

Demand Estimation with Social Interactions and the Implications for Targeted Marketing*

Wesley R. Hartmann

Stanford Graduate School of Business
518 Memorial Way
Stanford, CA 94305-5015
Hartmann_Wesley@gsb.stanford.edu

June 19, 2009

Abstract

This paper develops a model for the estimation and analysis of demand in the context of social interactions. Decisions made by a group of customers are modeled to be an equilibrium outcome of an empirical discrete game, such that all group members must be satisfied with chosen outcomes. The game-theoretic approach assists estimation by allowing us to account for the endogeneity of group members' decisions, while also serving as a managerial tool that can simulate equilibrium outcomes for the group when the firm alters the marketing mix to the group. The model builds upon the existing literature on empirical models of discrete games by introducing a random coefficients heterogeneity distribution. Monte Carlo simulations reveal that including the heterogeneity resolves the endogenous group formation bias commonly noted in the social interactions literature. By estimating the heterogeneous equilibrium model using Bayesian hierarchical MCMC, we can also recover some parameters at the individual-level to evaluate group specific characteristics and targeted marketing strategies. To validate the model and illustrate its implications, we apply it to a data set of groups of golfers. We find significant social interaction effects, such that 65% of the median customer value is attributable to the customer, while the other 35% is attributable to the customer's affect on members of his group. We also consider targeted marketing strategies and show that group-level targeting increases profit by 1%, while targeting within groups can increase profitability by 20%. We recognize that customer backlashes to targeting could be greater when group members receive different offers, so we suggest some alternatives that could retain some of the profitability of within group targeting while avoiding customer backlashes.

Keywords: decision-making, interdependent preferences, consumption, discrete choice, social interactions, targeted marketing, customer relationships.

* Thanks to Greg Allenby, Lanier Benkard, Gunter Hitsch, Sanjog Misra, Harikesh Nair, Sridhar Narayanan and Christophe Van den Bulte, as well as seminar participants at Ohio State University and The Wharton School for their useful comments. I would also like to thank Taylan Yildiz for research assistance on an early variant of this paper titled "A Structural Analysis of Joint Decision-Making." The usual disclaimer applies.

1 Introduction

With the recent growth of social networking websites and information technology that can track consumption behavior, marketers are increasingly conducting research in the area of social interactions. While most of the work has been in the areas of diffusion (e.g. Van den Bulte and Stremersch, 2004) and word-of-mouth (Godes et.al., 2005), there is relatively little work extending choice models to analyze data in which the decision of one customer is directly influenced by the decisions of, as opposed to the referrals from, one or more of the customer's peers. Such interactions are however common. Individuals are often observed to eat together at restaurants, watch movies together, play games online together or choose common wireless phone plans to take advantage of in-network calling discounts. In each of these examples, individuals may differ in their preferences, but prefer to coordinate the decisions to yield either social or pecuniary benefits.

From the marketer's perspective, there are unique implications when individuals within a group coordinate their decisions. Like a word-of-mouth social interaction, the lifetime value of any given customer will depend on the customer's demand and others' demand that is generated by the customer's social interactions. However, unlike word-of-mouth where the social interaction is primarily relevant for customer acquisition, coordinated decision-making implies that the social interactions are relevant for every purchase occasion. In such cases, a marketer cannot invest heavily in an initial social interaction such as a referral with the hope that it will

generate a long stream of future revenue.¹ Rather, the marketer's focus is on incorporating the social interaction into all marketing mix decisions.

This paper develops a model that can be used to estimate social interactions and analyze their implications. One important challenge in developing such a model is to extend the typical discrete choice model to include the decisions of a customer's peers. To do this we recognize that coordinating decisions within a group involves individuals weighing their personal preferences against those of their peers to find decisions that none of the group members has an incentive to deviate from. In other words, we model their decisions as the equilibrium outcome of a coordination game. This characterization of group decision-making is common in the economics literature and is most closely related to the "Battle of the Sexes" game.² To make this empirically tractable, we incorporate the coordination game directly into an econometric model. Hence, the model falls within the broad class of empirical models of discrete games defined by Bresnahan and Reiss (1991), and is most comparable to equilibrium models of social interactions defined by Brock and Durlauf (2001) and Soetevent and Kooreman (2007).

The equilibrium model serves two purposes. First, the model provides an econometric benefit by accounting for the fact that when a partner's decision enters the payoffs of a focal customer's choice equation, it is an endogenous variable.³ The model deals with this by treating all of the group members' choices as a single outcome of a simultaneous move game with

¹ Kumar, Petersen and Leone (2007) provide a discussion of how referrals can be incorporated into customer lifetime value estimates.

² Fudenberg and Tirole (1991) describe an example of the "battle of the sexes" in which two players "wish to go to an event together, but disagree about whether to go to a football game or the ballet." Our model is similar, except that from the perspective of the firm, the interest is whether one, both, or neither buy. The no purchase option may or may not have a coordination benefit. In the applications we consider, no purchase occurs with a high frequency such that it is safe to assume that if a customer does not purchase, her utility is unlikely to be affected by whether or not the other person purchases (i.e. social interaction benefits only arise when both purchase).

³ i.e. the partner's decision is a function of the unobserved preferences of the focal customer and hence not exogenous to the error term in the equation for the focal customer.

complete information.⁴ Second, the equilibrium model allows us to conduct counterfactuals to measure customer values and the effect of marketing mix activities in the presence of social interactions. Therefore, the model is a useful way for managers to evaluate potential social interactions, even if customer data is not available. Customer data provide the benefit of measuring preference parameters to focus on empirically relevant predictions of the model, rather than the full range of comparative statics that the parameter space may yield.

Another concern that arises when developing models to estimate social interactions is that unobserved customer heterogeneity may be correlated between the members of a group. As is well documented in the sociology literature, individuals tend to form groups with similar individuals, i.e. homophily. This implies a natural correlation in the behavior of individuals that could be incorrectly interpreted as a causal social interaction. Nair et.al. (2007) account for this potential confound by including individual-level fixed effects in their linear model. In this paper, we also exploit the panel structure of data, but do so using a hierarchical Bayesian MCMC approach. Through model simulations, we illustrate that incorporating a rich random coefficients heterogeneity specification into equilibrium models of social interactions resolves a large potential bias. The Bayesian approach has the added advantage of allowing us to recover individual-level parameter estimates to evaluate targeted marketing strategies as in Rossi, McCulloch and Allenby (1996).

To validate our model and analyze the marketing implications of the social interactions, we apply it to a data set of groups of golfers. An important aspect of the data is that we observe individuals purchasing with one or more of their golfing partners, but also purchasing alone. This helps us obtain posterior inference on how much of an individual's demand is attributable to

⁴ This endogeneity problem has been referred to in the social interactions literature by Manski (1993) and Moffitt (2001) as the simultaneity problem.

their own innate preferences as opposed to the social interaction effect conveyed by their partner(s). The golfing data also allows us to analyze the behavior of individuals that may be partnered with two different individuals who are themselves not partnered with one another (i.e. group members by association).

After estimating the model, we explore features of the model that can be used to guide a firm's marketing decisions. First, we decompose a customer's value into their own purchases and the purchases they generate in their partners. We find that 65% of the median customer value is attributable to the customer, while the other 35% is attributable to the customer's effect on his partners. This distribution varies across the customers with substantial mass at both 10% and 50% of the customer value attributable to effects on partners. Second, we consider targeting advertised pricing in the context of social interactions. We find that targeting at the group level increases profit by 1%, while targeting within groups can increase profitability by 20%. We also illustrate that targeting without recognizing the social interaction leads to 16% less profit. Finally, we note that targeting within groups has the ability to alienate some customers. We therefore suggest that firms may want to make offers to the group that allow them to allocate the price between members. An example might be a two-for-one price. While within group bargaining is beyond the scope of this paper, the lower the price one partner pays, the less elastic is the demand of his peers.

The remainder of the paper is structured as follows. Section 2 defines the equilibrium model and discusses its features. Section 3 describes the econometric specification and identification. Section 4 describes the empirical application. Section 5 reports model estimates, measures of customer values and pricing implications. Section 6 concludes.

2 An Empirical Equilibrium Model of Social Interactions

In this section we define a model of how individuals choose their actions, when the payoff of their actions is contingent on the decisions of others in their group. To be clear, the focus of this paper is the analysis of actions within a group, rather than the formation of the group itself. We therefore take as given a predetermined set of dyadic relationships that form a group.⁵ We define Ω to be an $N \times N$ matrix in which element Ω_{ij} equals one if individual i and j are partners and zero otherwise. The diagonals of this matrix are consequently irrelevant. All relationships are assumed to be symmetric such that $\Omega_{ij} = \Omega_{ji}$. A group that includes an individual i also includes all individuals j for which $\Omega_{ij}=1$, and all individuals k for which $\Omega_{jk}=1$ and so on. This is essential because even though i may not be partnered with k , i 's decisions are influenced by k through k 's effect on j . The N individuals observed in the data will therefore be divided into G mutually exclusive groups of varying sizes (i.e. groups are defined large enough that an individual can only be a member of a single group).

From this point forward, we index an individual by their group, g , and which individual they are within that group, i . We therefore define the preference parameters of individual i in group g to be γ_{ig} . In each time period t , individuals are modeled to make a binary purchase decision $y_{igt} \in \{0,1\}$, where $y_{igt}=1$ indicates a purchase. The decision is an outcome of a discrete choice random utility model with the following latent utilities:

$$u(y_{igt}, y_{-igt}, \varepsilon_{igt}; \gamma_{ig}, \Omega) = \begin{cases} v_0 + \varepsilon_{0,igt} & \text{if } y_{igt} = 0 \\ v_{1igt}(y_{-igt}; \gamma_{ig}, \Omega) + \varepsilon_{1,igt} & \text{if } y_{igt} = 1 \end{cases} \quad (1)$$

⁵ Firms may identify relationships through a variety of means. For example, internet social networking sites may observe whether or not two individuals have put one another in their friend list. It is important to note that group identification and relationship formation is a challenging research area of its own (see for example Narayan, 2007).

v_{igt} represents the non-stochastic portion of the indirect utility from choosing a firm's good or service. The primary feature of this model, and its deviation from the standard discrete choice model, is that the indirect utility is a function of y_{-igt} , a vector of the decisions of others in the group.⁶ Ω also affects the indirect utility because it identifies which group members directly affect i 's decision. Due to our interest in individuals that prefer to coordinate consumption decisions, we focus on models in which the following holds:

Assumption 1: $v_{igt}(y_{jgt} = 1, y_{-i, jgt}; \gamma_{ig}, \Omega) > v_{igt}(y_{jgt} = 0, y_{-i, jgt}; \gamma_{ig}, \Omega)$ if $\Omega_{ij} = 1$ and

$$v_{igt}(y_{jgt} = 1, y_{-i, jgt}; \gamma_{ig}, \Omega) = v_{igt}(y_{jgt} = 0, y_{-i, jgt}; \gamma_{ig}, \Omega) \text{ if } \Omega_{ij} = 0.$$

In words, a group member purchasing increases i 's utility of purchasing if that group member is i 's partner, but has no direct effect on i 's purchase decision if not a partner, rather just a group member by association. When $\Omega_{ji} = 0$, j does still indirectly affect i , through any other partners that i and j have in common.

v_0 is the utility of the outside good, which is normalized to 0. ε_{igt} is an extreme value distributed individual- and time-specific shock to preferences. We assume this is independent across individuals and time.

This latent utility structure applies to all individuals in the group. Furthermore, we assume complete information such that individuals in a group know others' γ 's, ε 's, and any other

⁶ In some cases, it may also be appropriate to define the outside alternative as being a function of y_{-igt} . For example, two golfers might be playing tennis together when not at the course. Without broader knowledge of group activities beyond those with the specific firm, it is impossible to identify such behavior. Furthermore, because the majority of no purchase occasions likely do not involve a joint activity of the individuals, the model as defined in the paper should provide a more realistic estimate and is perhaps the best a manager could do in trying to analyze the behavior observed in its customer purchase database. In applications where coordination on the outside alternative occurs, the later definition that a (1,1) outcome Pareto dominates a (0,0) outcome would not necessarily hold.

variables that may enter the indirect utilities. Given that individuals care about the decisions of others in the group, we model them to coordinate on a set of decisions, y_{gt} , from which none of the individuals will have an incentive to deviate. Specifically, we define y_{gt} to be a pure strategy Nash equilibrium, i.e $u(y_{igt}, y_{-igt}, \varepsilon_{igt}; \gamma_{ig}, \Omega) \geq u(y'_{igt}, y_{-igt}, \varepsilon_{igt}; \gamma_{ig}, \Omega) \forall y'_{igt}$.

2.1 Multiple Equilibria

A problem that commonly arises in discrete games, and also in the decisions of groups, is that there may be more than one decision, y_{gt} , that is sufficiently satisfactory that all individuals will go along with it. For example, if none of the individuals has a strong enough preference to unilaterally purchase, but all of them would be willing to do so if the rest purchased, none purchasing and all purchasing are both equilibria.

Multiple equilibria can present problems for both of the model's purposes: estimation and counterfactual analysis. In the case of estimation, the lack of a unique equilibrium can prevent us from defining a likelihood function. When conducting counterfactuals, the lack of a unique equilibrium prevents us from predicting outcomes when the market environment implies multiple equilibria.

Consider an interaction with two players, A and B . There are four possible outcomes of the game: $(y_A, y_B) \in \{(1,1), (1,0), (0,1), (0,0)\}$. The equilibrium conditions for each are:

$$\begin{aligned}
 &(1,1) \text{ if } (v_{11A} + \varepsilon_{1A} > \varepsilon_{0A}) \text{ and } (v_{11B} + \varepsilon_{1B} > \varepsilon_{0B}) \\
 &(1,0) \text{ if } (v_{10A} + \varepsilon_{1A} > \varepsilon_{0A}) \text{ and } (v_{11B} + \varepsilon_{1B} < \varepsilon_{0B}) \\
 &(0,1) \text{ if } (v_{11A} + \varepsilon_{1A} < \varepsilon_{0A}) \text{ and } (v_{10B} + \varepsilon_{1B} > \varepsilon_{0B}) \\
 &(0,0) \text{ if } (v_{10A} + \varepsilon_{1A} < \varepsilon_{0A}) \text{ and } (v_{10B} + \varepsilon_{1B} < \varepsilon_{0B})
 \end{aligned} \tag{2}$$

where $v_{1y_B A}$ abbreviates $v_{1A}(y_B; \gamma_A)$. Multiple equilibria arise when there is an overlap in ε_{1A} - ε_{0A} and ε_{1B} - ε_{0B} in the conditions for two different equilibria. In this two player example, outcomes (1,1) and (0,0) are both equilibria when $(-v_{10A} > \varepsilon_{1A} - \varepsilon_{0A} > -v_{11A})$ and $(-v_{10B} > \varepsilon_{1B} - \varepsilon_{0B} > -v_{11B})$. To illustrate this, we apply Bresnahan and Reiss (1991)'s graphical depiction to our model.

Figure 1 depicts the inequalities that form these equilibria over the space defined by

$\zeta_A = \varepsilon_{1A} - \varepsilon_{0A}$ on the horizontal axis and $\zeta_B = \varepsilon_{1B} - \varepsilon_{0B}$ on the vertical axis.⁷ Equilibrium (1,1) corresponds to regions *II*, *III*, *V*, and *VI*. Equilibria (1,0) and (0,1) correspond to regions *IX* and *I* respectively. Equilibrium (0,0) corresponds to regions *IV*, *V*, *VII*, and *VIII*. Notice that region *V* arises in both (1,1) and (0,0). This implies that under the specific set of ε s defined by region *V*, there are multiple equilibria.

Multiple equilibria are problematic in estimation because the ε 's must integrate to one, yet we would double count the ε 's in region *V* if we merely used the equilibrium conditions to define the probability of each outcome. Researchers have often circumvented this problem by assuming that individuals will always select the equilibrium that maximizes the total surplus of the agents (e.g. Berry, 1992). This assumption is even more plausible in the social interaction game we have defined because the equilibrium that maximizes total surplus is also the equilibrium that maximizes individual surplus. Specifically, equilibrium (1,1) Pareto dominates equilibrium (0,0), because individuals in a group either prefer to play with their partner or are indifferent, i.e.

⁷ This graphical approach follows that introduced by Bresnahan and Reiss (1991) and has been used in the entry literature to illustrate equilibria and the potential for multiple equilibria. In our case, complementarities between consumers imply that the region of multiple equilibria involves neither or both consuming, whereas in the case of competition between potential entrants the region of multiple equilibria involves a single entrant.

$v_{1i} \geq v_{10i}$. Therefore, when we define our likelihood function below, we will define the Pareto dominant equilibrium to occur when the ε 's leave the group in a position of multiple equilibria.

The Pareto equilibrium selection rule will also assist us when defining counterfactuals. For example, an advertisement or price decrease could shift the v 's into a region where multiple equilibria are much more likely. This could leave substantial ambiguity in terms of potential outcomes. However, by turning to the notion of Pareto dominance, we can narrow down the potential outcomes further by recognizing that the individuals would have little incentive to settle on an equilibrium that leaves both players worse off than if they had selected the other equilibrium.

3 Empirical Model and Identification

3.1 Data Likelihood

Construction of the likelihood for empirical models of discrete games is very similar to that for typical discrete choice models, but it differs in that, rather than dividing the ε space into two mutually exclusive regions for a binary choice, the space of ε 's across all group members is divided into mutually exclusive regions based on equilibrium conditions and selection rules. Specifically, we construct the likelihood of an observed choice by integrating over the ε 's that (i) satisfy the equilibrium conditions for the observed choice and (ii) are not associated with another equilibrium that is favored by the selection rule.

For simplicity, we will define the likelihood for the two player game here, but we extend it up to four players in the empirical application below and in Appendix 1. The likelihoods for the four

possible outcomes, $(y_A, y_B) \in \{(0,0), (1,0), (0,1), (1,1)\}$, can be expressed as function of inequalities as follows:

$$\begin{aligned}
\Pr(1,1) &= \Pr(v_{11A} + \varepsilon_{1A} > \varepsilon_{0igt}) \Pr(v_{11B} + \varepsilon_{1B} > \varepsilon_{0-igt}) \\
\Pr(1,0) &= \Pr(v_{10A} + \varepsilon_{1A} > \varepsilon_{0igt}) \Pr(v_{11B} + \varepsilon_{1B} < \varepsilon_{0-igt}) \\
\Pr(0,1) &= \Pr(v_{11A} + \varepsilon_{1A} < \varepsilon_{0igt}) \Pr(v_{10B} + \varepsilon_{1B} > \varepsilon_{0-igt}) \\
\Pr(0,0) &= 1 - \Pr(1,1) - \Pr(1,0) - \Pr(0,1)
\end{aligned} \tag{3}$$

Notice that the primary difference between these probabilities and the equilibrium conditions defined above in Equation (2) is that the probability of outcome $(0,0)$ involves integration over a smaller space because this outcome is Pareto dominated in region V . The likelihood function for a given group, g , over the T_g periods that we observe it is therefore:

$$L_g(y_{Ag}, y_{Bg}; \gamma_g) = \prod_{t=1}^{T_g} \left[\sum_{A=0}^1 \sum_{B=0}^1 \left\{ \Pr(A, B; \gamma_g) 1\{y_{Agt} = A\} 1\{y_{Bgt} = B\} \right\} \right] \tag{4}$$

If agents are homogenous, there will only be two parameters to estimate in the simple model without covariates: $\{v_{11}, v_{10}\}$. As depicted in Equation (3) above, there are three separate moments to identify these two variables. This implies that exogenous variables are not necessary to identify the model parameters. However, exogenous variables can easily be included in Equation (1) to provide less parametric identification, which we do in simulations below.

We now define the parameterization we use for the indirect utilities. We define the indirect utilities of purchasing in a additively linear form:

$$v_{igt}(y_{-igt}; \gamma_{ig}, \Omega) = \gamma_{0ig} + \gamma_{1ig} 1\left\{\sum_{j \neq i} \Omega_{ij} y_{jgt} \geq 1\right\} + \alpha_{ig} \gamma_{1ig} 1\left\{\sum_{j \neq i} \Omega_{ij} y_{jgt} \geq 2\right\} + \alpha_{ig}^2 \gamma_{1ig} 1\left\{\sum_{j \neq i} \Omega_{ij} y_{jgt} \geq 3\right\} + \dots + \beta_{ig} X_{igt} \quad (5)$$

where γ_{0ig} is the individual-specific taste for purchasing the firm's product. γ_{1ig} is the individual's preference for a partner also purchasing from the firm in the same time period. $1\{\}$ above represents an indicator of the expression inside. We allow the individual to obtain additional utility increases, γ_{1ig} , when more partners join in the purchase occasion. However, we scale these utility increases by the term α which is assumed to be positive. If it is less than one, then the marginal utility of additional partners is decreasing. Note that the representation above assumes that group members are substitutable so long as a partnership between them exists, i.e. $\Omega_{ij}=1$. X_{igt} is a vector of potentially individual specific marketing mix or other variables that affect the individual's utility and β_{ig} measures the impact of these variables. The specification could be generalized for example by allowing γ_{1ig} to differ by partner (e.g. γ_{1igt}), but we use this specification because it easily scales up and illustrates a minimal amount of data and parameters required to estimate such a model. Similarly, α restricts the relation between different group structures and this could also be relaxed when appropriate.

3.2 Heterogeneity Structure

In practice, consumers may differ in their utilities both across and within groups. For example, a vast literature notes the existence of homophily, the process through which similar individuals systematically group together. If a researcher incorrectly used a homogenous version of the model above, homophily would imply a positive correlation in the ε 's across members within a

group. With no other way to manifest itself, this correlation would lead to a positive bias in

$$v_{igt} (y_{jgt} = 1, y_{-i,jgt}; \gamma_{ig}, \Omega) - v_{igt} (y_{jgt} = 0, y_{-i,jgt}; \gamma_{ig}, \Omega).$$

We account for homophily by allowing customers' preferences to include unobserved heterogeneity that is potentially correlated between the individuals. We define the group specific parameter vector $\theta_g = \{\gamma_{01g}, \dots, \gamma_{0Ng}, \ln \gamma_{11g}, \dots, \ln \gamma_{1Ng}, \beta_{1g}, \dots, \beta_{Ng}\}$, where $\theta_g \sim N(\theta, \Sigma)$. The preference for purchasing with a partner, γ_{1ig} , is modeled to be log-normally distributed to impose the restriction that individuals in a group prefer to have a partner also purchasing.⁸ θ is the population mean vector of the parameters that includes the feature that all individuals within group have the same expectation of the parameters, i.e. $\theta = \{\gamma_0, \dots, \gamma_0, \tilde{\gamma}_1, \dots, \tilde{\gamma}_1, \beta, \dots, \beta\}$. Note that we define the mean of the underlying normally distributed parameter for the Partner Effect to be $\tilde{\gamma}_1$. The covariance matrix for the parameters is most easily expressed decomposed into a matrix of standard deviations and correlations $\Sigma = SRS$, where

$$S = \text{diag}\left(\{\sigma_{\gamma_0}, \dots, \sigma_{\gamma_0}, \sigma_{\gamma_1}, \dots, \sigma_{\gamma_1}, \sigma_{\beta}, \dots, \sigma_{\beta}\}\right) \text{ and}$$

$$R = \begin{bmatrix} R_{\gamma_0} & R_{\gamma_0\gamma_1} & R_{\gamma_0\beta} \\ R_{\gamma_0\gamma_1} & R_{\gamma_1} & R_{\gamma_1\beta} \\ R_{\gamma_0\beta} & R_{\gamma_1\beta} & R_{\beta} \end{bmatrix} = \begin{bmatrix} \overline{1} & \overline{\rho_0} & \cdot & \overline{\rho_0} & \overline{\rho_{01}} & \overline{\tilde{\rho}_{01}} & \cdot & \overline{\tilde{\rho}_{01}} & \overline{\rho_{0\beta}} & \overline{\tilde{\rho}_{0\beta}} & \cdot & \overline{\tilde{\rho}_{0\beta}} \\ \overline{\rho_0} & 1 & \cdot & \cdot & \overline{\tilde{\rho}_{01}} & \overline{\rho_{01}} & \cdot & \cdot & \overline{\tilde{\rho}_{0\beta}} & \overline{\rho_{0\beta}} & \cdot & \cdot \\ \cdot & \cdot & 1 & \overline{\rho_0} & \cdot & \cdot & \overline{\rho_{01}} & \overline{\tilde{\rho}_{01}} & \cdot & \cdot & \overline{\rho_{0\beta}} & \overline{\tilde{\rho}_{0\beta}} \\ \overline{\rho_0} & \cdot & \overline{\rho_0} & \overline{1} & \overline{\tilde{\rho}_{01}} & \cdot & \overline{\tilde{\rho}_{01}} & \overline{\rho_{01}} & \overline{\tilde{\rho}_{0\beta}} & \cdot & \overline{\tilde{\rho}_{0\beta}} & \overline{\rho_{0\beta}} \\ \overline{\rho_{01}} & \overline{\tilde{\rho}_{01}} & \cdot & \overline{\tilde{\rho}_{01}} & \overline{1} & \overline{\rho_1} & \cdot & \overline{\rho_1} & \overline{\rho_{1\beta}} & \overline{\tilde{\rho}_{1\beta}} & \cdot & \overline{\tilde{\rho}_{1\beta}} \\ \overline{\tilde{\rho}_{01}} & \overline{\rho_{01}} & \cdot & \cdot & \overline{\rho_1} & 1 & \cdot & \cdot & \overline{\tilde{\rho}_{1\beta}} & \overline{\rho_{1\beta}} & \cdot & \cdot \\ \cdot & \cdot & \overline{\rho_{01}} & \overline{\tilde{\rho}_{01}} & \cdot & \cdot & 1 & \overline{\rho_1} & \cdot & \cdot & \overline{\rho_{1\beta}} & \overline{\tilde{\rho}_{1\beta}} \\ \overline{\tilde{\rho}_{01}} & \cdot & \overline{\tilde{\rho}_{01}} & \overline{\rho_{01}} & \overline{\rho_1} & \cdot & \overline{\rho_1} & \overline{1} & \overline{\tilde{\rho}_{1\beta}} & \cdot & \overline{\tilde{\rho}_{1\beta}} & \overline{\rho_{1\beta}} \\ \overline{\rho_{0\beta}} & \overline{\tilde{\rho}_{0\beta}} & \cdot & \overline{\tilde{\rho}_{0\beta}} & \overline{\rho_{1\beta}} & \overline{\tilde{\rho}_{1\beta}} & \cdot & \overline{\tilde{\rho}_{1\beta}} & \overline{1} & \overline{\rho_{\beta}} & \cdot & \overline{\rho_{\beta}} \\ \overline{\tilde{\rho}_{0\beta}} & \overline{\rho_{0\beta}} & \cdot & \cdot & \overline{\tilde{\rho}_{1\beta}} & \overline{\rho_{1\beta}} & \cdot & \cdot & \overline{\rho_{\beta}} & 1 & \cdot & \cdot \\ \cdot & \cdot & \overline{\rho_{0\beta}} & \overline{\tilde{\rho}_{0\beta}} & \cdot & \cdot & \overline{\rho_{1\beta}} & \overline{\tilde{\rho}_{1\beta}} & \cdot & \cdot & 1 & \overline{\rho_{\beta}} \\ \overline{\tilde{\rho}_{0\beta}} & \cdot & \overline{\tilde{\rho}_{0\beta}} & \overline{\rho_{0\beta}} & \overline{\tilde{\rho}_{1\beta}} & \cdot & \overline{\tilde{\rho}_{1\beta}} & \overline{\rho_{1\beta}} & \overline{\rho_{\beta}} & \cdot & \overline{\rho_{\beta}} & \overline{1} \end{bmatrix} \quad (6)$$

⁸ Relaxation of this restriction could lead to a lack of an equilibrium that would unnecessarily complicate our analysis. See Bresnahan and Reiss (1991) for more on this.

where a ‘ \sim ’ over the correlation coefficient means that it is a cross-partner-cross-parameter correlation such as $\tilde{\rho}_{01}$ which is the relationship between A’s intercept and B’s partner effect as opposed to ρ_{01} which describes the relationship between A’s intercept and A’s partner effect, or ρ_0 which would describe the relationship between A’s intercept and B’s intercept. While it is possible to estimate this entire matrix, we focus only on the upper left set of correlations, R_{γ_0} , to capture the correlations between individuals inherent tastes for the good or service. This set of correlations is that most likely affected by homophily.

We estimate this model and heterogeneity structure using a hierarchical Bayesian model. It is useful to point out that a simulated maximum likelihood approach could also be used, but would provide less direct inference of the individual-level parameters. The heterogeneity distribution above forms the priors over individual-level parameters, which are drawn using a Metropolis-Hastings algorithm due to the extreme value errors. The population mean parameters are defined to have normal or log-normal priors. The standard deviations of the population-level γ -s are defined to be log-normal, while the variances of the β -s are inverse-Gamma. The distributions and steps of the MCMC process are more thoroughly laid out in Appendix 2.

3.3 Identification and Parameter Recovery in Model Simulations

3.3.1 Identification of Homogenous Model

We now return to identification. As noted at the end of section 3.1 above, the homogeneous model with two players is identified based on the fact that there are two unknown parameters and three separate probabilities observed in Equation (3) to identify these parameters. A three player version implies seven separate probabilities to identify the parameters, while only adding the α parameter. As specified above, four or more players add additional observed probabilities but

imply no additional parameters. Given these additional moments, one could clearly make the specification for three or more players even more complex. It is important to note that we have defined the likelihood assuming the ε 's are independent. If there is plausible concern of correlation among the unobservables of the players, identification would absorb an additional moment. Furthermore, the exclusion restrictions would become very useful in identification. Without obvious exclusion restrictions, the researcher must either find other restrictions to facilitate identification or conduct robustness checks to verify that correlated unobservables are not a concern. In the estimation below, we test the robustness of the estimates to the inclusion and exclusion of an obvious correlated unobservable and find that the social interaction effect remains stable.

Table 1 reports Monte Carlo simulations and parameter recovery. In all models reported in the table, the likelihood for a group's decision is defined based on Equation (3) and Appendix 1. This equilibrium-based derivation of the likelihood accounts for the endogeneity of partners' decisions. We illustrate its value in resolving biases below.

The left panel of Table 1 reports Monte Carlo simulations and parameter recovery for the homogenous model. We first only estimate the model with pairs and see that all parameters are recovered. Then we estimate the model with pairs, threesomes and foursomes and recover parameters. Note that as described previously, all groups are non-overlapping.

3.3.2 Identification of Heterogeneous Model

Intuitively, the heterogeneity structure is identified based on the panel structure of the data in which we repeatedly observe each of the groups. Under the Bayesian estimation approach, as we observe the group more times, the estimates of group-level parameters will shrink around the true group-level values within the population distribution of the parameters. Keep in mind that

in referring to group level parameters, there are potentially separate parameters for each individual within the group (e.g. A, B, C etc.).

It is first useful to consider which parameters could be identified with a sufficiently long time series for the group. If there are no covariates and only two players, then based on Equation (3), we can identify three parameters. Lets call these the intercept for A, the intercept for B, and a Partner Effect that is assumed to be the same for both A and B. By observing a time series for multiple groups, we can also estimate a variance for each of these parameters in the population. We can also identify correlations between the parameters, but lets hold off on that for now. One useful thing to point out now is that three intercepts and three variances are actually too many parameters, because the mean and variance of A's and B's parameters are the same, since definition as A or B is purely random. Essentially, we only need the mean and variance of both the intercept and Partner Effect parameters. The discussion above, illustrates that we have that.

It is now useful to return to two interrelated points in the preceding paragraph: i) we ignored the discussion of the correlations between the parameters and ii) the Partner Effect was assumed identical for A and B. The correlation between the intercepts is identified from the fact that we could get a group specific value for both A's and B's intercepts. If these both do not appear to be independent draws from the distribution defined by the mean and variance parameters of the intercept, then there is an identifiable correlation. When dealing with the Partner Effect, we assumed it was perfectly correlated (i.e. identical) within the group. This was because we cannot identify group specific values for A's and B' Partner Effect. Following the discussion we just had for the intercepts, this implies that we cannot identify the correlation between the Partner Effects within a group. However, the correlation need not be assumed to be 1. We could similarly assume the correlation is zero. In which case, the Partner Effect is allowed to differ

within the group.⁹ To summarize, for the two player case without covariates, we can identify both the mean and variance of the intercept and the Partner Effect and the correlation between A's and B's intercepts, but we do need to assume a correlation for A's and B's Partner Effects.¹⁰ Adding covariates to this specification would create additional moments and allow the full heterogeneity structure to be estimated. Ideally, these covariates would be individual specific to create exclusion restrictions, but a single common dummy variable would double the number of moments to 6, such that two intercepts, two Partner Effects and two coefficients on the dummy could be estimated, so long as we did not add an additional parameter, such as allowing the dummy to have a different effect when playing with a partner.¹¹

The upper right panel of Table 1 simulates data for this heterogeneous model of pairs of individuals and recovers parameters. Notice that means and standard deviations on both parameters are recovered, and that a correlation parameter for γ_{0Ag} and γ_{0Bg} is recovered. The bottom right portion of Table 1 illustrates parameter recovery of the full model with pairs, threesomes and foursomes. Identification is straightforward in these cases because, as described above, all parameters would be identified with a long enough time series at the group level, so cross-group comparisons allow a full covariance matrix of heterogeneity to be estimated.

⁹ In some simulations, we found that assuming a zero correlation actually fits better than assuming it is one. This likely arises from the case of not observing one of the four potential outcomes for a group. Suppose (1,0) is never observed for a group. One way to explain this is for B to have an infinitely positive utility for playing with A. However, if the Partner Effects are assumed identical, then A must also have an infinitely positive utility for playing with B. But, that would suggest we should not have observed (0,1) either. By allowing the Partner Effects to be different with a fixed correlation of zero, some (0,1) observations occurring together with zero (1,0) observations is accommodated.

¹⁰ In terms of the heterogeneity distribution, we can identify $\gamma_0, \tilde{\gamma}_1, \sigma_{\gamma_0}, \sigma_{\gamma_1}$, and ρ_0 . $\rho_{01} = \tilde{\rho}_{01}$ can also be identified based on whether intercept and Partner Effect parameters that could be estimated at the group level have a systematic within group relationship.

¹¹ Essentially we can estimate the effect of X on (1,1), (1,0) and (0,1) separately for each value of X. Yet, under the model structure, we only need to assume that the effect of X on (1,1) is derivable from the Partner Effect and the effect of X on (1,0) and (0,1).

3.3.3 Bias When Ignoring Endogeneity of Partner's Choice

As motivated above, an advantage of the equilibrium model is that it helps resolve an endogeneity bias resulting from the partner's decision not being exogenous. While instruments (i.e. exclusion restrictions in this case) are always preferable, researchers might not have access to them, such that a parametric approach like the equilibrium model becomes the best alternative. To illustrate the value of the equilibrium model in resolving this bias, we re-estimate the scenario in the upper-right corner of Table 1. Recall that the data for the pairs are generated with the equilibrium model and the heterogeneous individuals. We now estimate that data with the heterogeneity specification defined above, but falsely assuming that the partner's decision is exogenous. Table 2a reports these estimates. We see that by ignoring the endogeneity of a partner's decision, the social interaction (Partner Effect) is estimated to be between 1.17 and 1.23, when it is really only 0.5.

3.3.4 Bias When Ignoring Heterogeneity

One motivation for including unobserved heterogeneity in the model is the bias that is introduced when groups are endogenously formed to include similar types of individuals. To illustrate this bias and the role of the heterogeneity distribution in removing it, we compare the estimates of the heterogeneous model in Table 1, with estimates of a homogenous model applied to the simulated heterogeneous data. We report this in Table 2b. While in the right side of Table 1, we illustrated that we were able to recover the social interaction effect (i.e. Partner Effect) in both sets of estimates with heterogeneity, we are unable to recover this effect when ignoring heterogeneity in Table 2b. We see that the Partner Effect is estimated to be greater than 0.57 instead of being equal to 0.5. This positive bias in the estimation of the social interaction effect results from the fact that the underlying correlation in the heterogeneous tastes is positive (i.e. homophily). Specifically, the heterogeneity in the model captures the time invariant unobserved tastes that are

similar within groups. Without the heterogeneity, the model can only capture this by inferring the social interaction to be larger than it really is.

3.4 Comparison with Other Approaches to Identifying Social Interactions

Here we briefly describe some related models that can put the contribution of our model and the identification approach in a context. For a more thorough coverage of the estimation and identification of social interactions, see Hartmann et.al. (2008). Most of the existing literature studying social interactions has estimated linear models with a continuous outcome variable and a continuous measure aggregating the decisions of others as the social interaction variable of interest. Identification concerns in these models are discussed at length in Manski (1993) and Moffitt (2001). Nair et.al. (2007) is a marketing application that provides a thorough consideration of each of these issues in a marketing application analyzing how opinion leading physicians influence the prescription behavior of other physicians. Yang et.al. (2006) also uses a linear model to estimate the interrelatedness of spouses' television viewing behavior.

The equilibrium approach we use falls within the scope of empirical models of discrete games as defined by Bresnahan and Reiss (1991). Most applications of this approach have been in the context of entry games in which the actions of one agent decrease the probability the other agent will take the same action. One primary contribution of our model relative to this literature is that we have included extensive unobserved customer heterogeneity that is allowed to be correlated among the agents. The necessity of this in our model arises from the fact that unobserved correlations in agents' payoffs will bias the interaction effect upward, potentially leading to an estimated effect when one does not exist. Moffitt (2001) notes the importance of this consideration as resulting from a problem of endogenous group formation. The same problem

typically does not exist in entry games because the players in the game are generally assumed to be all potential entrants (i.e. there is one single group observed repeatedly across all markets and/or time). Furthermore, if the problem did exist in entry games, it would bias the interaction effect to be positive and thus only understate the degree of competition between firms, rather than falsely identify competitive effects that do not exist. There are a few models of equilibrium firm behavior that are more closely related to the present paper. Both Ellickson and Misra (2007) and Vitorino (2007) explicitly allow for complementarities between the decisions of firms. Models of direct and indirect network effects are also related in that they allow for complementarities (e.g. Akerberg and Gowrisankaran, 2006, Tucker (2006), Ryan and Tucker (2007) and Dube, Hitsch and Chintagunta (2008).

Specific to social interactions, Brock and Durlauf (2003 and 2001) consider structural estimation of discrete choice models of interactions between large numbers of agents. These are models of incomplete information in which individuals are uncertain about the decision that others will make. Therefore, the individual's choice is affected by his or her belief about others. They specifically acknowledge that this assumption is "clearly problematic in describing interactions between a pair of best friends," where it is likely that they know each others preferences, unobservables and decisions. Incomplete information models such as these involve a number of empirical challenges. The specification of expectations in the model has typically involved a fixed-point solution that settles on all agents' expectations aligning with one of multiple potential equilibria. While recent empirical developments can more easily consider all potential equilibria (e.g., Su and Judd, 2008), the models still have little to say about how all agents manage to coordinate their expectations on only one of the possible equilibria.

One additional limitation of incomplete information models of social interactions is that they typically involve a two-step estimation approach (e.g., Bajari et.al., 2005) that creates difficulties in the presence of unobserved heterogeneity. This implies that the bias described above arising from endogenous group formation could be present and could identify a social interaction when one did not actually exist. Recent work has however been directed at resolving this specific problem. Nevertheless, the choice between complete and incomplete information should primarily depend on the application and feasibility of the modeling.

One relevant paper considering discrete games of complete information is Bajari, Hong and Ryan (2004). Its focus is on developing an alternative approach for dealing with multiple equilibria. Specifically, they show that an equilibrium selection equation can be identified when there is a variable that affects the equilibrium selected, but not the payoffs of the agents. Their empirical application is an entry game, but they consider that the model could be extended to social interactions. However, like other papers in the literature they also do not allow for persistent unobserved heterogeneity. Finally, Soetevent and Kooreman (2007) estimate a discrete choice model of social interactions applied to high school teen behavior, such as smoking. Their equilibrium model closely resembles our model, though they do not recognize that these models follow directly from the pioneering work of Bresnahan and Reiss (1991). Their contribution is therefore in defining a simulation based estimator that can be applied to large groups, however we have solved our model analytically for up to a group size of four, which is most relevant for our empirical application. They also do not allow for unobserved heterogeneity.

One final problem which potentially arises in social interaction models is correlated unobservables. Our treatment of endogenous group formation above, controls for time invariant

unobservables that are correlated within the group. However, it is possible that if time-varying unobservables are correlated, one might infer a social interaction when one does not exist. For example, Van den Bulte and Lillien (2001) find that previously unobserved marketing efforts explain what previous researchers had inferred to be contagion in the diffusion of the drug tetracycline.

We deal with correlated unobservables through the richness of our scanner panel data rather than through a modeling feature. The idea is that we observe the exact time and dates of purchases, such that unobservables would have to be correlated at the minute level to falsely infer a social interaction between group members. This might still be a problem in some applications where there is a potentially large peak of demand at any given minute in time. Consider the rush of students at a lunch line at 11:45am. This certainly does not reflect that they all want to eat together, rather it reflects the fact that their classes end at 11:30am. Our application of golf described below avoids this problem because no more than four people can purchase at any given time, so if two members of a group are observed purchasing at the same time, it is pretty safe to assume they actually wanted to purchase together.

While this approach for dealing with correlated unobservables avoids inference of an interactions that does not really exist, we have not fully accounted for correlated unobservables because we do not have an exclusion restriction. We therefore check the robustness of the results below by including an obvious correlated unobservable, such as weather in outdoor activities.

4 Empirical Application

To validate our model, explore how model parameters are derived from data, and evaluate the marketing implications of social interactions, we apply our model to a sample of groups of golfers.¹² Golf is certainly an activity where individuals obtain more utility when playing with a friend or colleague, as opposed to alone or with an unknown partner. Our data consist of individuals whose purchase decisions at a southern California golf course were revealed by swiping a membership card that recorded their identity and exact time of purchase. We also observe weekly variation in whether an ad was run in a local newspaper and what price was reported in that ad.¹³ We will include these variables to evaluate the role of social interactions in the effectiveness of marketing mix activities.

4.1 Groups

As stated above, our model and estimation approach is designed for groups that are known ex-ante.¹⁴ Our treatment of groups is through the matrix of binary indicators of whether a dyadic relationship exists. In practice, the golf course could form this matrix based on which individuals booked their reservation together and a social networking website could base this on

¹² The empirical example of golf has also been analyzed in a dynamic decision context by Hartmann (2006). Including forward-looking behavior in the present model is intractable. We therefore ignore the dynamic aspects of the data to explore identification of the model presented in this paper.

¹³ We do not know the determinants or reasoning behind these marketing mix variables, so it is possible that they may have advertised when demand at the course was anticipated to be systematically low. We do not however have instruments to account for this so it is possible that the effectiveness of these marketing mix variables will be biased downward. That said, the focus of this paper is to illustrate the social interactions effects of our model rather than solve an endogeneity bias that might have arisen in the marketing mix decisions of the course. When our model is applied by managers, they are likely to be more familiar with potential instruments that affect their actual pricing or can use experimentation to create the necessary exogenous variation in marketing mix variables.

¹⁴ The identification of groups and the process through which groups are formed are interesting research questions with substantial technological challenges that are currently being explored by others (e.g. Narayan, 2007). Once the technological ability to identify group formation processes has evolved further, we expect researchers will merge our modeling with this literature.

friend lists. We do not have the reservation log at the course, so from our perspective, Ω is considered an unobserved matrix of states. Because our estimation will not integrate over all possible values of that matrix, Ω , (the dimension is too large), we assume that we know the matrix. To define the matrix, we therefore use a set of values that are the probably the most likely posterior values of that state matrix that could be inferred from the precise purchase timings in the data.

The exact process by which we uncover groups in the data is as follows. If two individuals are partners and if they arrive on the same day, they will purchase at almost the same time. In fact, this golf course, like most others, calls upon groups to “check-in” together when the group is two or three groups behind the group on the tee. From this, we can infer the values Ω as follows. If two individuals are observed to play on the same day but purchase more than an hour apart, these individuals clearly are not close golfing partners, so we assume their elements in Ω equal zero. On the other hand, if we see them purchase at exactly the same time, or within minutes they may be partners. Yet, before setting their elements of Ω to one, we recognize that they could have been randomly paired together at the course. We therefore require that the individuals are observed to purchase within minutes of each other on at least two occasions before we declare them partners. This is an admittedly imperfect approach, but serves as a group structure within which we can explore the model defined above.

Of the 3,151 individuals we observe golfing at the course, this grouping process yielded 199 mutually exclusive groups¹⁵ ranging in size from 2 to 6. Only six of these groups were larger than the maximum capacity of four that the course could accommodate at a time, so we decided

¹⁵ The implication is that an individual can only be a member of a single group. Such groups could be of very large size, however, the maximum capacity of four at a course probably prevents

to drop these groups and ignore the issue that capacity constraints could create a crowding phenomenon. While this is an interesting dynamic it is just not prevalent enough in our data to explore in more detail.¹⁶ We were therefore left with 193 groups consisting of 152 pairs, 34 threesomes, and 7 foursomes. This number of foursomes may appear small given that a typical time slot has a capacity of four. This is because pairs of golfers are often randomly joined at the course to fill the course's time slot capacity and there may be some group members that do not have a card for us to observe their behavior.

Summary statistics of the purchase patterns within groups are included in Table 3. We see that within the 152 observed pairs, the individuals purchased together an average of 10 times, with a minimum of two and a maximum of 114. We observed only one of the golfers playing an average of 16.7 times, with the minimum of zero indicating that at least one pair played together but never played separately. Within threesomes, all individuals played together an average of 3 times with a minimum of zero and a maximum of 17. The minimum of zero suggests that there is at least one threesome in which a given individual purchased with either one of two partners, but never with both. Across the 7 foursomes we observe, they all played together an average of 1.85 times, with a minimum of zero and maximum of 4. Similarly at least one of the foursomes likely involves an individual that plays with different group members, but not all at once. The data consist of daily observations between January 4, 2000 and December 31, 2001. While we observe some groups over the entire time period, one group is only observed for 83 days. This is due to the timing in which the firm began recording purchases through the individuals' swipe cards.

¹⁶ In fact, the maximum capacity may be most relevant to the process through which golfers form groups.

Table 3 also provides some model-free evidence that a social interaction might exist. Specifically, the average number of purchases per individual was 27 (16.7+10.5) out of roughly 347 opportunities (i.e 7.8% of the time). If the individuals purchase decisions were independent, that would imply they would play on the same day about 2.1 times (0.6% of the time). Yet, we see that they purchased together over 10 times on average. This suggests that the joint probability of purchase is certainly more than the product of independent probabilities of purchase.

The price and advertising data, as well as rain, over this time is summarized in Table 4. We see that there was an ad in the paper for 34 percent of the 728 days observed. The advertisements were published on Tuesdays and typically included a price discount which we have incorporated into the price variable. We maintain the advertisement indicator at 1 throughout the entire week because individuals may make arrangements on Tuesday to play on any one of the days that the advertised price covers. The price varies from \$35 to \$75 with a mean of just under \$60. The variation in the price is due to the advertised price and weekend vs. weekday pricing. Due to the fact that much of the variation in the price is coming through a newspaper ad which some golfers may not see, we will expect measures of sensitivity to this price to be relatively low, reflecting awareness of the ad as well as price sensitivity. We see that there was an average of 0.04 inches of precipitation during the sample.

5 Model Estimates and Implications

To estimate the model we use a hierarchical Bayesian MCMC approach. This has two primary benefits. First, it allows us to include unobserved heterogeneity that is allowed to be correlated

between the individuals within a group. Second, it allows for inference of individual-level parameters so that we can evaluate targeted marketing strategies.

5.1 Estimates

Table 5 reports estimates for three different specifications of the model. Specification I restricts heterogeneity to the intercept and Partner Effect. The primary reason for restricting heterogeneity in this case is that this model best illustrates the targeting implications of the social interaction (see section 5.3 below). We see that the intercept is on average -3.64 across golfers. This reflects the fact that golfers play infrequently (i.e. because the utility of the outside option is normalized to zero). There is significant heterogeneity in purchase frequency reflected by the standard deviation of 0.90. The median increase in the utility of purchasing that a golfer obtains when a partner joins him is 1.92 (i.e. $\exp(0.65)$). There is also substantial heterogeneity in this parameter suggesting that some golfers obtain significantly higher and lower benefits of a partner joining them on the course. The last parameter reported describes how this social interaction utility changes when additional partners join in the round. We see that this is significantly less than one, with a value of 0.67 (i.e. $\exp(-.40)$), suggesting that the utility of additional partners joining at the course is diminishing. The weekend parameter measures the additional utility golfers obtain when purchasing on the weekend. We see that the price coefficient is significantly negative (this also holds past the 99th percentile of the distribution). The advertisement has a median effect of 0.02, but is not significant. This is likely due to the difficulty of separately identifying an advertisement and price effect when the price changes are advertised prices. When price is excluded from the analysis, the advertisement coefficient is significantly positive. Finally, we see that there is a positive correlation in the intercepts within groups of 0.45.

Specification II adds heterogeneity on the other covariates. Estimates change slightly though the pattern resembles those in Specification I. We see that there is significant heterogeneity for each of the parameters and the correlation between intercepts remain significantly positive, indicating homophily. Specification III adds the variable Rain to the analysis. Rain clearly affects individuals' abilities to play golf, but serves a broader role here in evaluating the role of correlated unobservables in the parameter estimates. Rain is certainly a correlated unobservable, but as we see from the specification, this does not seem to bias the Partner Effect estimates. It does increase the intercept, but this is because rain explains more of the occasions in which players do not golf.

5.2 Estimates of Customer Value

To quantify the importance of the social interaction effects, we measure the value of a customer to the firm over a one year horizon, then divide the fraction of these values between demand from the customer himself and demand from the customers' partners. The value is based on the profit¹⁷ stream from the customer over a year, plus the additional profit from the customer's partners that is attributable to the customer being in the group. Specifically, the latter component is calculated as the difference between the profit from the customer's group members when the individual is in the group and the profit from the customer's group members when the customer is not in the group. Figure 2 is a histogram of the customer value for the 434 golfers observed in the data.

Figure 3 is a histogram of the fraction of a customer's value that is composed of demand from others in the customer's group. We see substantial heterogeneity in this measure ranging up to 0.68, with a median of 0.35. An important implication of these customer values is that the sum

¹⁷ Profit in this case is treated as revenue as the marginal cost of a round of golf is practically zero. While this changes if the capacity constraint binds, we cannot calculate the course's opportunity cost of the round.

of all these values is likely to be greater than the value of the group. This implies that it may be almost as costly to lose one or two members of a group as it is to lose the entire group.

5.3 Targeted Advertised Prices

In this section, we consider strategies in which the firm targets price advertisements to either consumer groups or individuals within groups. For simplicity and ease of exposition, we focus on pricing only to customer pairs. We also use the estimates from Specification I in Table 5 so that we can explore targeting implications arising primarily from the social interactions, as opposed to different sensitivities to the marketing mix. Before considering the profitability of targeting we must first examine the profitability at the current pricing and the profitability at what would be an optimal uniform (e.g. blanket) advertised price.¹⁸ We find that an optimal uniform advertised price to the pairs of customers in our data would be \$95.65. Note that this price is about \$20 greater than the current weekend price to residents of \$75, but it may be reasonable for two reasons. First, the price at the course for customers that are not residents is \$98. Our analysis may therefore suggest that offering a discount to local customers may not make sense. Second, the predicted price of \$95.65 may reflect the fact that the sample of golfers we observe are more frequent golfers who are less price elastic, such that if the firm began discriminating in prices, these would be the customers it should charge its highest prices to. We find that this uniform price increase would increase profitability relative to the current \$75 price by about 2.9 percent.

¹⁸ The optimal price is calculated as the price that maximizes the revenue from the group. This includes the assumption that the capacity constraint is not binding, i.e., $MC=0$.

5.3.1 Targeting Groups of Customers

We now compare the profitability of targeting price at the group level to profitability under the uniform price of \$95.65. The average targeted price would be \$91, with a standard deviation of 9.55. The price to the group is generally increasing in the average of the golfers' intercepts, but may also be high for groups with small average intercepts but large complementarities. These prices targeted at the group level increase profitability relative to the uniform price by 1% (the profit increase relative to the current price would be 3.9%).

5.3.2 Targeting Individuals within Groups

While targeting prices at the group level increases profitability, there may be additional gains from targeting advertised prices within the group. When we target within the group, we find that average price is about \$100, with a standard deviation of 84. The large standard deviation arises because a large social interaction in the group can give the course an incentive to drastically lower the price to one customer, while exploiting the inelastic demand of his partner. In Figure 4a, we depict the relationship between the maximum social interaction effect (on the horizontal axis) and the lowest price charged to a member of the group (on the vertical axis). We see that when the social interaction effect is large, the price to one of the group members goes down to zero. In Figure 4b, we change the vertical axis to depict the highest price charged to a member of the group. Here we see that at really large social interaction effects, the firm would ideally like to charge more than \$400 to exploit the inelasticity generated by offering his partner a free round. This within group discriminatory pricing predicts profitability that is 20% greater than the blanket offering and 19% greater than only targeting at the group level.

While this seems like an attractive possibility, we must balance this against the model's inability to incorporate the fact that the high paying customer might reasonably feel cheated. Customer

backlash is obviously a concern with any targeted pricing policy, but it should be even greater within social groups where individuals are likely to be aware of the preferential treatment their friends are receiving. A better alternative likely involves offering a two-for-one coupon that allows the individuals to split the price accordingly. In fact a course located about 20 miles north of this course regularly offers such a coupon in the same newspaper. While exploring the within group bargaining that is involved in such an offer is beyond the scope of this paper, we can infer from the analysis above that if one customer is willing to incur a larger price than their partner, a package price predicts an average optimal price per customer that is about \$9 greater than if the firm were forced to charge both customers the same price. One other strategy to mitigate the backlash would be to target groups with strong complementarities by randomizing or alternating which customer receives the zero price. If the firm used this alternating zero price strategy for the 28 pairs for which it is more profitable than group-level targeting, while retaining group-level targeting for the other 124 pairs, it would increase its profit by 7% relative to pure group-level targeting.¹⁹ It should be noted that the effectiveness of this strategy relies on customers ignoring intertemporal budget constraints. Finally, the firm might offer a discounted annual membership to an individual that is in a group with people with strong social interaction effects, because that would lower that customer's marginal price of attending to zero, thereby decreasing the elasticity of the customers' partners throughout the year.

Another alternative to the highly discriminatory prices within groups would be to focus targeting on non-price items which lead to less backlash. For instance, if we had estimated a significant advertising effect and knew the costs of advertising to a given individual, we could consider

¹⁹ While one concern of the alternating zero price strategy is that the price to the paying customer might still be quite large, we find that for 12 of these 28 groups, the price to the paying customer would be less than the twice the group-level targeted price.

strategies which just increased the email or mail advertisements disproportionately to one of the customers without offering an explicit price decrease.

5.3.3 Targeting when Ignoring Groups

To evaluate the importance of considering social interactions when targeting, we now evaluate targeting these customers if the firm had obtained individual-level parameter estimates from a model that ignored the inherent relationships between its customers. Population-level estimates of a model without the social interaction are reported in Table 6 below. Most coefficients are of similar magnitudes to those reported in Table 5, except for the intercept which is larger because it is picking up preferences for purchasing both alone and with partners.

Before estimating the profitability of targeting based on this model it is important to point out that this model will by nature target within groups of customers. Therefore, it is subject to all of the potential backlashes as the targeting within groups considered above. This fact alone, may indeed be enough to avoid targeting whenever we expect customers to purchase in groups, unless we know the exact group structure. First, we compare the profitability of this approach with the existing price of \$75. It predicts an average price of \$96.41, with a profitability improvement of 3.9%. It therefore appears almost as profitable as pure group-level targeting, except that the price differences within groups implied by this strategy could lead to a backlash. When we compare this targeting based on a model without social interactions to the targeting within groups based on our model with social interactions, we find that it is 16% less profitable. We therefore conclude that if a firm targets within groups, it should consider group structure to avoid this loss in profitability.

5.4 Potential Future Applications and Model Extensions

The introduction mentions a few other potential applications of this modeling but we elaborate on these in more detail here. Visits to movie theaters would be a great application and some theaters, through their introduction of loyalty programs, now have data that is quite similar to that analyzed here. Restaurants would be an ideal application, but it may be difficult to acquire the individual-level data to estimate the model defined above. However, the substantive implication that a restaurant might want to offer two-for-ones or other marketing programs that lower the marginal price to one individual to attract others does carry over. Another obvious application would be online gaming. For example, Xbox-Live has histories of its customers similar to those observed here. Furthermore, gamers are observed to both systematically play against common friends or foes, while also playing against opponents that are practically selected at random (at least within a skill-class). Finally, the growth of social networking websites provides a rich area for future applications of this type of modeling. While work until now has focused on understanding group structure (i.e. Ω), relevance to marketers requires analysis of actual behavior that can generate revenue. The model here is designed for purchase data, but other choice data that generates advertising revenue (such as the desire to view the same sites or YouTube videos as friends) could also be accommodated through extensions of this model.

From a modeling perspective, extensions could involve more complex choices, dynamics or the integration of group formation. The movie theater situation would be one that would be assisted by moving from modeling the choice of whether or not to see a movie to which movie to see. Dynamics could also be relevant if past choices affect current decisions. While this is clearly a valuable extension, forward-looking dynamic equilibrium models are currently unable

to accommodate the extent of unobserved customer heterogeneity that is likely to arise in customer purchase databases.

Other modeling extensions may further explore asymmetries in the complementarities within groups. For example, does leader or follower behavior arise and what are its implications. Also, future research could explore within group bargaining that may arise if a price is offered at the group level, yet participants differ in their purchase intentions (e.g. Fehr and Schmidt, 1999). Along these lines, is fairness of different prices within groups really a concern, and if so, what steps can be taken to mitigate it.

Another potential extension could involve the incorporation of attribute level influence as in Arora and Allenby (1999). In their paper, individuals exogenously make a joint decision or individual decision and the importance weight of the attributes are affected by whether or not the decision is joint. In the model specified in this paper, whether or not both parties coordinate on a single decision is an endogenous outcome, but the effects of attributes such as price or advertising are assumed to remain unchanged. Relaxing this latter assumption by interacting y_i with attributes could, in the case of movies mentioned above, accommodate the fact that tastes for some movie characteristics such as romance are more desirable by both parties when consuming as a group. These interactions would certainly affect identification, and depending on the dimension of the attribute space, could make analytical solution of the equilibrium and likelihood very complicated.

Finally, an important step in the development of this literature will involve the combination of demand modeling for the group and group formation processes. One example where this would be particularly useful is the calculation of customer lifetime values. In the present case, we only calculate values over a year, because it is quite possible that group

structure could change after a year. A model of group formation could allow the firm to understand how likely it is that a lost customer will be replaced within his group.

While this is a valuable extension it has two related limitations. First, up to date group structures are rarely observed. Even in the case of social networking sites that observe links between individuals, it is unlikely that these are regularly updated if the relationship between two individuals dissolves. Furthermore, many of the links observed in the data may have existed well before the website observes them ratifying their relationship online (e.g. two high school friends may have formed a link recently in the database despite a relationship that has existed for many years). Second, because true group structure is likely to be unobserved, researchers may need to integrate over potential relationships, but as we saw with Ω in the present model, this is a very high dimensional integration which is currently too computationally intensive.

6 Conclusion

This paper has sought to accomplish three primary goals. First, it develops a model that can be used to measure demand and social interactions within groups of customers. Second, it estimates the model using a Bayesian approach that allows us to uncover substantial heterogeneity and to obtain individual-level parameter estimates that could be used for counterfactuals. Finally, the paper computes customer values and is able to account for how much of that value derives from the people the focal customer influences, either directly or indirectly, in the group. These social interactions were also shown to magnify the effect of targeted marketing activity, such as a price decrease, by increasing the demand of a customer's fellow group members. Group-level targeted pricing policies increase profitability and within group targeting can provide an additional 19% increase in profitability. For groups with strong social interactions, we find that it may be profitable to set one customer's price to zero, while raising the price to his peers whose demand

will be very inelastic due to their desire to purchase with him. While generally impractical, a real-world example of this might be a free tuxedo offer to a groom, whose groomsmen's price elasticity will be very low. In most practical examples, such pricing will however create a backlash. We therefore suggest a variety of mechanisms such as two-for-one coupons or free offers that alternate through the group members. The latter policy can increase profitability relative to group-level targeting, without treating one member less favorably.

The golf application in this paper is just one of many examples where firms face demand from groups of customers that coordinate their decisions. Similar issues arise in industries such as restaurants, online gaming, and movies. In fact, Weinberg (2005) notes the importance of understanding joint decisions for analyzing movie demand. He suggests analyses that could be enabled by extending our model to a context in which the group selects between a set of available movies. An important additional component of the model he notes would be whether or not any of the group members had previously seen the movie. This feature is absent from golf in which customers are willing to repeatedly play the same course. However, the recent introduction of customer loyalty programs in movie theaters may soon generate the individual-level data to analyze this model extension and answer some of Weinberg's interesting questions about social coordination in movie choice.

References

- Ackerberg, D.A. and G. Gowrisankaran (2006), "Quantifying Equilibrium Network Externalities in the ACH Banking Industry," *Rand Journal of Economics*, 37:738-61.
- Arora, N. and G. Allenby (1999), "Measuring the Influence of Individual Preference Structures in Group Decision Making," *Journal of Marketing Research*, XXXVI: 476-487.
- Bajari, P., H. Hong and S. Ryan (2004), "Identification and Estimation of Discrete Games of Complete Information," working paper.
- Bell, D. and S. Song (2004). "Social Contagion and Trial on the Internet: Evidence from Online Grocery Retailing," working paper.
- Berry, S. (1992). "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), 889-917.
- Bresnahan, T. and P. Reiss (1991). "Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57-81.
- Brock, W.A. and S.N. Durlauff (2001), "Discrete Choice with Social Interactions," *Review of Economic Studies*, 68, 235-260.
- Brock, W.A. and S.N. Durlauff (2003), "Multinomial Choice with Social Interactions," working paper.
- Dube, J.P., G. Hitsch and P. Chintagunta (2008), "Tipping and Concentration in Markets with Indirect Network Effects," working paper.
- Ellickson, P. and S. Misra (2007), "Supermarket Pricing Strategies," forthcoming at *Marketing Science*.
- Fehr, E. and K. Schmidt (1999), "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114(3), 817-868.
- Godes et.al. (2005). "The Firm's Management of Social Interactions," *Marketing Letters*, 16(3-4), 415-428.
- Hartmann, W.R. (2006). "Intertemporal Effects of Consumption and Their Implications for Demand Elasticity Estimates," *Quantitative Marketing and Economics*, 4(4).
- Hartmann, W.R. et.al. (2008). "Modeling Social Interactions: Identification, Empirical Methods, and Policy Implications," forthcoming at *Marketing Letters*.
- Heckman, J. (1979). "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153-162.

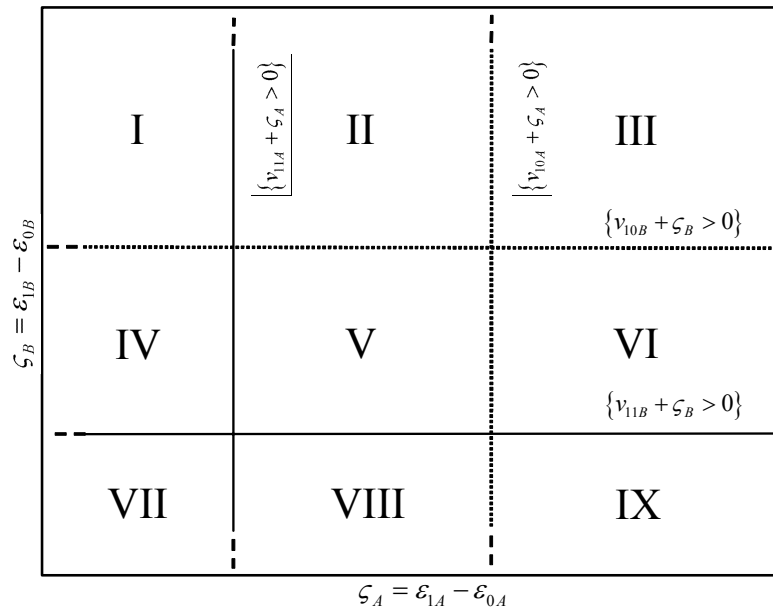
- Kumar, V. J.A. Petersen, R.P. Leone (2007), "How Valuable is Word of Mouth," *Harvard Business Review*.
- Manchanda, P., Y. Xie, and N. Youn (2004), "The Role of Targeted Communication and Contagion in New Product Adoption," working paper.
- Manski, C.F. (1993). "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531-542.
- Moffitt, R.A. (2001). "Policy Interventions, Low-Level Equilibria, and Social Interactions," in *Social Dynamics*, edited by Steven N. Durlauf and H. Peyton Young, Brookings Institution Press, Washington, D.C., 45-82.
- Nair, H., P. Manchanda, T. Bhatia (2006), "Asymmetric Peer Effects in Physician Prescription Behavior: The Role Of Opinion Leaders," working paper.
- Nam, S., P. Manchanda, and P. Chintagunta (2006), "The Effects of Service Quality and Word-of-Mouth on Customer Acquisition, Retention and Usage," working paper.
- Narayan, Vishal (2007). "Modeling the Formation of Dyadic Relationships Between Consumers in Online Communities," Working Paper.
- Rossi, P.E, R.E. McCulloch and G.M. Allenby (1996), "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15(4), 321-340.
- Ryan, Stephen and Catherine Tucker (2006). "Heterogeneity and the Dynamics of Technology Adoption," Working Paper, MIT.
- Soetevent, A.R. and P. Kooreman (2007). "A Discrete-Choice Model with Social Interactions: With an Application to High School Teen Behavior," *Journal of Applied Econometrics*.
- Su, Che-lin and K. Judd (2008). "Constrained Optimization Approaches to Estimation of Structural Models," Working Paper.
- Train, K.E. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.
- Tucker, C. (2006), "Formal and Informal Influence in the Adoption of Technologies with Network Effects," working paper.
- Van den Bulte, C. and G. Lilien (2001). "Medical Innovation Revisited: Social Contagion versus Marketing Effort," *American Journal of Sociology*, 106(5), 1409-1435.
- Van den Bulte, C. and S. Stremersch (2004). "Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test," *Marketing Science*, 23(4),530-544.
- Vitorino, M. (2007). "Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry," working paper.

Yang, S. and G.M. Allenby (2003). "Modeling Interdependent Consumer Preferences," *Journal of Marketing Research*, XL, 282-294.

Weinberg, C.B. (2005). "Profits out of the Picture: Research Issues and Revenue Sources Beyond the North American Box Office," C.C. Moul, ed. *A Concise Handbook of Movie Industry Economics*. Cambridge University Press, New York, 163-202.

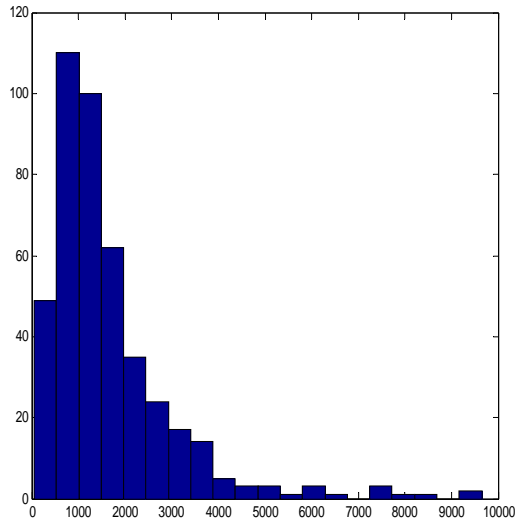
Yang, S., V. Narayan, and H. Assael (2006). "Estimating the Interdependence of Television Program Viewership Between Spouses: A Bayesian Simultaneous Equation Model," *Marketing Science*, 336-349.

Figure 1: Regions of Equilibria



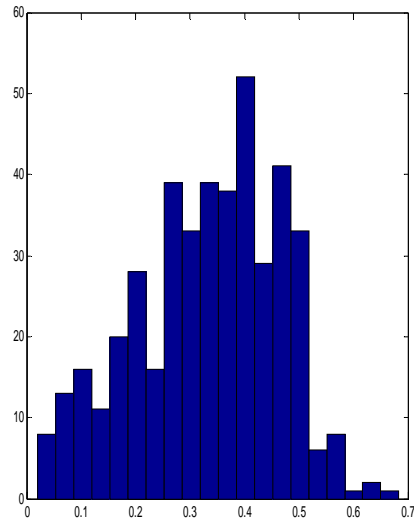
The horizontal and vertical axes in the figure define the net shock to preferences for consuming, $\zeta_i = \varepsilon_{1i} - \varepsilon_{0i}$, for an individual A and B respectively. The dashed lines define indifference between consuming and not when an individual's partner is not consuming. The solid lines define indifference between consuming and not when an individual's partner is consuming. Regions II, V, and VI represent coordinated consumption in that at least one partner is consuming only because the other is consuming. In region V, it is also possible that neither consume, but this is Pareto dominated by both consuming. Region III defines cases in which partners coincidentally consume together. Regions IV, VII, and VIII correspond to neither consuming. Regions I and IX correspond to only one partner consuming.

Figure 2



One Year Customer Value (in dollars)

Figure 3



Fraction of Customer Value from Others in Group

Figure 4a

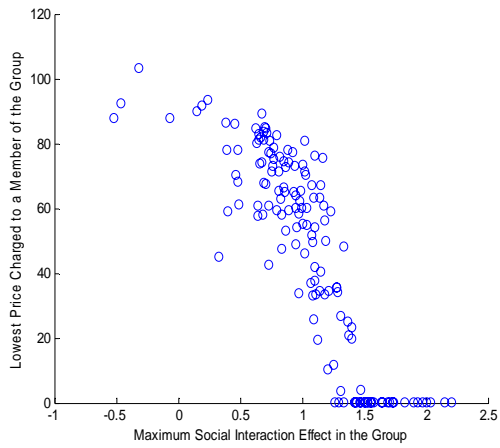


Figure 4b

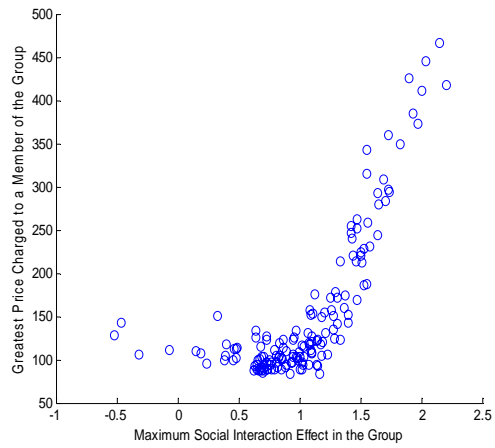


Table 1

Parameter Recovery in Monte Carlos

	Homogeneous Model					Heterogenous Model			
	Estimates	2.5 Percentile	True Values	97.5 Percentile		Median	2.5 Percentile	True Values	97.5 Percentile
Pairs Only					Pairs Only				
Intercept	-3.03	-3.09	-3.00	-2.97	Intercept	-3.02	-3.39	-3.00	-2.67
Partner Effect	0.52	0.48	0.50	0.56	Partner Effect	0.49	0.43	0.50	0.55
X	0.29	0.25	0.25	0.33	X	0.27	0.22	0.25	0.32
Pairs and Threesomes					SD Intercept				
Intercept	-3.04	-3.09	-3.00	-2.98	SD Intercept	1.95	1.74	2.00	2.23
Partner Effect	0.52	0.48	0.50	0.55	SD Partner Effect	0.20	0.14	0.20	0.28
X	0.27	0.24	0.25	0.30	SD X	0.15	0.11	0.10	0.20
Additional Partner	-1.04	-1.23	-1.00	-0.88	Corr Intercepts	0.72	0.60	0.67	0.82
Pairs, Threesomes, and Foursomes					Pairs, Threesomes, and Foursomes				
Intercept	-3.03	-3.08	-3.00	-2.99	Intercept	-3.18	-3.39	-3.00	-2.95
Partner Effect	0.51	0.47	0.50	0.54	Partner Effect	0.47	0.41	0.50	0.52
X	0.28	0.25	0.25	0.31	X	0.28	0.24	0.25	0.32
Additional Partner	-0.99	-1.11	-1.00	-0.89	Additional Partner	-1.07	-1.23	-1.00	-0.88
					SD Intercept	1.70	1.51	2.00	1.92
					SD Partner Effect	0.21	0.15	0.20	0.28
					SD X	0.15	0.11	0.10	0.19
					SD Add Partner	0.21	0.13	0.10	0.37
					Corr Intercept	0.75	0.66	0.67	0.82

*The Partner Effect and Additional Partner parameters are transformed in the likelihood by taking logs. The estimates here are the pre-transformed values.

Table 2a

Bias from Ignoring Endogeneity

	Homogenous Model and Estimation			
	Estimates	2.5 Percentile	True Values	97.5 Percentile
Pairs Only				
Intercept	-3.19	-3.24	-3.00	-3.12
Partner Effect	1.20	1.17	0.50	1.23
X	0.25	0.20	0.25	0.30

Table 2b

Bias from Ignoring Heterogeneity

	Heterogenous Model, Homogeneous Estimation			
	Estimates	2.5 Percentile	True Values	97.5 Percentile
Pairs Only				
Intercept	-2.26	-2.31	-3.50	-2.22
Partner Effect	0.60	0.57	0.50	0.63
X	0.15	0.13	0.25	0.19

Table 3

Summary Statistics of Purchase Data

	N	Number of Times Observed by Group			
		Mean	Standard Deviation	Min	Max
Individuals in a Pair Purchasing					
0	152	319.4	148.2	64	681
1	152	16.7	24.0	0	140
2	152	10.5	12.7	2	114
Individuals in a Threesome Purchasing					
0	34	299.9	121.1	59	589
1	34	22.1	41.4	0	232
2	34	12.1	10.7	2	54
3	34	3.0	3.4	0	17
Individuals in a Foursome Purchasing					
0	7	298.7	98.6	165	454
1	7	34.9	47.0	0	132
2	7	8.9	4.7	1	14
3	7	3.1	3.0	0	9
4	7	1.9	1.7	0	4

Table 4

Summary Statistics of Price, Advertising and Weather

	N	Standard			
		Mean	Deviation	Min	Max
Weekly Ad in Paper	728	0.34	0.46	0	1
Price	728	59.08	11.54	35	75
Rain	728	0.04	0.17	0	2.13

Table 5

Model Estimates

	I			II			III		
	2.5	50	97.5	2.5	50	97.5	2.5	50	97.5
Posterior of Parameter Means									
Intercept	-3.82	-3.64	-3.43	-3.94	-3.78	-3.58	-3.78	-3.63	-3.44
Social Interaction¹	0.61	0.65	0.70	0.65	0.73	0.79	0.64	0.72	0.79
Advertisement in Paper	-0.03	0.02	0.07	-0.06	0.00	0.06	-0.08	-0.02	0.05
Price¹	-0.01	-0.01	-0.01	-4.77	-4.62	-4.47	-4.66	-4.49	-4.37
Weekend	0.33	0.41	0.47	0.01	0.18	0.36	0.07	0.25	0.42
Rain							-0.86	-0.63	-0.47
Fraction of Interaction for Additional Partner¹	-0.54	-0.40	-0.25	-0.87	-0.65	-0.42	-0.74	-0.56	-0.33
Posterior Parameter Standard Deviations									
Intercept	0.83	0.90	1.00	0.93	1.04	1.14	0.91	1.01	1.13
Social Interaction¹	0.70	0.78	0.87	0.50	0.57	0.67	0.51	0.61	0.72
Advertisement in Paper				0.11	0.18	0.27	0.14	0.21	0.27
Price¹				0.13	0.46	0.72	0.52	0.62	0.73
Weekend				1.49	1.64	1.80	1.51	1.66	1.80
Rain							0.12	0.21	0.35
Fraction of Interaction for Additional Partner¹				0.17	0.38	0.87	0.21	0.37	0.75
Correlations Between Parameters									
Intercepts within Group	0.28	0.45	0.58	0.23	0.43	0.62	0.35	0.57	0.75

1. These parameters represent the mean and standard deviation for the underlying normal distribution of these log-normally distributed parameters. In the case of the price coefficient, the first model does not transform the price coefficient so that its sign could be verified before assuming the sign for the heterogeneous specification which is log-normally distributed.

Table 6

Estimates of Model w/o Social Interaction

	Percentiles of Distribution		
	Median	2.5	97.5
Intercept	-2.68	-2.91	-2.56
Std Dev of Intercept	0.84	0.78	0.90
Weekend	0.49	0.42	0.56
Price	-0.01	-0.01	-0.01
Advertisement in Paper	0.02	-0.03	0.08

Appendix 1: Four Player Equilibrium and Likelihood Derivation

Defining a likelihood for a group's outcome involves integration over a set of unobservables, where each vector of unobservables in the set i) justifies the outcome as an equilibrium and ii) does not generate an alternative equilibrium outcome that is favored by the equilibrium selection rule. We focus only on pure strategy Nash equilibria. When dealing with multiple equilibria, we select equilibria by assuming that whenever multiple equilibria exist, the group chooses the outcome with the most players purchasing. In the case of golf, the implicit assumption is that the players would rather purchase with a group member than the random person they may be assigned to complete their foursome at the course. This always provides a Pareto dominant equilibrium because at least one player will always be better off and no player can be worse off, given Assumption 1 in the text.

Throughout this appendix, we use the following notation. $v_{1,-i1,-i2,-i3,i}$ is the observable component of the purchase utility for player i , conditional on the actions of the other three players. For example, v_{1011C} is player C's utility from purchasing when player A does not purchase, but players B and D do purchase. The net unobserved purchase utility for player i is expressed as $\zeta_i = \varepsilon_{1i} - \varepsilon_{0i}$.

In the case of four players, there are 16 potential outcomes. Of these 16, there are 4 distinct types for which we must separately consider the likelihood derivations: i) all players purchasing, ii) three of the players purchasing; iii) two players purchasing, iv) one player purchasing. We do not need to consider the case of no players purchasing because the likelihood of that is defined to be one minus the sum of the other likelihoods.

A three player likelihood can be derived in a similar manner, or one can assume that player D in the following always has an infinitely negative utility from purchasing.

1 Four Players Purchasing

All four players purchasing is an equilibrium if the following conditions holds:

$$(v_{1111A} + \varsigma_A \geq 0); (v_{1111B} + \varsigma_B \geq 0); (v_{1111C} + \varsigma_C \geq 0); (v_{1111D} + \varsigma_D \geq 0) \quad (\text{A1.1})$$

Given the equilibrium selection assumption described above, there cannot exist another equilibrium that will be chosen instead of this outcome. The equilibrium condition therefore defines the likelihood for this outcome:

$$\Pr(1111) = \Pr[(v_{1111A} + \varsigma_A \geq 0)(v_{1111B} + \varsigma_B \geq 0)(v_{1111C} + \varsigma_C \geq 0)(v_{1111D} + \varsigma_D \geq 0)] \quad (\text{A1.2})$$

2 Three Players Purchasing

Three players purchasing (e.g. all except D) is an equilibrium if the following condition holds:

$$(v_{1110A} + \varsigma_A \geq 0); (v_{1110B} + \varsigma_B \geq 0); (v_{1110C} + \varsigma_C \geq 0); (v_{1111D} + \varsigma_D < 0) \quad (\text{A1.3})$$

Our equilibrium selection assumption that equilibria with more purchasers always dominates tells us that we only need to consider other 3 purchaser equilibria and 4 purchaser equilibria. We consider them in descending order of the number of purchasers.

- ***There does not exist a four purchaser equilibrium that is also a three purchaser equilibrium (e.g. (1,1,1,0) vs. (1,1,1,1)***: To show this we establish that outcome (1,1,1,0) does not have a vector of unobservables in common with outcome (1,1,1,1). A necessary condition for a multiple equilibria problem to arise between the outcomes (1,1,1,0) and (1,1,1,1) is that there must be overlapping ζ_i 's for all i . Therefore, showing a lack of overlap of ζ_i 's for at least one player establishes that multiple equilibria cannot arise between the outcomes. For these two outcomes, there are no overlapping unobservables for the best responses of player D:

$$(1,1,1,1) : v_{1111D} + \zeta_D \geq 0$$

$$(1,1,1,0) : v_{1111D} + \zeta_D < 0$$

The above reveals that ζ_D cannot exist in both regions. Note that we have assumed that if a customer is indifferent to purchasing, he purchases. We assume this throughout our analysis.

- ***For each three purchaser equilibrium , there does not exist any other three purchaser outcome that is an equilibrium generated by a common set of unobservables (e.g. (1,1,1,0) vs. (0,1,1,1)***: To illustrate, we focus on the inequalities for player D in these two hypothetical 3 purchaser outcomes.

$$(1,1,1,0) : v_{1111D} + \zeta_D < 0$$

$$(0,1,1,1) : v_{1011D} + \zeta_D \geq 0$$

$$\Rightarrow v_{1111D} + \zeta_D \geq 0 \text{ because } v_{1111D} \geq v_{1011D}$$

The last line in the preceding indicates that because of Assumption 1 in the text, if D is willing to purchase when B and C are purchasing, then D also must be willing to purchase when A is also purchasing. However, this inequality directly contradicts the first inequality. Therefore, (1,1,1,0) and (0,1,1,1) cannot both be equilibria generated by the same set of unobservables.

The equilibrium condition therefore defines the likelihood for this outcome:

$$\Pr(1110) = \Pr[(v_{1110A} + \varsigma_A \geq 0)(v_{1110B} + \varsigma_B \geq 0)(v_{1110C} + \varsigma_C \geq 0)(v_{1111D} + \varsigma_D < 0)] \quad (\text{A1.4}).$$

3 Two Players Purchasing

Two players purchasing (e.g. all except C and D) is an equilibrium if the following condition holds:

$$(v_{1100A} + \varsigma_A \geq 0); (v_{1100B} + \varsigma_B \geq 0); (v_{1110C} + \varsigma_C < 0); (v_{1110D} + \varsigma_D < 0) \quad (\text{A1.5})$$

The equilibrium selection assumption that equilibria with more purchasers always dominates tells us that we only need to consider other 2 purchaser equilibria and 3 and 4 purchaser equilibria. We consider them in descending order of the number of purchasers.

- ***For every two purchaser outcome there exists a four player equilibrium that would be chosen instead because of the greater number of players purchasing (e.g. (1,1,0,0) vs. (1,1,1,1)):*** The vector of unobservables in common to both a (1,1,1,1) and (1,1,0,0) equilibrium is defined as follows:

$$\begin{aligned}
(1,1,0,0): & \quad v_{1100A} + \varsigma_A \geq 0 & v_{1100B} + \varsigma_B \geq 0 & v_{1110C} + \varsigma_C < 0 & v_{1110D} + \varsigma_D < 0 \\
(1,1,1,1): & \quad v_{1111A} + \varsigma_A \geq 0 & v_{1111B} + \varsigma_B \geq 0 & v_{1111C} + \varsigma_C \geq 0 & v_{1111D} + \varsigma_D \geq 0 \\
\text{Overlap:} & \quad \varsigma_A \geq -v_{1100A} > -v_{1111A} & \varsigma_B \geq -v_{1100B} > -v_{1111B} & -v_{1111C} \leq \varsigma_C < -v_{1110C} & -v_{1111D} \leq \varsigma_D < -v_{1110D}
\end{aligned}$$

(A1.6)

- **There does not exist a three purchaser equilibrium that is also a two purchaser equilibrium (e.g. (1,1,0,0) vs. (1,1,1,0) or (0,1,1,1)):** This is shown by illustrating that there are no overlapping unobservables for at least one of the players when

comparing (1,1,0,0) to either (1,1,1,0) or (0,1,1,1):

$$\begin{aligned}
(1,1,0,0): & \quad v_{1100A} + \varsigma_A \geq 0 & v_{1100B} + \varsigma_B \geq 0 & v_{1110C} + \varsigma_C < 0 & v_{1110D} + \varsigma_D < 0 \\
(1,1,1,0): & \quad v_{1110A} + \varsigma_A \geq 0 & v_{1110B} + \varsigma_B \geq 0 & v_{1110C} + \varsigma_C \geq 0 & v_{1111D} + \varsigma_D < 0 \\
\text{Overlap:} & \quad \varsigma_A \geq -v_{1100A} \geq -v_{1110A} & \varsigma_B \geq -v_{1100B} \geq -v_{1110B} & \text{NONE} & -v_{1111D} \leq -v_{1110D} < \varsigma_D
\end{aligned}$$

(A1.7)

(1,1,0,0):	$v_{1100A} + \varsigma_A \geq 0$	$v_{1100B} + \varsigma_B \geq 0$	$v_{1110C} + \varsigma_C < 0$	$v_{1110D} + \varsigma_D < 0$
(0,1,1,1):	$v_{1111A} + \varsigma_A < 0$	$v_{1011B} + \varsigma_B \geq 0$	$v_{1011C} + \varsigma_C \geq 0$	$v_{1111D} + \varsigma_D < 0$
Overlap:	NONE: $\varsigma_A \geq -v_{1100A}$, but $-v_{1100A} \geq -v_{1111A}$ so $v_{1111A} + \varsigma_A \geq 0$	$\varsigma_B \geq -v_{1100B} \geq -v_{1011B}$	<i>only if</i> $v_{1011C} > v_{1110C}$ <i>because</i> $\Omega_{AC} = 0$ <i>and</i> $\Omega_{CD} = 1$	<i>if</i> $-v_{1011D} \leq -v_{1110D}$ $-v_{1011D} \leq -v_{1110D} < \varsigma_D$ <i>if</i> $-v_{1110D} \leq -v_{1011D}$ $-v_{1110D} \leq -v_{1011D} < \varsigma_D$

(A1.8)

The likelihood for a two purchaser equilibrium is therefore the region of unobservables defined by the equilibrium condition in (A1.5), less the region of unobservables that overlaps with equilibrium (1,1,1,1), i.e. that defined in (A1.6):

$$\begin{aligned}
\Pr(1100) = & \Pr\left[(v_{1100A} + \varsigma_A \geq 0)(v_{1100B} + \varsigma_B \geq 0)(v_{1110C} + \varsigma_C < 0)(v_{1110D} + \varsigma_D < 0)\right] \\
& - \Pr\left[(v_{1100A} + \varsigma_A \geq 0)(v_{1100B} + \varsigma_B \geq 0)(-v_{1111C} \leq \varsigma_C < -v_{1110C})(-v_{1111D} \leq \varsigma_D < -v_{1110D})\right]
\end{aligned}$$

(A1.9)

A similar calculation can be done for all other two purchaser equilibria.

4 One Player Purchasing

One player purchasing (e.g. only A) is an equilibrium if the following condition holds:

$$(v_{1000A} + \varsigma_A \geq 0); (v_{1100B} + \varsigma_B < 0); (v_{1100C} + \varsigma_C < 0); (v_{1100D} + \varsigma_D < 0) \quad (A1.10)$$

The equilibrium selection assumption that equilibria with more purchasers always dominates tells us that we only need to consider other 1 purchaser equilibria and 2, 3 and 4 purchaser equilibria. We consider them in descending order of the number of purchasers.

- ***For every one purchaser outcome there exists a four player equilibrium that would be chosen instead because of the greater number of players purchasing (e.g. (1,0,0,0) vs. (1,1,1,1)):*** The vector of unobservables in common to both a

(1,1,1,1) and (1,0,0,0) equilibrium is defined as follows:

$$\begin{array}{llll} (1,0,0,0): & v_{1000A} + \varsigma_A \geq 0 & v_{1100B} + \varsigma_B < 0 & v_{1100C} + \varsigma_C < 0 & v_{1100D} + \varsigma_D < 0 \\ (1,1,1,1): & v_{1111A} + \varsigma_A \geq 0 & v_{1111B} + \varsigma_B \geq 0 & v_{1111C} + \varsigma_C \geq 0 & v_{1111D} + \varsigma_D \geq 0 \\ \text{Overlap:} & \varsigma_A \geq -v_{1000A} > -v_{1111A} & -v_{1111B} \leq \varsigma_B < -v_{1100B} & -v_{1111C} \leq \varsigma_C < -v_{1100C} & -v_{1111D} \leq \varsigma_D < -v_{1100D} \end{array}$$

(A1.11)

- ***For every one purchaser outcome there exists 3 three player equilibria that would be chosen instead because of the greater number of players purchasing***

(e.g. $(1,0,0,0)$ vs. $(1,1,1,0)$, $(1,0,1,1)$, $(1,1,0,1)$): The vector of unobservables in common to are defined as follows:

$$\begin{array}{l}
(1,0,0,0): \quad v_{1000A} + \zeta_A \geq 0 \quad v_{1100B} + \zeta_B < 0 \quad v_{1100C} + \zeta_C < 0 \quad v_{1100D} + \zeta_D < 0 \\
(1,1,1,0): \quad v_{1110A} + \zeta_A \geq 0 \quad v_{1110B} + \zeta_B \geq 0 \quad v_{1110C} + \zeta_C \geq 0 \quad v_{1111D} + \zeta_D < 0 \\
\text{Overlap:} \quad \zeta_A \geq -v_{1000A} > -v_{1110A} \quad -v_{1110B} \leq \zeta_B < -v_{1100B} \quad -v_{1110C} \leq \zeta_C < -v_{1100C} \quad \zeta_D < -v_{1111D} < -v_{1100D} \\
\\
(1,0,0,0): \quad v_{1000A} + \zeta_A \geq 0 \quad v_{1100B} + \zeta_B < 0 \quad v_{1100C} + \zeta_C < 0 \quad v_{1100D} + \zeta_D < 0 \\
(1,0,1,1): \quad v_{1011A} + \zeta_A \geq 0 \quad v_{1111B} + \zeta_B < 0 \quad v_{1101C} + \zeta_C \geq 0 \quad v_{1101D} + \zeta_D \geq 0 \\
\text{Overlap:} \quad \zeta_A \geq -v_{1000A} > -v_{1011A} \quad \zeta_B < -v_{1111B} < -v_{1100B} \quad -v_{1101C} \leq \zeta_C < -v_{1100C} \quad -v_{1101D} \leq \zeta_D < -v_{1100D} \\
\\
(1,0,0,0): \quad v_{1000A} + \zeta_A \geq 0 \quad v_{1100B} + \zeta_B < 0 \quad v_{1100C} + \zeta_C < 0 \quad v_{1100D} + \zeta_D < 0 \\
(1,1,0,1): \quad v_{1101A} + \zeta_A \geq 0 \quad v_{1101B} + \zeta_B \geq 0 \quad v_{1111C} + \zeta_C < 0 \quad v_{1110D} + \zeta_D \geq 0 \\
\text{Overlap:} \quad \zeta_A \geq -v_{1000A} > -v_{1101A} \quad -v_{1101B} \leq \zeta_B < -v_{1100B} \quad \zeta_C < -v_{1111C} < -v_{1100C} \quad -v_{1110D} \leq \zeta_D < -v_{1100D}
\end{array}$$

(A1.12)

- **There does not exist a two purchaser equilibrium that is also a one purchaser equilibrium (e.g. $(1,0,0,0)$ vs. $(1,1,0,0)$ or $(0,1,1,0)$):** This is shown by illustrating that there are no overlapping unobservables for at least one of the players when comparing $(1,0,0,0)$ to either $(1,1,0,0)$ or $(0,1,1,0)$:

$(1,0,0,0):$	$v_{1000A} + \zeta_A \geq 0$	$v_{1100B} + \zeta_B < 0$	$v_{1100C} + \zeta_C < 0$	$v_{1100D} + \zeta_D < 0$
$(0,0,1,1):$	$v_{1011A} + \zeta_A < 0$	$v_{1011B} + \zeta_B < 0$	$v_{1001C} + \zeta_C \geq 0$	$v_{1001D} + \zeta_D \geq 0$
Overlap:	NONE: $\zeta_A \geq -v_{1000A}$, but $-v_{1000A} \geq -v_{1011A}$ so $v_{1011A} + \zeta_A \geq 0$	if $\Omega_{BC} = 1$ or $\Omega_{BD} = 1$ $\zeta_B < -v_{1011B} \leq -v_{1100B}$ otherwise $\zeta_B < -v_{1100B} < -v_{1011B}$	<i>onlyif</i> $v_{1001C} > v_{1100C}$ because $\Omega_{AC} = 0$ and $\Omega_{CD} = 1$	<i>onlyif</i> $v_{1001D} > v_{1100D}$ because $\Omega_{AD} = 0$ and $\Omega_{CD} = 1$
$(1,0,0,0):$	$v_{1000A} + \zeta_A \geq 0$	$v_{1100B} + \zeta_B < 0$	$v_{1100C} + \zeta_C < 0$	$v_{1100D} + \zeta_D < 0$
$(1,1,0,0):$	$v_{1100A} + \zeta_A \geq 0$	$v_{1100B} + \zeta_B \geq 0$	$v_{1110C} + \zeta_C \geq 0$	$v_{1110D} + \zeta_D \geq 0$
Overlap:	$\zeta_A \geq -v_{1000A} \geq -v_{1011A}$	NONE	$-v_{1110C} \leq \zeta_C < -v_{1100C}$	$-v_{1110D} \leq \zeta_D < -v_{1100D}$

(A1.13)

The likelihood for a one purchaser equilibrium is therefore the region of unobservables defined by the equilibrium condition in (A1.10), less the region of unobservables that overlaps with equilibria (1,1,1,1), (1,1,1,0), (1,0,1,1), (1,1,0,1), i.e. those defined in (A1.11) and (A1.12). :

$$\begin{aligned}
\Pr(1000) = & \Pr\left[(v_{1000A} + \varsigma_A \geq 0)(v_{1100B} + \varsigma_B < 0)(v_{1100C} + \varsigma_C < 0)(v_{1100D} + \varsigma_D < 0)\right] \\
& - \Pr\left[(\varsigma_A \geq -v_{1000A})(-v_{1111B} \leq \varsigma_B < -v_{1100B})(-v_{1111C} \leq \varsigma_C < -v_{1100C})(-v_{1111D} \leq \varsigma_D < -v_{1100D})\right] \\
& - \Pr\left[(\varsigma_A \geq -v_{1000A})(-v_{1110B} \leq \varsigma_B < -v_{1100B})(-v_{1110C} \leq \varsigma_C < -v_{1100C})(\varsigma_D < -v_{1111D})\right] \\
& - \Pr\left[(\varsigma_A \geq -v_{1000A})(\varsigma_B < -v_{1111B})(-v_{1101C} \leq \varsigma_C < -v_{1100C})(-v_{1101D} \leq \varsigma_D < -v_{1100D})\right] \\
& - \Pr\left[(\varsigma_A \geq -v_{1000A})(-v_{1101B} \leq \varsigma_B < -v_{1100B})(\varsigma_C < -v_{1111C})(-v_{1110D} \leq \varsigma_D < -v_{1100D})\right]
\end{aligned}$$

A similar calculation can be done for all other one purchaser equilibria.

Appendix 2: MCMC Steps

Estimation of the model proceeds by iteratively applying the following 7 steps after defining starting values of the parameters. I draw the group-level parameters in three separate sets: the intercepts, the social interaction parameters, and the coefficients on the covariates. When drawing the population parameters, I draw the mean and variance of the intercepts using Metropolis Hastings because of restrictions in the covariance matrix. I draw the remaining population means and variances using the conjugate normal-chi-squared specification. I draw the correlation parameter using the Griddy Gibbs Sampler.

1 Draw Intercepts for Each Group

The posterior of a group's intercepts, conditional on all other parameters and data is expressed as follows:

$$P\left(\gamma_{01g}, \dots, \gamma_{0N_g} \mid \begin{matrix} y_{gt}, X, \gamma_{11g}, \dots, \gamma_{1N_g}, \\ \beta_{1g}, \dots, \beta_{N_g}, \theta, S, R \end{matrix}\right) \propto \prod_t L(y_{gt} \mid X, \gamma_{1g}, \dots, \gamma_{N_g}) p(\gamma_{01g}, \dots, \gamma_{0N_g} \mid \theta, S, R) \quad (\text{A2.1})$$

where L is as defined in Section 3 of the paper or Appendix 1 depending on whether there are 2, 3, or 4 players in a group. The prior is defined to be the multivariate normal heterogeneity distribution defined in the paper. Given the logit specification, I use Metropolis Hastings by drawing a candidate value of $\{\gamma_{01g}, \dots, \gamma_{0N_g}\}$ from a random walk around the existing value $\{\gamma_{01g}, \dots, \gamma_{0N_g}\}$. A comparison of the right-hand sides of Equation (A2.1) determines whether to accept the new candidate value or retain the old value. In the paper, I only estimate the correlation within intercepts, so (A2.1) does not need to condition on the other group-level parameters, but this would be an obvious extension if deemed necessary. In the present case, the correlations among the intercepts

capture homophily, so I reduced computational time by assuming other correlations to be zero. As described in the text, the prior on the group level parameters is assumed to be multivariate normal with mean vector θ and a covariance matrix defined to be *SRS*.

2 Draw Partner Effects for Each Group

The posterior of a group's Partner Effects, conditional on all other parameters and data is expressed as follows:

$$p\left(\ln \gamma_{11g}, \dots, \ln \gamma_{1N_g} \mid \begin{matrix} y_{gt}, X, \gamma_{01g}, \dots, \gamma_{0N_g}, \\ \beta_{1g}, \dots, \beta_{N_g}, \theta, S, R \end{matrix}\right) \propto \prod_t L(y_{gt} \mid X, \gamma_{1g}, \dots, \gamma_{N_g}) p(\ln \gamma_{1g}, \dots, \ln \gamma_{1N_g} \mid \theta, S) \quad (\text{A2.2})$$

where the prior is defined to be the multivariate normal heterogeneity distribution defined in the paper. As above, I use Metropolis Hastings by drawing a candidate value of $\{\ln \gamma_{11g}, \dots, \ln \gamma_{1N_g}\}$ from a random walk around the existing value $\{\ln \gamma_{11g}, \dots, \ln \gamma_{1N_g}\}$. R does not affect the posterior in this case because I am not estimating the correlations between the Partner Effects or any other parameters. Most of the groups are pairs, so as described in the text, the correlation between Partner Effects is poorly identified without exclusion restrictions.

3 Draw Covariate Effects for Each Group

The posterior of a group's covariate effects, conditional on all other parameters and data is expressed as follows:

$$P\left(\beta_{1g}, \dots, \beta_{Ng} \mid \begin{matrix} y_{gt}, X, \gamma_{01g}, \dots, \gamma_{0Ng}, \\ \gamma_{11g}, \dots, \gamma_{1Ng}, \theta, S, R \end{matrix}\right) \propto \prod_t L(y_{gt} \mid X, \beta_{1g}, \dots, \beta_{Ng}) P(\beta_{1g}, \dots, \beta_{Ng} \mid \theta, S)$$

(A2.3)

where the prior is defined to be the multivariate normal heterogeneity distribution defined in the paper. As above, I use Metropolis Hastings by drawing a candidate value of $\{\beta_{1g}, \dots, \beta_{Ng}\}$ from a random walk around the existing value $\{\beta_{1g}, \dots, \beta_{Ng}\}$.

4 Draw Mean of Population Distribution of Intercepts

The posterior of the mean of the population distribution of intercepts is expressed as follows:

$$P(\gamma_0 \mid \gamma_{011}, \dots, \gamma_{0N_1}, \dots, \gamma_{01G}, \dots, \gamma_{0N_G}, S, R) \propto \left[\prod_g N(\gamma_{01g}, \dots, \gamma_{0N_g} \mid \gamma_0, SRS) \right] P(\gamma_0)$$

(A2.4)

where N is the normal distribution of the parameter conditional on its mean and variance, and the prior, $p(\gamma_0)$, is defined to be normal with mean zero and variance 1000. I also use Metropolis Hastings by drawing a candidate value of γ_0 from a random walk around the existing value γ_0 .

5 Draw Variance of Population Distribution of Intercepts

The posterior of the variance of the population distribution of intercepts is expressed as follows:

$$P(\ln \sigma_0 \mid \gamma_{011}, \dots, \gamma_{0N_1}, \dots, \gamma_{01G}, \dots, \gamma_{0N_G}, \gamma_0, R) \propto \left[\prod_g N \left(\begin{bmatrix} \gamma_{01g} \\ \vdots \\ \gamma_{0N_g} \end{bmatrix} \mid \begin{bmatrix} \gamma_0 \\ \vdots \\ \gamma_0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \rho_0 & \rho_0 \\ \rho_0 & \ddots & \rho_0 \\ \rho_0 & \rho_0 & \sigma_0^2 \end{bmatrix} \right) \right] P(\ln \sigma_0)$$

(A2.5)

where the prior $p(\ln \sigma_0)$ is distributed normal with mean zero and variance 1000. I also use Metropolis Hastings by drawing a candidate value of $\ln \sigma_0$ from a random walk around the existing value $\ln \sigma_0$.

6 Draw Means and Variances of all other Population Distribution Parameters

The posterior of the mean and variance of the population distribution of parameters is expressed as follows:

$$p(\theta_m, \sigma_m^2 | \theta_{m11}, \dots, \theta_{mN_11}, \dots, \theta_{m1G}, \dots, \theta_{mN_GG}) \propto \left[\prod_{g=1}^G \prod_{i=1}^{N_g} N(\theta_{mig} | \theta_m, \sigma_m^2) \right] p(\theta_m, \sigma_m^2 | A, v_0, s_0) \quad (\text{A2.6})$$

where the priors on the mean and variance are the standard conjugate priors with an inverse chi-squared prior on the variance and a conditional normal prior on the mean., i.e.

$$p(\theta_m, \sigma_m^2 | A, v_0, s_0) = p(\theta_m | \sigma_m^2, A) p(\sigma_m^2 | v_0, s_0) \text{ and}$$

$$p(\theta_m | \sigma_m^2, A) \sim N(\theta_m | 0, \sigma_m^2 A^{-1}), \quad p(\sigma_m^2 | v_0, s_0) \sim (v_0 s_0^2 / \chi_{v_0}^2) \text{ where } v_0 = 4,$$

$s_0^2 = 0.05$, and $A = 0.001$. We therefore obtain the following analytic expressions for the posteriors of these parameters:

$$p(\theta_m | \theta_{m11}, \dots, \theta_{mN_11}, \dots, \theta_{m1G}, \dots, \theta_{mN_GG}, \sigma_m^2, A) \sim N\left(\tilde{\theta}_m, \sigma_m^2 \left(\left(\sum_{g=1}^G \sum_{i=1}^{N_g} 1 \right) + A \right)^{-1}\right)$$

$$p(\sigma_m^2 | \theta_{m11}, \dots, \theta_{mN_11}, \dots, \theta_{m1G}, \dots, \theta_{mN_GG}, A, v_0, s_0) \sim \frac{\left[v_0 + \left(\sum_{g=1}^G \sum_{i=1}^{N_g} 1 \right) \right] \left[v_0 s_0^2 + \left[\sum_{g=1}^G \sum_{i=1}^{N_g} (\theta_{mig} - \tilde{\theta}_m)^2 \right] + \tilde{\theta}_m^2 A \right]}{\chi_{v_0 + \left(\sum_{g=1}^G \sum_{i=1}^{N_g} 1 \right)}^2}$$

7 Draw Correlation Parameter

The posterior of the correlation of the population distribution of intercepts is expressed as follows:

$$p(\rho_0 \mid \gamma_{011}, \dots, \gamma_{0N_11}, \dots, \gamma_{01G}, \dots, \gamma_{0N_GG}, \gamma_0, \mathbf{S}) \propto \left[\prod_g N \left(\begin{bmatrix} \gamma_{01g} \\ \vdots \\ \gamma_{0N_g} \end{bmatrix} \mid \begin{bmatrix} \gamma_0 \\ \vdots \\ \gamma_0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \rho_0 & \rho_0 \\ \rho_0 & \ddots & \rho_0 \\ \rho_0 & \rho_0 & \sigma_0^2 \end{bmatrix} \right) \right] p(\rho_0)$$

where $p(\rho_0)$ is assumed to be uniformly distributed between the interval -1 to 1. To draw these I use a Griddy Gibbs sampler.