker C (2008) Modeling social interactions: Identification, empirical methods and policy implication. *Marketing Lett.* 19(3):287–304.

Hartmann WR (2010) Demand estimation with social interactions and the implications for targeted marketing. *Marketing Sci.* 29(4):585–601.

Hill S, Provost F, Volinsky C (2006) Network-based marketing: Identifying likely adopters via consumer networks. *Statist. Sci.* 21(2):256–276.

Hoff PD, Raftery AE, Handcock MS (2002) Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* 97(460): 1090–1098.

Hunter DR (2007) Curved exponential family models for social networks. *Soc. Networks* 29(2):216–230.

Iacobucci D, Hopkins N (1992) Modeling dyadic interactions and networks in marketing. *J. Marketing Res.* 29(1):5–17.

Jacoby J, Hoyer WD (1981) What is opinion leaders didn't know more? A question of nomological validity. Monroe KB, ed. *Advance in Consumer Research*, Vol. 8 (Association for Consumer Research, Ann Arbor, MI), 299–303.

Katz E, Lazarsfeld PF (1955) *Personal Influence: The Part Played by People in the Flow of Mass Communications* (Free Press, New York).

Keeney RL, Raiffa H (1993) *Decision with Multiple Objectives* (Cambridge University Press, New York).

Kelman HC (1958) Compliance, identification, and internalization: Three processes of attitude change. *J. Conflict Resolution* 2(1):51–60.

Kelman HC (1961) Process of opinion change. *Public Opinion Quart.* 25:57–78.

Koskinen JH, Snijders TAB (2007) Bayesian inference for dynamic social network data. *J. Statist. Plan. Inference* 137(12):3930–3938.

Kuhfeld WF, Tobias RD, Garratt M (1994) Efficient experimental design with marketing research applications. *J. Marketing Res.* 31(4):545–557.

Leider S, Mobius MM, Rosenblat T, Do Q-A (2009) Directed altruism and enforced reciprocity in social networks. *Quart. J. Econom.* 124(4):1815–1851.

Manchanda P, Packard G, Pattbhiramaiah A (2013) Social dollars: The economic impact of customer participation in a firm-sponsored online community. Working paper, University of Michigan, Ann Arbor.

Manski CF (1993) Identification of endogenous social effects: The reflection problem. *Rev. Econom. Stud.* 60(3):531–542.

Manski CF (2000) Economic analysis of social interactions. *J. Econom. Perspect.* 14(3):115–136.

McGuire WJ (1969) The nature of attitudes and attitude change. Lindzey G, Aronson E, eds. *The Handbook of Social Psychology*, 2nd ed., Vol. 3 (Addison-Wesley Publishing, Reading, MA), 136–314.

Midgley DF (1976) A simple mathematical theory of innovation behavior. *J. Consumer Res.* 3(1):31–41.

Miller CM, McIntyre SH, Mantrala MK (1993) Toward formalizing fashion theory. *J. Marketing Res.* 30(2):142–157.

Miller DT, Prentice DA (1996) The construction of social norms and standards. Higgins ET, Kruglanski AW, eds. *Social Psychology: Handbook of Basic Principles* (Guilford Press, New York), 799–829.

Moffitt R (2001) Policy interventions, low-level equilibria, and social interactions. Durlauf S, Young P, eds. *Social Dynamics* (Brookings Institution Press, Washington, DC), 45–82.

Moon S, Russell GJ (2008) Predicting product purchase from inferred customer similarity: An autologistic model approach. *Management Sci.* 54(1):71–82.

Murray I, Ghahramani Z, MacKay DJC (2006) MCMC for doubly-intractable distributions. *22nd Conf. Uncertainty in Artificial Intelligence (UAI 2006), July 13–16, Cambridge, MA*.

Myers JH, Robertson TS (1972) Dimensions of opinion leadership. *J. Marketing Res.* 9(1):41–46.

Nair H, Manchanda P, Bhatia T (2010) Asymmetric social interactions in physician prescription behavior: The role of opinion leaders. *J. Marketing Res.* 47(5):883–895.

Nam S, Manchanda P, Chintagunta PK (2010) The effect of signal quality and contiguous word of mouth on customer acquisition for a video-on-demand service. *Marketing Sci.* 29(4): 690–700.

Narayan V, Rao VR, Saunders C (2011) How peer influence affects attribute preferences: A Bayesian updating mechanism. *Marketing Sci.* 30(2):368–384.

Niraj R, Padmanabhan V, Seetharaman PB (2008) Cross-category model of households' incidence and quantity decisions. *Marketing Sci.* 27(2):225–235.

Park CW, Lessig VP (1977) Students and housewives: Differences in susceptibility to reference group influence. *J. Consumer Res.* 4(2):102–110.

Reingen PH, Kernan JB (1986) Analysis of referral networks in marketing: Methods and illustration. *J. Marketing Res.* 23(4): 370–378.

Reingen PH, Foster BL, Brown JJ, Seidman SB (1984) Brand congruence in interpersonal relations: A social network analysis. *J. Consumer Res.* 11(3):771–783.

Robertson TM, Zielinski J, Ward S (1984) *Consumer Behavior* (Scott, Foresman and Company, Glenview, IL).

Robins G, Pattison P, Kalish Y, Lusher D (2007) An introduction to exponential random graph (p^*) models for social network. *Soc. Networks* 29(2):173–191.

Rogers EM, Cartano DG (1962) Methods of measuring opinion leadership. *Public Opinion Quart.* 26(3):435–441.

Ross SM (1996) *Stochastic Processes* (John Wiley & Sons, New York).

Rubin DB (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Ed. Psych.* 66(5): 688–701.

Russell G, Peterson A (2000) Analysis of cross category dependence in market basket selection. *J. Retailing* 76(3):367–392.

Schelling T (1971) Dynamic models of segregation. *J. Math. Soc.* 1(2):143–186.

Shalizi CR, Thomas AC (2011) Homophily and contagion are generically confounded in observational social network studies. *Soc. Methods Res.* 40(2):211–239.

Snijders TAB, Steglich C, Schweinberger M (2007) Modeling the coevolution of networks and behavior. van Montfort K, Oud H, Satorra A, eds. *Longitudinal Models in the Behavioral and Related Sciences* (Lawrence Erlbaum Associates, Mahwah, NJ), 41–71.

Steglich C, Tom EG, Snijders AB, Pearson M (2010) Dynamic networks and behavior: Separating selection from influence. *Soc. Methodol.* 40(1):329–393.

Stephen AT, Toubia O (2010) Deriving value from social commerce networks. *J. Marketing Res.* 47(2):215–228.

Song I, Chintagunta P (2006) Measuring cross-category price effects with aggregate store data. *Management Sci.* 52(10):1594–1609.

Toubia O, Stephen AT, Freud A (2011) Viral marketing: A large-scale field experiment. *Econom., Management, Financial Markets* 6(3):43–65.

Trusov M, Bodapati AV, Bucklin RE (2010) Determining influential users in Internet social networks. *J. Marketing Res.* 47(4): 643–658.

Tucker C (2008) Identifying formal and informal influence in technology adoption with network externalities. *Management Sci.* 54(12):2024–2038.

Van den Bulte C, Joshi YV (2007) New product diffusion with influentials and imitators. *Marketing Sci.* 26(3):400–421.

Van den Bulte C, Stremersch S (2004) Social contagion and income heterogeneity in new product diffusion: A meta-analytic test. *Marketing Sci.* 23(4):530–544.

Van den Bulte C, Wuyts S (2007) *Social Network and Marketing. Marketing Science Institute Relevant Knowledge Series* (Marketing Science Institute, Cambridge, MA).

Wang J, Atchadé YF (2014) Approximate Bayesian computation for exponential random graph models for large social networks. *Commun. Statist. Simulat.* 43(2). Forthcoming.

Wasserman S, Faust K (1994) *Social Network Analysis: Methods and Applications* (Cambridge University Press, New York).

Wasserman S, Pattison P (1996) Logit models and logistic regressions for social networks: An introduction to Markov random graph and p^*. *Psychometrika* 61(3):401–425.

Watts DJ, Dodds PS (2007) Influentials, networks, and public opinion formation. *J. Consumer Res.* 34(4):441–458.

Yang S, Allenby GM (2003) Modeling interdependent consumer preferences. *J. Marketing Res.* 40(3):282–294.

Yang S, Zhao Y, Erdem T, Zhao Y (2010) Modeling the intra-household behavioral interaction. *J. Marketing Res.* 47(3): 470–484.

Zhang J (2010) The sound of silence: Observational learning in the U.S. kidney market. *Marketing Sci.* 29(2):315–335.

Zhu F, Zhang X (2010) Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *J. Marketing.* 74(2):133–148.