



Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Revealed Social Networks

*Christopher P. Chambers, Yusufcan Masatlioglu, and
Christopher Turansick*

Working Paper n. 712

This Version: January, 2025

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
<http://www.igier.unibocconi.it>

The opinions expressed in the working papers are those of the authors alone, and not those of the Institute, which takes non institutional policy position, nor those of CEPR, NBER or Università Bocconi.

Revealed Social Networks*

Christopher P. Chambers[†] Yusufcan Masatlioglu[‡] Christopher Turansick[§]

January 5, 2025

Abstract

People are influenced by their peers when making decisions. In this paper, we study the linear-in-means model which is the standard empirical model of peer effects. As data on the underlying social network is often difficult to come by, we focus on data that only captures an agent's choices. Under exogenous agent participation variation, we study two questions. We first develop a revealed preference style test for the linear-in-means model. We then study the identification properties of the linear-in-means model. With sufficient participation variation, we show how an analyst is able to recover the underlying network structure and social influence parameters from choice data. Our identification result holds when we allow the social network to vary across contexts. To recover predictive power, we consider a refinement which allows us to extrapolate the underlying network structure across groups and provide a test of this version of the model.

Keywords: Revealed Preference, Social Interactions, Linear-in-Means, Peer Effects

1 Introduction

People rarely make decisions in isolation and are often influenced by their neighbors and peers. This influence can take the form of a social norm, a common convention, or conformism. The amount of influence a person faces may be heterogeneous across their peers with friends and family exerting higher influence than distant connections. A common model of conformism based influence is the

*

[†]Department of Economics, Georgetown University, ICC 580 37th and O Streets NW, Washington DC 20057. E-mail: Christopher.Chambers@georgetown.edu.

[‡]Department of Economics, University of Maryland, 3147E Tydings Hall, 7343 Preinkert Dr., College Park, MD 20742. E-mail: yusufcan@umd.edu.

[§]Department of Decision Sciences, Bocconi University and IGIER E-mail: christopher.turansick@unibocconi.it

linear-in-means model. The linear-in-means model supposes that a person’s realized choice is a convex combination of their ideal point and a weighted average of every other person’s choice. We study the linear-in-means model and answer two questions. First, we characterize datasets that are consistent with the linear-in-means model when the set of people making decisions varies. Second, using this data, we provide sufficient conditions under which we can recover both an agent’s ideal point and the underlying social network, including the intensity of each connection.

Social influence has been studied in an array of contexts including test scores (Sacerdote, 2001), worker productivity (Mas and Moretti, 2009), alcohol use (Kremer and Levy, 2008), risky behavior (Card and Giuliano, 2013), and tax compliance (Fortin et al., 2007). Testing models of peer effects is difficult as many people choose peers who are similar to them in observable characteristics. As such, many studies which aim to test models of peer effects, including the linear-in-means model, do so through natural/quasi/pure experimental methods (Sacerdote, 2014; Basse et al., 2024).¹ Our approach differs from these tests. We consider data that captures an agent’s choice frequencies in the context of different groups.² In the spirit of Afriat (1967), our test characterizes datasets of the prior form which are consistent with the linear-in-means model via an easily solvable linear program. We interpret this linear program as a no money-pump condition on an outside observer who is making bets on the choices of an agent. Unlike standard no money-pump conditions, which are typically given by two conditions, feasibility of a bet and (expected) profitability of the bet, our condition has a third part which imposes incentive compatibility of a bet. Here incentive compatibility captures the idea that if the outside observer is betting on one decision maker across each group, there is no group where they would prefer to bet on a different decision maker.

In addition to the problem of testing, since the work of Manski (1993), it has been known that identifying social influence parameters is difficult. At the core of the reflection problem of Manski (1993) is the difficulty of recovering and differentiating the impact of exogenous group effects and endogenous peer effects. As De Paula et al. (2024) notes, another aspect that leads to difficulty in identifying peer effect parameters is (a lack of) knowledge of the underlying social network. Much of the work on identifying parameters in the linear-in-means model assumes (partial) observability of the underlying network structure (Bramoullé et al., 2009; Blume et al., 2015). De Paula et al. (2024) provides sufficient conditions under which the underlying network structure can be recovered without data on the network itself. Our identification results complement this work. Unlike De Paula et al.

¹Experimental methods are also often used when quantifying the impact of peer effects. See Agranov et al. (2021) as an example.

²Our main focus in this paper is on group variation. However, our main results can be applied in a setting where there is no group variation but there is variation of some observable variable that causes unobservable variation in the underlying network structure of a fixed group. In this sense, we can apply both our testing and identification results when analysts observe an instrumental variable for the underlying network structure.

(2024) which requires characteristic variation, our results require no characteristic variation but rely on group variation. Further, our focus is on the case when an agent's outcome variable is a distribution over some finite set X . This can be thought of as the average consumption bundle of an agent or an agent's distribution of time usage. With sufficient group variation and putting no restrictions on the network structure across different groups, we provide sufficient conditions which allow us to recover the (weighted) network structure in each of these groups. A key insight of our analysis is that a necessary condition for recovery of the underlying network structure is that the number of agents in each of our groups, given by $|N|$, must be no more than $|X|$, the dimension of our outcome variable.

Thus far we have made no assumptions on the underlying social network structure across groups. While this is a strong point of our testing and identification results, this poses a problem for the predictive power of our model. As such, we consider a refinement of the linear-in-means model which assumes a common social network structure across groups. That is to say, the absolute importance of one person to another is fixed across groups. The strength of connection from one person to another only varies across groups due to renormalization. Under this assumption, we provide sufficient conditions which allow us to predict choices in any possible group. To test the validity of this assumption, we develop an extension of our testing procedure for the general case. This version of the linear-in-means model is characterized by a type of no incentive compatible money pump condition. This time incentive compatibility corresponds to the idea that if the outside observer is betting on one decision maker across each group, there is no other agent that the outside observer would rather bet on across all of the same groups.

Finally, we consider a version of the linear-in-means model more in line with the original reflection problem posed by Manski (1993). In this version of the model, each person influences each other person uniformly. This corresponds to the unweighted average choice of the group being a common social norm within the group. In this case, we provide a test in terms of a finite set of linear inequalities. In the context of linear social influence models, such as the linear-in-means model, the influence one person exerts on another is proportional to the difference of their choices. As such, we call the difference between agent i 's choice and agent j 's choice the peer effect of agent j on agent i . This version of the model is characterized by three restrictions on the peer effects between agents. Our first axiom asks that the peer effect of agent j on agent i is group invariant. Our second axiom is a condition about the symmetry of agent i 's peer effect across groups. The last axiom asks that the total peer effect on agent i in group N is bounded above by agent i 's choice in group N .

The rest of this paper is organized as follows. In Section 2 we formally introduce the linear-in-means model and our notation. In Section 3 we introduce and discuss our testing and identification results for the three different specifications of the linear-in-means model. Finally, we conclude and discuss the related literature in Section 4.

2 Model and Preliminaries

Our interest is in studying the linear-in-means model of social interaction. We build on the base model in two meaningful ways. First, instead of restricting an agent's choice to be from an interval, we allow agents to choose a distribution over a finite set of goods. Second, we consider stochastic choice data that arises when the set of agents present varies.

2.1 Preliminaries and Notation

Denote by \mathcal{A} the grand set of agents. Assume that $|\mathcal{A}| \geq 2$. A typical group of agents will be denoted N , where $\emptyset \neq N \subseteq \mathcal{A}$. We let $\mathcal{N} \subseteq 2^{\mathcal{A}}$ be any set of *groups*. For any agent $i \in \mathcal{A}$, let $\mathcal{N}_i = \{N \in \mathcal{N} : i \in N\}$ denote the set of groups to which i belongs. Note that we sometimes abuse notation and use $N \setminus i$ to denote the group of agents formed by removing agent i from group N . Let X be some finite set of alternatives. Enumerate these alternatives from 1 to $|X|$. We use $(0, \dots, 0, 1, 0, \dots, 0)$, where 1 is in the n th dimension, to denote the n th alternative and use $x, y \in X$ to denote arbitrary alternatives. The data in our model consist of a probability distribution over X , for each $N \in \mathcal{N}$ and each $i \in N$. Formally, for $i \in N \in \mathcal{N}$, this is denoted $p_i^N \in \Delta(X)$. We use $p_i^N(x)$ to denote the choice probability of good x by agent i in group N and p^N to denote the matrix where each row corresponds to p_i^N for a different $i \in N$.

2.2 The Model

In the linear-in-means model, agents take actions to maximize their utility. Each agent's utility depends on some agent-specific parameter and the actions of each other agent. We use $v_i \in \Delta(X)$ to denote agent i 's ideal point. This corresponds to the action agent i would take in isolation. We introduced p_i^N as our data, but it also corresponds to the action taken by agent i in group N . The amount that agent j 's action impacts agent i 's action may differ from the amount that agent k 's action impacts agent i 's action. In fact, the impact of agent j 's action on agent i may depend on the context or the group of agents currently present. We use $\pi_i^N(j) \geq 0$ to denote the impact of agent j 's action on agent i in the context of group N . Similarly, we use $\pi_i^N(i) > 0$ to denote the impact of agent i 's ideal point on agent i in group N . We assume that $\sum_{j \in N} \pi_i^N(j) = 1$. These impact or influence weights enter into agent i 's action in a linear manner.

$$p_i^N = \pi_i^N(i)v_i + \sum_{j \in N \setminus i} \pi_i^N(j)p_j^N \quad (1)$$

Equation 1 tells us that agent i 's action in group N is given as a convex combination of their ideal point and a weighted average of the other agents' actions. In the standard setup of the linear-in-means model, when $|X| = 2$, which we call the one dimension case, p_i^N is often interpreted as an effort level.³ Alternatively, instead of just splitting their time between leisure and labor (i.e. effort level), an agent can potentially split their time between leisure, labor, and volunteering. Our higher dimensional model allows us to capture this finer granularity in each agent's decision.⁴ More in line with classic stochastic choice, we can interpret each alternative $x \in X$ as some good, in which case p_i^N corresponds to the average good or bundle consumed by the agent in the context of group N . In this case, an agent can be thought of as having a preference over their choice frequencies rather than repeatedly maximizing (potentially different) static utility functions. This corresponds to an agent who is deliberately stochastic (Cerrei-Vioglio et al., 2019) or faces a perturbed utility function (Fudenberg et al., 2015) subject to social influence. Specifically, the choices described in Equation 1 arise from the group N playing a perfect information game where each agent's utility is given by the following.

$$u(p_i^N, p_{-i}^N) = -\pi_i^N(i) \sum_{x \in X} (p_i^N(x) - v_i(x))^2 - \sum_{j \in N \setminus i} \pi_i^N(j) \sum_{x \in X} (p_i^N(x) - p_j^N(x))^2 \quad (2)$$

This is in line with the observation made in Blume et al. (2015), Boucher and Fortin (2016), Kline and Tamer (2020), and Ushchev and Zenou (2020) that the one dimensional linear-in-means model arises from agents maximizing quadratic loss functions. Since we are working in a multidimensional setup, we make assumptions on our quadratic loss function which have no content in the one dimensional case. Notably, in the one dimensional case, the amount of weight agent i gives to agent j does not depend on the good that agent i is considering. However, in our setup, since we are considering multiple dimensions, the importance of agent j to agent i could in theory depend on the dimension or good in consideration. The linear-in-means model imposes that these influence parameters $\pi_i^N(j)$ are constant across each dimension. We also note that Golub and Morris (2020) offers an interpretation of the linear-in-means model as arising from an incomplete information game where, in Equation 1, v_i corresponds to agent i 's expectation of some underlying random variable and $p_j^N(j)$ corresponds to agent i 's expectation of agent j 's action.

We now discuss how group variation may be modeled through different assumptions on π_i^N . Intuitively, π^N is a matrix that captures weighted directed influence. By assuming that $\sum_{j \in N} \pi_i^N(j) = 1$, we are assuming that π^N is a stochastic matrix. When we observe group variation, each group N is

³As we will see later, the reflection problem of Manski (1993) is partially a result of the one dimension case being the standard case. We are able to provide conditions for identifications that generically do not hold in the one dimension case but generally hold in higher dimensions.

⁴In the same line of thought, we could partition leisure time and labor time into specific actions. In this case we would once again be able to capture this higher level of granularity.

subject to their own π^N . We consider three cases in which π^N and π^M are related between groups N and M .

1. **The General Model:** $\pi_i^N(j)$ is not assumed to satisfy any hypotheses across N except $\pi_i^N(i) > 0$.⁵
2. **The Luce Model:** For each $i, j \in \mathcal{A}$, there is $w_i(j)$ such that for all $i \in \mathcal{A}$, $w_i(i) > 0$, and $\sum_{j \in \mathcal{A}} w_i(j) = 1$ such that $\pi_i^N(j) = \frac{w_i(j)}{\sum_{k \in \mathcal{A}} w_i(k)}$.
3. **The Uniform Model:** For each $i, j \in N$, $\pi_i^N(j) = \frac{1}{|N|}$ (i.e. an unweighted average).

We interpret these three models in reverse order. In the case of the uniform linear-in-means model (ULM), an agent can be thought of as caring about each agent (including themselves) in group N equally. In this case, a sufficient statistic for agent i 's action is their ideal point v_i and the unweighted average of each other agent's action. As $|N|$ grows large, each agent only cares about matching their action to the average action in the rest of group N . This can be thought of as analogous to a Keynesian beauty contest (Keynes, 1937). In the case of the Luce linear-in-means model (LLM), each agent i can be thought of as having, for each agent j , an invariant importance weight $w_i(j)$. The relative importance of agent j to agent i in population N is simply the renormalization of each of these importance weights so that they sum to one. Finally, in the general linear-in-means model (GLM), the relative importance of each agent j to agent i is allowed to be fully context/group dependent. The interpretation here is that an agent j could act as a complement or substitute for agent k in the context of a group.⁶

Our goal in the next section is to study data that arises from these three cases of the linear-in-means model. Our focus is on testing these models and identification of v_i and π_N . With this in mind, we introduce our definition of consistency.

Definition 1. We say that a dataset $\{p^N\}_{N \in \mathcal{N}}$ is **consistent** with GLM/LLM/ULM if there exists $v_i \in \Delta(X)$ for each $i \in \mathcal{A}$ and π^N satisfying the conditions of GLM/LLM/ULM such that Equation 1 holds for each $N \in \mathcal{N}$ and $i \in N$.

These three models are nested: ULM \subset LLM \subset GLM.

⁵In Appendix C we consider the case when $\pi_i^N(i)$ is only assumed to satisfy non-negativity and extend all of our results to this case.

⁶Consider a setting where agent j has some noisy information about an underlying state. Agent i finds it important to match their action with agent j in order to coordinate with the underlying state. Now suppose that agent k is added to the group. Agent k has perfect information about the underlying state. In this case, the relative importance of agent j to agent i would go to zero when agent k is introduced as the action of agent k is a better signal for the underlying state.

3 Results

In this section, we study behavioral implications of the general, the Luce, and the uniform linear-in-mean models. We begin with the general case and provide two characterizations of datasets consistent with the general case as well as partial and point identification results for v_i and π^N . Our first characterization is via the non-empty intersection of a collection of convex sets. Our partial identification of v_i builds on this result. Our second characterization provides an existential linear program which fails to hold if and only if the dataset is consistent with the general linear-in-means model. We then proceed to the Luce and uniform linear-in-mean models and provide refinements of these results.

3.1 The General Model

We now begin our analysis of the general linear-in-means model. Recall that in GLM, for each group of agents N , each agent's choice p_i^N is written as a convex combination of v_i and p_j^N across all $j \in N \setminus i$. Further, there are no restrictions across groups about the weights assigned to each agent other than each agent i puts some weight on v_i . This means that testing GLM amounts to finding if there is some v_i which can induce, for all $N \in \mathcal{N}_i$, p_i^N as a convex combination of v_i and p_j^N across all $j \in N \setminus i$. This observation tells us two things. First, testing GLM can be done agent by agent. That is to say, whether or not agent i has a feasible rationalizing v_i can be tested independently of agent j . Second, a key part of testing GLM is finding the set of feasible v_i for agent i given p^N for each $N \in \mathcal{N}_i$.

With this in mind, suppose we observe data p^N and we see that agent i 's choices, p_i^N , lie on the interior of the convex hull of the other agents' choices, which we write $\text{int}\Delta(p_{-i}^N)$. In this case, no matter what $v_i \in \Delta(X)$ we consider, p_i^N can be written as a convex combination of the other agents' choices. Thus we can simply ask that agent i puts a vanishingly small amount of weight on v_i in their convex combination and rationalize this v_i . Now suppose we observe data such that p_i^N does not lie in $\text{int}\Delta(p_{-i}^N)$. In this case, a v_i is feasible if and only if p_i^N can be written as a convex combination of v_i and some convex combination of p_j^N for $j \in N \setminus i$. Thus the set of points on the "opposite side" of p_i^N from $\Delta(p_{-i}^N)$ correspond to the set of feasible v_i . Formally, this is given by the following equation.

$$\text{co}^{-1}(\Delta(p^N), p_i^N) = \{v \in \Delta(X) | v = \sum_{j \in N} \gamma_j p_j^N, \gamma_j \leq 0 \forall j \in N \setminus i, \sum_{j \in N} \gamma_j = 1\} \quad (3)$$

Figure 1 visualizes the set of feasible v_i for agent 1 when there are three goods and three agents. These two observations tell us that the testable content of GLM amounts to checking, for each i , if there is some feasible v_i in each $\text{co}^{-1}(\Delta(p^N), p_i^N)$.

Our prior observations tell us that testing GLM amounts to checking if, for each i , a collection of convex sets has a point of mutual intersection. Keeping in line with Samet (1998) and Morris (1994),

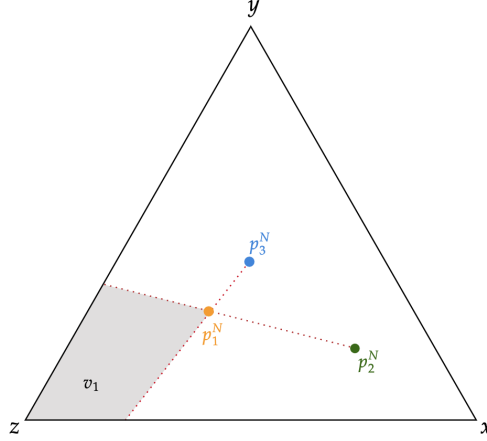


Figure 1: The figure considers three agents choosing when three alternatives are available. In the general linear-in-means model, p_1^N is an arbitrary convex combination of v_1 , p_2^N , and p_3^N . The shaded region corresponds to the set of feasible v_1 which induces p_1^N given p_2^N and p_3^N . This corresponds to $co^{-1}(\Delta(p^N), p_1^N)$ from Equation 3.

we can transform this condition into a no money-pump condition. With this in mind, we consider a setting where an outside observer is able to make bets on an agent and their choices in each group N .

Definition 2. A set of vectors $\{b^N\}_{N \in \mathcal{N}_i}$ with $b^N \in \mathbb{R}^X$ for each $N \in \mathcal{N}_i$ is called a **bet on agent i** .

We restrict attention to bets which we call feasible. Effectively, feasibility is a statement about initial investment and says that, once we aggregate across each group N in \mathcal{N}_i , placing a bet on agent i choosing alternative x should be ex-ante costly.

Definition 3. A bet on agent i is **strictly feasible** if $\sum_{N \in \mathcal{N}_i} b^N \ll 0$.⁷

However, for an outside observer to make a bet, it should be ex-post profitable to them. We ask that this individual rationality condition holds for each group $N \in \mathcal{N}_i$.

Definition 4. A bet on agent i is **individually rational** if $b^N \cdot p_i^N > 0$ for each $N \in \mathcal{N}_i$.

We impose one last condition on these bets. Notably, we have defined these bets as bets on a specific agent i . As such, we restrict to bets which are incentive compatible. That is to say, the outside observer cannot gain by placing the bet on agent j instead of agent i at any $N \in \mathcal{N}_i$.

⁷When working with vectors, we use $b \ll c$ to denote that the vector b is strictly less than c in each of its dimensions. We use $b < c$ to denote that vector b is weakly less than c in each of its dimensions and strictly less in at least one dimension. We use $b \leq c$ to denote that vector b is weakly less than c in each of its dimensions.

Definition 5. A bet on agent i is **incentive compatible** if $b^N \cdot (p_i^N - p_j^N) \geq 0$ for each $N \in \mathcal{N}_i$ and each $j \in N \setminus i$.

Finally, we say that there is no incentive compatible money-pump if there is no bet that satisfies these three conditions. This effectively asks that there is no way for an outside observer to guarantee that they make an expected profit using an incentive compatible bet.

Definition 6. We say that a dataset $\{p^N\}_{N \in \mathcal{N}}$ satisfies **no incentive compatible money pump** if for each $i \in \mathcal{A}$ there are no strictly feasible, individually rational, and incentive compatible bets on agent i .

We are now ready to state our characterization of GLM.

Theorem 1. For a dataset $\{p^N\}_{N \in \mathcal{N}}$, the following are equivalent.

1. $\{p^N\}_{N \in \mathcal{N}}$ is consistent with the general linear-in-means model.
2. For every $i \in \mathcal{A}$, the collection of sets $\{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ has a point of mutual intersection.
3. $\{p^N\}_{N \in \mathcal{N}}$ satisfies no incentive compatible money pump.

We leave all proofs to the appendix. The equivalence between (1) and (2) follows from our discussion at the start of Section 3.1. If $co^{-1}(\Delta(p^N), p_i^N)$ corresponds to the set of feasible v_i for agent i in group N , then there needs to be some v_i common to this set across all groups. The equivalence between (1) and (3) is partially a result of (2). As mentioned previously, since testing for GLM amounts to testing for a point of mutual intersection, we can transform this via linear programming duality to get our no money pump condition. However, we note that our no money pump condition does not follow immediately from (2) and an application of the result of Samet (1998) and Billot et al. (2000) as, in our proof, we work with polyhedral sets rather than compact sets. This variation allows us to recover the exact form of our no money pump condition. In Appendix D we discuss the relation and application of Samet (1998) to our Theorem 1. In this case, we get a type of no trade condition which characterizes GLM.

Before moving on, we note that condition (2) from Theorem 1 reduces to an easily checkable condition in the one dimensional case. In the one dimensional case, we consider two goods x and y . The choice probability of x is a sufficient statistic for the choice probability of y . In the one dimension case, we use p_i^N to denote the scalar which corresponds to the probability that agent i chooses x in group N . Now let $\mathcal{N}_i^- \subseteq \mathcal{N}_i$ denote the set of groups N satisfying $p_i^N \leq p_j^N$ for each $j \in N \setminus i$. Similarly, let $\mathcal{N}_i^+ \subseteq \mathcal{N}_i$ denote the set of groups N satisfying $p_i^N \geq p_j^N$ for each $j \in N \setminus i$.

Corollary 1. In the one dimension case, a dataset $\{p^N\}_{N \in \mathcal{N}}$ is consistent with the general linear-in-means model if and only if, for all $i \in \mathcal{A}$, the following conditions hold.

1. $\min_{N \in \mathcal{N}_i^-} p_i^N \geq \max_{N \in \mathcal{N}_i^+} p_i^N$ when both \mathcal{N}_i^- and \mathcal{N}_i^+ are non-empty,
2. $\{p_i^N\}_{N \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-}$ contains at most one value (which we denote p_i^-),
3. $\min_{N \in \mathcal{N}_i^-} p_i^N = p_i^- = \max_{N \in \mathcal{N}_i^+} p_i^N$ when $\mathcal{N}_i^- \cap \mathcal{N}_i^+$ is non-empty.

3.1.1 Identification

We now turn our attention to the identification properties of GLM. Going back to Manski (1993), it is known that recovering influence parameters is generally a hard problem. Here we provide tight partial identification bounds for each v_i and give conditions under which these v_i are point identified. When v_i is point identified, we are further able to give conditions under which each $\pi_i^N(j)$ is point identified. These correspond to the endogenous influence parameters in the context of Manski (1993).

Definition 7. For parameter v_i , we say that A is the **tight identified set** if $v \in A$ if and only if there exists π_i^N for each $N \in \mathcal{N}_i$ such that $p_i^N = \pi_i^N(i)v + \sum_{j \in \mathcal{N} \setminus i} \pi_i^N(j)p_j^N$.

Our first goal is to characterize the tight identified set of v_i . Recall our discussion of $co^{-1}(\Delta(p^N), p_i^N)$ prior to Theorem 1. $co^{-1}(\Delta(p^N), p_i^N)$ corresponds to the set of feasible v_i given observed choices p^N . Theorem 1 tells us that our dataset is consistent with GLM if and only if, for each $i \in \mathcal{A}$, $\{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ have a point of mutual intersection. This amounts to testing if there is some ideal point v_i which is feasible in each group $N \in \mathcal{N}_i$. It follows from similar logic that any point in the mutual intersection of $\{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ is a feasible value for v_i .

Proposition 1. In GLM, the tight identified set for v_i is given by $\bigcap_{N \in \mathcal{N}_i} co^{-1}(\Delta(p^N), p_i^N)$.

Proposition 1 formally states the observation from the prior paragraph. As mentioned earlier, part of our aim is to give conditions under which v_i is point identified. Here, point identified means that the tight identified set is a singleton. Our next result gives a sufficient condition for point identification of v_i . Let $\mathcal{N}_i^{ext} \subseteq \mathcal{N}_i$ denote the set of groups N with $p_i^N \notin \Delta(p_{-i}^N)$.

Corollary 2. Let $N_j = \{i, j\}$ and $N_k = \{i, k\}$. Suppose that $\{p^N\}_{N \in \mathcal{N}}$ is consistent with GLM, $N_j, N_k \in \mathcal{N}_i^{ext}$, and that the vectors $(p_i^{N_j} - p_j^{N_j})$ and $(p_i^{N_k} - p_k^{N_k})$ are linearly independent. Then v_i is point identified.

Figure 2 offers a visualization of Corollary 2. While we do not formally show it here, Corollary 2 naturally extends. Consider $co^{-1}(\Delta(p^N), p_i^N)$ for a group of agents N . Note that, in the case of $N \in \mathcal{N}_i^{ext}$, if we drop the non-negativity restriction on v in our definition of $co^{-1}(\Delta(p^N), p_i^N)$, then $co^{-1}(\Delta(p^N), p_i^N)$ defines a polyhedral cone. The extremal rays of this cone take the form $p_i^N - p_j^N$ plus some location translation. In our setting, linear independence of the extremal rays of two of

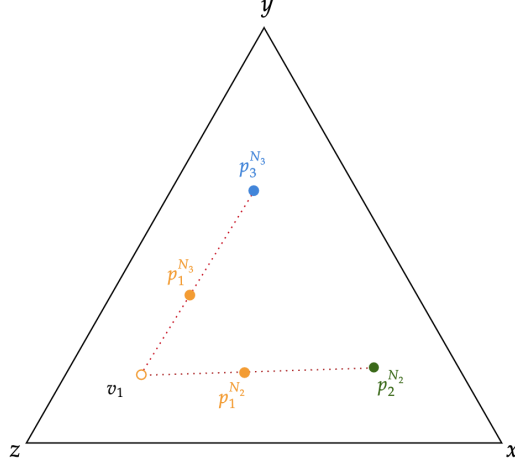


Figure 2: The figure considers choice by agents 1, 2, and 3. We observe choice in two groups; $N_2 = \{1, 2\}$ and $N_3 = \{1, 3\}$. Consider the two rays, one for each group. These rays intersect, so, by Theorem 1, this dataset is consistent with GLM. The linear independence of $(p_1^{N_2} - p_2^{N_2})$ and $(p_1^{N_3} - p_3^{N_3})$ gives us that these rays intersect at a single point. This single point corresponds to the uniquely feasible v_1 .

these n dimensional cones will make the intersection of these two sets $(n - 1)$ dimensions. Thus, when we observe n groups of n agents, with each group being contained in \mathcal{N}_i^{ext} , and the set of extremal rays of $co^{-1}(\Delta(p^N), p_i^N)$ across all n groups of n agents being linearly independent, $\bigcap_{N \in \mathcal{N}_i^{ext}} \{co^{-1}(\Delta(p^N), p_i^N)\}$ is a $(n - n) = 0$ dimensional set. By Theorem 1 this set is non-empty and thus we get point identification. We now observe that identification of v_i is a key step in identifying the influence parameter π_i^N .

Proposition 2. *Suppose that v_i is point identified. Then π_i^N is point identified if the set of vectors including v_i and $\{p_j^N\}_{j \in N \setminus i}$ is linearly independent.*

The identification of π_i^N is pinned to the identification of v_i . Fixing $\{p_j^N\}_{j \in N \setminus i}$ and maintaining linear independence, each potential value of v_i induces a different value for π_i^N . Proposition 2 is informative about the reflection problem of Manski (1993) in our context. Our condition for identification is the linear independence of v_i and $\{p_j^N\}$. In the one dimension case, both v_i and p_j^N are scalars and thus cannot be linearly independent. We can still recover $\pi_i^N(j)$ when $N = \{i, j\}$ and v_i is point identified. However, v_i is point identified only if $\min_{N \in \mathcal{N}_i^-} p_i^N = \max_{N \in \mathcal{N}_i^+} p_i^N$ when both the min and max exist or when \mathcal{N}_i^- is non-empty. This tells us that, generically, we are unable to recover π_i^N in the one dimension case. When we return to the higher dimension case, Proposition 2 gives restrictions on the size of a group N for when we can identify π_i^N . Since Proposition 2 asks for the linear independence of $|N|$ vectors, we can only get linear independence if we are working in $\mathbb{R}^{|N|}$ or higher dimensions. In terms of group size and X , this means we can potentially have linear independence if and only if

$|N| \leq |X|$. Returning to the reflection problem, Proposition 2 and our prior discussion tell us that, in our setting, the identification of social influence parameters is simply a problem of dimension.

3.2 Luce Linear-in-Means

As we just saw, there are conditions in GLM under which we are able to recover both an agent's ideal point v_i and their social interaction parameters π_i^N for groups that we observe. However, these identified parameters have little predictive power for groups we do not observe in GLM. Since GLM allows arbitrary variation in π_i^N across groups, the most we can predict for agent i 's choice in an unobserved group N is that it lies within the convex hull of $\{v_j\}_{j \in N}$. Our goal in this section is to consider the Luce linear-in-means model which allows us to connect our social interaction parameters across groups. LLM allows analysts to predict choices in unobserved groups subject to the conditions for identification from Section 3.1.1 holding.

Recall that LLM supposes that each agent i has a weighting function, $w_i(j)$, which corresponds to the absolute importance of agent j to agent i . In a group N , the relative importance of agent j to agent i is given by the renormalization of this weighting function, $\pi_i^N(j) = \frac{w_i(j)}{\sum_{k \in N} w_i(k)}$. The interpretation here is that there are no second order interaction effects between agents j and k which impact their relative importance to agent i . We first discuss the identification and predictive properties of LLM. We focus on the case when v_i is known. For a specific group N , if the conditions of Proposition 2 hold, we can pin down π_i^N . This allows us to pin down the relative weights of each agent $j \in N \setminus i$. That is to say $\frac{w_i(j)}{w_i(k)} = \frac{\pi_i^N(j)}{\pi_i^N(k)}$. This is the exact condition for identification that is used in the Luce model of Luce (1959). This means that if $N = \mathcal{A}$, we can predict choice (for agent i) on every possible group of agents. However, we may not always observe choice on $N = \mathcal{A}$ or if we do the conditions of Proposition 2 may not be satisfied at N . To regain full predictive power in LLM, we need the conditions of Proposition 2 to hold for a collection of groups $\{N\}_l^L$ such that every two agents k and j can be compared through a string of groups.

Definition 8. Suppose v_i is point identified. Agents j and k are **comparable by agent i** if there exist two observed groups $N_1, N_2 \in \mathcal{N}$ such that the following hold.

- $i \in N_1 \cap N_2$,
- $j \in N_1$ and $k \in N_2$,
- For each $l \in \{1, 2\}$, v_i and $\{p_m^{N_l}\}_{m \in N_l \setminus i}$ are linearly independent.

Proposition 3. Suppose that v_i is point identified. $w_i(j)$ is point identified for all $j \in \mathcal{A}$ if every pair of agents $j \neq k$ are comparable by agent i .

Proposition 3 is an immediate result of our Proposition 2, Corollary 4 of Alós-Ferrer and Mihm (2024), and the fact that i is in both N_1 and N_2 . Further, Proposition 3 gives conditions under which we can predict the choice of agent i in any group N , conditional on knowing v_j for each $j \in N$. While we focus on LLM to recover predictive power in the linear-in-means model, one could instead consider any known mapping from π_i^N to π_i^M for two different groups N and M to recover predictive power. We focus on LLM for two reasons. First, LLM is an intuitive criterion that captures the case of no higher order interactions. Second, as we will now discuss, we can test LLM using a natural extension of our no incentive compatible money pump condition.

Recall our story about an outside observer making bets on an agent and their choices. In the case of GLM, we asked that these bets were strictly feasible, individually rational for each group containing i , and incentive compatible for each group containing i . When moving from GLM to LLM, we are moving to a model where the social influence parameters π_i^N and π_i^M are actually connected across groups. Our condition for testing LLM extends the no incentive compatible money pump condition taking this across group connection into account. Specifically, we weaken incentive compatibility so that the outside observer cannot gain by placing a bet on j instead of i across every $N \in N_i \cap N_j$.

Definition 9. A bet on agent i is **weakly incentive compatible** if, for each $j \neq i$, $\sum_{N \in N_i \cap N_j} b^N \cdot (p_i^N - p_j^N) \geq 0$.

Definition 10. We say that a dataset $\{p^N\}_{N \in \mathcal{N}}$ satisfies **no weakly incentive compatible money pump** if for each $i \in \mathcal{A}$ there are no strictly feasible, individually rational, and weakly incentive compatible bets on agent i .

Theorem 2. For a dataset $\{p^N\}_{N \in \mathcal{N}}$, the following are equivalent.

1. $\{p^N\}_{N \in \mathcal{N}}$ is consistent with the Luce linear-in-means model.
2. $\{p^N\}_{N \in \mathcal{N}}$ satisfies no weakly incentive compatible money pump.

Observe that if our dataset satisfies no weakly incentive compatible money pump, then it satisfies no incentive compatible money pump. It then follows from Theorem 1 that the collection of sets $\{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ have a point of mutual intersection. In the case of GLM, this is the set of feasible v_i . The strengthening of incentive compatibility to weak incentive compatibility is exactly what guarantees us that the social influence parameters π_i^N follow a Luce rule across groups.

Figure 3 compares GLM and LLM in terms of their prediction power on the Marschak-Machina triangle in a domain of three alternatives, $X = \{x, y, z\}$ and three agents $\{1, 2, 3\}$. This is the simplest environment in which we can illustrate the differences between these models. In the figure, stochastic choices for three binary groups ($\{1, 2\}$, $\{1, 3\}$ and $\{2, 3\}$) are fixed and denoted by different colored

dots on the edges of the dotted triangle. The blue dots indicate agent 3's choices, $p_3^{\{1,3\}}$ and $p_3^{\{2,3\}}$. We then illustrate the predictions for $p_3^{\{1,2,3\}}$ for GLM and LLM. The blue-shaded triangle identifies all possibilities for $p_3^{\{1,2,3\}}$ for GLM given these binary group choices. This figure illustrates both the predictive and explanatory power of GLM: If $p_3^{\{1,2,3\}}$ lies outside this triangle, the data cannot be explained by GLM.

LLM restricts choices for $p_3^{\{1,2,3\}}$ even further. Indeed, LLM predicts that there is only a single possibility for $p_3^{\{1,2,3\}}$ indicated by a blue dot inside GLM's prediction. If $p_3^{\{1,2,3\}}$ is not equal to this point, then the data cannot be explained by the Luce linear-in-means model. Note that the prediction of LLM belongs to the shaded triangle indicating that LLM is a special case of GLM.

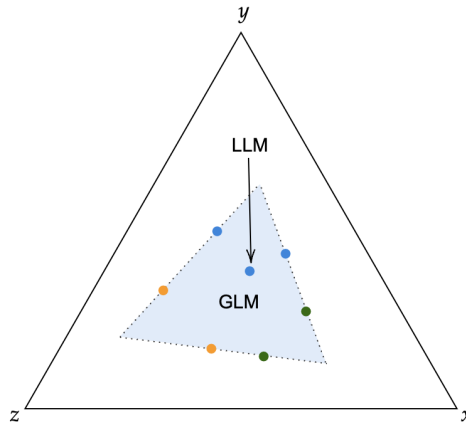


Figure 3: This figure illustrates predictions of GLM and LLM for three alternatives, $X = \{x, y, z\}$ and three agents: blue, green, and orange. The blue-shaded triangle identifies all possibilities for the choice probability of the blue agent consistent with GLM in the entire group, given three binary group choices for each agent depicted on three edges of the shaded region. Note that LLM predicts uniquely the choice probability of the blue agent.

3.3 Uniform Model

In the prior sections, we considered the linear-in-means model allowing for heterogeneous social interaction terms. The interpretation of the heterogeneity is that the importance of agent j and agent k may differ to agent i . In this section, we consider the hypothesis that agent i belonging to a group impacts their behavior but no one agent in that group is any more important than the other. We model this hypothesis through the uniform linear-in-means model. This is a special case of LLM when $w_i(j) = w_i(k)$ for each $j, k \in \mathcal{A}$. In this section, we maintain the assumption that each observed group has the same size, for all $N, M \in \mathcal{N}$, $|N| = |M|$. This is done so we can focus on the empirical content

of variation of group make-up rather than group size. We relax this assumption in Appendix E where we give all the axioms and results from this section allowing for group size variation.

In the context of a linear influence model, such as ours, $p_i^N - p_j^N$ corresponds to a (rescaling) of the influence agent j has on agent i 's choice. As such, we call $p_i^N - p_j^N$ the **peer effect** of agent j on agent i in group N . All of our axioms for ULM are stated in terms of peer effects. Before stating our axioms, we need one definition.

Definition 11. A cycle is a sequence of tuples $\{(i_k, j_k, N_k)\}_{k=1}^K$ such that $i_k, j_k \in N_k$, $j_k = i_{k+1}$, and $j_K = i_1$.

A cycle captures cycles of influence. Agent i_1 influences agent j_1 in group N_1 and then agent $j_1 = i_2$ influences agent j_2 in group N_2 . This is repeated until we return to agent i_1 . Our first axiom puts restrictions on the sum of peer effects across cycles.

Axiom 1 (Cyclically constant). A dataset $\{p_N\}_{N \in \mathcal{N}}$ is cyclically constant if for every cycle $\{(i_k, j_k, N_k)\}_{k=1}^K$ we have that

$$\sum_{k=1}^K (p_{i_k}^{N_k} - p_{j_k}^{N_k}) = 0. \quad (4)$$

To best understand cyclic constancy, we first introduce a second axiom implied by cyclic constancy.

Axiom 2 (Constant Peer Effects). A dataset $\{p_N\}_{N \in \mathcal{N}}$ satisfies constant marginal effects if for each N and M with $i, j \in N \cap M$ we have that $[p_i^N - p_j^N] = [p_i^M - p_j^M]$.

Observe that constant peer effects is implied by cyclic constancy when we consider cycles of length two. Constant peer effects tells us that the peer effect of agent j on agent i is group invariant. Returning to cyclic constancy, it says that the peer effect of agent i on agent j plus the peer effect of agent j on agent k should be equal to the peer effect of agent i on agent k (and so on for longer cycles). Further, there is no way to break this equality by going to different groups during a cycle. As such, cyclic constancy tells us that agent i to agent j peer effects are group invariant and that a long chain of peer effects corresponding to the indirect peer effect of agent i on agent j equals the direct peer effect of agent i on agent j . Our next axiom puts restrictions on the peer effect of agent i across different groups.

Axiom 3 (Symmetric Peer Effects). For $i \in N \cap M$,

$$p_i^N - \sum_{j \in N \setminus M} [p_j^N - p_i^N] = p_i^M - \sum_{k \in M \setminus N} [p_k^M - p_i^M]. \quad (5)$$

Symmetric peer effects tell us that, once we take agent i 's actual choice into account, the total peer effect of agent i on agents in group N which are not in M is equal to the total peer effect of agent i on agents in group M which are not in group N . The combination of cyclic constancy and symmetric peer effects tells us that the total peer effect of agent i is the same in groups N and M , once we take into account their actual choice in each group. Our last axiom restricts the total amount of peer effect that agent i has in a group.

Axiom 4 (Bounded Total Peer Effects). For $i \in N$,

$$p_i^N \geq \sum_{j \in N \setminus i} [p_j^N - p_i^N]. \quad (6)$$

Bounded total peer effects tell us that the total amount of peer effect agent i has in group N can be no more than their actual choice p_i^N . With this in mind, we are now ready to give our characterization of ULM. Recall that this result assumes that $|N| = |M|$ and that this assumption is relaxed in the appendix.

Theorem 3. For a dataset $\{p^N\}_{N \in \mathcal{N}}$, the following are equivalent.

1. $\{p^N\}_{N \in \mathcal{N}}$ is consistent with the uniform linear-in-means model.
2. $\{p^N\}_{N \in \mathcal{N}}$ satisfies cyclic constancy, symmetric peer effects, and bounded total peer effects.

We first note that, in Theorem 3, we can replace cyclic constancy with constant peer effects and the equivalence holds. Our focus on cyclic constancy is due to the following discussion. Consider the following the equation.

$$p_i^N = \frac{1}{|N|} \hat{v}_i + \sum_{j \in N \setminus i} \frac{1}{|N|} p_j^N + O^N \quad (7)$$

In Equation 7, \hat{v}_i is not restricted to have non-negative elements (but still satisfies $\sum_{x \in X} v_i(x) = 1$) and we ask $\sum_{x \in X} O^N(x) = 0$. Choice induced by Equation 7 differs from ULM in two important ways. First, an agent's ideal point \hat{v}_i no longer needs to lie within the simplex. Second, every agent i in group N is subject to some group specific shock to tastes given by O^N . Under one additional assumption on \mathcal{N} , we show in Appendix E that choice according to Equation 7 is characterized by cyclic constancy. The addition of symmetric peer effects is exactly what rules out group specific shocks, reducing O^N to be the zero vector across all $N \in \mathcal{N}$. Finally, bounded total peer effects is what induces each agent's ideal point to be within the simplex.

We conclude our analysis with the observation that ULM is trivially point identified. Since each agent’s social interaction parameters are pinned down to be $\frac{1}{|N|}$, the only parameter left to identify is each agent’s ideal point. This can be easily recovered from the following equation.

$$v_i = |N|p_i^N - \sum_{j \in N \setminus i} p_j^N \quad (8)$$

Equation 8 also offers an alternative characterization of ULM. A dataset is consistent with ULM if and only if the value of the right hand side of Equation 8 is group invariant and lies in the simplex.

4 Discussion

In this paper we study the testing and identification properties of the linear-in-means model when an analyst observes group variation. We first study the linear-in-means model in its most general setup, allowing for arbitrary variation of social influence parameters across groups. In this setting, we show that the linear-in-means model can be tested by checking if a group of convex sets has a point of mutual intersection. We also provide a linear programming formulation of this test. Further, we are able to provide a series of linear independence based conditions under which an agent’s social influence parameters are point identified, thus identifying the underlying network structure of the group. As part of our analysis, we show that the reflection problem of Manski (1993) is a generic problem when each agent’s decision space is one dimension but stops being generic when we move to higher dimensions. We then consider the Luce linear-in-means model as a way to recover the predictive power not present in the general linear-in-means model. Finally, we study the uniform linear-in-means model and provide normative axioms which characterize group effects when an agent cares about each agent’s choices equally.

Remark 1 (Network Variation). *Throughout the course of this paper our focus has been on group variation. We could instead consider network variation. That is to say, fix a group N and consider two realizations of some observable variable or regime, R and R' . Suppose that an agent’s social influence parameters are given by π_i when R is observed and π'_i when R' is observed. Neither π_i nor π'_i are observable to the analyst. If no restrictions are placed on π_i or π'_i , testing for the linear-in-means model with network variation can be done using our Theorem 1. Further, all of the identification results of Section 3.1.1 also extend to this arbitrary network variation setup.*

Remark 2 (Characteristic Variation and Exogenous Effects). *In our setup, we do not assume observation of any characteristics other than agent identity. Part of the reflection problem of Manski (1993) is disentangling exogenous and endogenous group effects. If we add characteristic variation, we can distinguish*

between exogenous and endogenous effects in a two step procedure. In the first step, we proceed with the identification arguments from Section 3.1.1. Upon doing so, we have recovered v_i for each agent. In the second step, we treat v_i as the outcome variable and proceed with standard analysis of the relationship between v_i and observed characteristics. Notably, our first stage captures the endogenous effects of the reflection problem and the second stage captures the exogenous effects.

Remark 3 (Product-Attribute Variation). Another variation that could be introduced in our framework is product attribute variations. Suppose that there is a set of K observable attributes, with a_x denoting the vector of attributes for product x . a_x includes not only things that affect product quality, but also things like price, advertising, etc. We assume that the ideal point of each agent is a linear function of observable product attributes. This is captured by an agent-specific real vector β_i such that $u_i(x) = \beta_i a_x$ for each i and x . This means that we assume that attributes affect utility in the same way for all products. Then the ideal point is calculated as $v_i(x) = \frac{u_i(x)}{\sum_{y \in X} u_i(y)}$. Identification of β_i then proceeds with a two-step procedure. As above, we first recover v_i for each agent. In the second step, we treat v_i as the outcome variable and proceed with standard analysis of the relationship between v_i and observed attributes.

4.1 Related Literature

Our paper is related to several strands of literature. We begin by discussing the strand which studies the linear-in-means model of social interactions. Predating the modern literature on the linear-in-means model, Keynes (1937) considers a model of financial markets via a story of beauty contests. In this setting an agent wishes to take the action that coincides with the average action of the rest of the population. In our setup, this corresponds to $\pi_i^N(i) = 0$ and $\pi_i^N(j) = \frac{1}{|N|-1}$. More recently, Ushchev and Zenou (2020) studies the microfoundations and comparative statics of the linear-in-means model allowing for arbitrary network structure. As mentioned earlier, Ushchev and Zenou (2020) along with Blume et al. (2015), Boucher and Fortin (2016), and Kline and Tamer (2020) show that the linear-in-means choice rule can be achieved as the best response to a quadratic loss utility functions in a perfect information game where each agent knows each v_i and π_i^N . Golub and Morris (2020) consider an extension of this setup where agents have incomplete information and relate the linear-in-means model to higher-order expectations as well as conventions in networks. We build on this literature by studying the empirical content of the linear-in-means model.

More closely related to our paper is the strand of literature which focuses studying the identification properties of the linear-in-means model as well as other related models of peer effects. A seminal contribution in this literature is Manski (1993) and his discussion of the reflection problem. Important to our analysis is the following takeaway from the reflection problem of Manski (1993). It is in general difficult to identify the social influence of a group on an agent due to the endogenous nature

of outcomes. In our setting, this corresponds to identifying both the underlying network structure and the corresponding weights on directed edges in this network. Much of the literature following Manski (1993) aims to identify social interaction parameters when the underlying network structure is (partially) known. This literature is extensive, so we list the following, all of which provide various conditions in order to recover identification of social interaction parameters in linear-in-means style models; Graham (2008), Bramoullé et al. (2009), De Giorgi et al. (2010), Blume et al. (2011), Blume et al. (2015), De Paula (2017). More recently, Boucher et al. (2024) extends the linear-in-means to a CES in means model of social influence and provides identifications in their setting. Perhaps most closely related to our analysis within this literature is the work of De Paula et al. (2024). In their analysis they consider both the identification of social interaction parameters as well as the identification of the underlying network structure. A common tool for identification in each of the previously mentioned studies is the use of characteristic variation in order to recover identification. Our analysis differs from this literature in that we use group variation rather than characteristic variation to recover identification. A second difference in our analysis is that our outcome vector p_i^N has across dimension restrictions. While a standard assumption in stochastic choice and discrete choice analysis, the assumption that p_i^N lies in a multi dimensional simplex seems underutilized in the literature on identification of the linear-in-means model.

Our paper also contributes to the choice theoretic strand of literature studying social interactions. To our knowledge, this literature begins with Cuhadaroglu (2017) who studies a two period model with social influence taking effect in the second period. Borah et al. (2018) also considers a two period model of social influence. In the first period, social influence is used to form an agent’s consideration set and the second period is used for choice. Kashaev et al. (2023) consider a model of social influence where the choices of an agent’s peers directly form their consideration set. One of the main goals of this work is the actual identification of underlying social parameters, which they are able to achieve through dynamic choice variation. Bhushan et al. (2023) considers a model of choice where agents influence each other through their beliefs. An agent’s beliefs are determined through a process similar to ULM. Choices then correspond to subjective expected utility given these beliefs. Most closely related to our work in this literature is the work of Chambers et al. (2023) who consider a version of the linear-in-means model. They focus on a setting with menu variation (i.e. variation of X) with a fixed group and network structure. This differs from our analysis in that our focus is on group variation and we can accommodate network variation.

More generally, our paper is related to the literature studying the empirical content of strategic settings. Part of this literature focuses on testing the empirical content of specific solution concepts. Sprumont (2000) studies the testable content of Nash equilibrium. The work of Haile et al. (2008) finds that quantal response equilibrium has no empirical content. Similarly, Bossert and Sprumont

(2013) and Rehbeck (2014) find that backwards induction has no empirical content when we only observe the induced choice function. Another portion of this literature focuses on characterizing the empirical content of specific (types of) games. To begin, Lee (2012) characterizes the testable content of zero-sum games. Carvajal et al. (2013) provides a revealed preference style characterization of the Cournot model of competition. Finally, Lazzati et al. (2023) characterizes the empirical content of Nash equilibrium in monotone games. As mentioned earlier, the linear-in-means model can be thought of as arising from Nash equilibrium play where each agent's utility is given by Equation 2. As such, we characterize Nash equilibrium play in the corresponding game when we observe player variation.

Finally, our paper is also related to the literature in stochastic choice studying agents who have preferences for non-deterministic bundles. The idea of deliberately stochastic preferences goes back to at least Machina (1985). Recently, Cerreia-Vioglio et al. (2019) axiomatizes data that arises from agents choosing with deterministic preferences over lotteries. Similarly, Fudenberg et al. (2015) characterizes stochastic choice data that arises from agents who have cardinal preferences over each alternative but face a perturbation to their utility function within the simplex. Allen and Rehbeck (2019) studies the identification properties of a similar class of perturbed utility functions. There has been little work on these types of perturbed utility functions with social influence components. Hashidate and Yoshihara (2023) considers a perturbed utility function where the perturbation corresponds to a norm. To reiterate one final time, the choices in the linear-in-means model corresponds to a perturbed utility model where an agent's base utility corresponds to a quadratic loss function with reference to their ideal point. The perturbation then corresponds to a sum of quadratic loss functions with each one referencing another agent's choice.

References

- AFRIAT, S. N. (1967): “The construction of utility functions from expenditure data,” *International economic review*, 8, 67–77.
- AGRANOV, M., M. ELLIOTT, AND P. ORTOLEVA (2021): “The importance of Social Norms against Strategic Effects: The case of Covid-19 vaccine uptake,” *Economics Letters*, 206, 109979.
- ALLEN, R. AND J. REHBECK (2019): “Identification with additively separable heterogeneity,” *Econometrica*, 87, 1021–1054.
- ALÓS-FERRER, C. AND M. MIHM (2024): “A Characterization of the Luce Choice Rule for an Arbitrary Collection of Menus,” *Journal of Economic Theory*, 105941.
- BASSE, G., P. DING, A. FELLER, AND P. TOULIS (2024): “Randomization tests for peer effects in group formation experiments,” *Econometrica*, 92, 567–590.
- BHUSHAN, P., A. BORAH, AND S. VENKATESH (2023): “Beliefs in echo chambers: Identification and characterization,” .
- BILLOT, A., A. CHATEAUNEUF, I. GILBOA, AND J.-M. TALLON (2000): “Sharing Beliefs: Between Agreeing and Disagreeing,” *Econometrica*, 68, 685–694.
- BLUME, L. E., W. A. BROCK, S. N. DURLAUF, AND Y. M. IOANNIDES (2011): “Identification of Social Interactions,” in *Handbook of Social Economics*, Elsevier, vol. 1, 853–964.
- BLUME, L. E., W. A. BROCK, S. N. DURLAUF, AND R. JAYARAMAN (2015): “Linear Social Interactions Models,” *Journal of Political Economy*, 123, 444–496.
- BORAH, A., C. KOPS, ET AL. (2018): “Choice via Social Influence,” Unpublished.
- BOSSERT, W. AND Y. SPRUMONT (2013): “Every choice function is backwards-induction rationalizable,” *Econometrica*, 81, 2521–2534.
- BOUCHER, V. AND B. FORTIN (2016): “Some challenges in the empirics of the effects of networks,” *Handbook on the Economics of Networks*, 45, 48.
- BOUCHER, V., M. RENDALL, P. USHCHEV, AND Y. ZENOU (2024): “Toward a general theory of peer effects,” *Econometrica*, 92, 543–565.
- BRAMOULLÉ, Y., H. DJEBBARI, AND B. FORTIN (2009): “Identification of peer effects through social networks,” *Journal of econometrics*, 150, 41–55.

- CARD, D. AND L. GIULIANO (2013): “Peer Effects and Multiple Equilibria in the Risky Behavior of Friends,” *Review of Economics and Statistics*, 95, 1130–1149.
- CARVAJAL, A., R. DEB, J. FENSKE, AND J. K.-H. QUAH (2013): “Revealed Preference Tests of the Cournot Model,” *Econometrica*, 81, 2351–2379.
- CERREIA-VIOGLIO, S., D. DILLENBERGER, P. ORTOLEVA, AND G. RIELLA (2019): “Deliberately Stochastic,” *American Economic Review*, 109, 2425–45.
- CHAMBERS, C. P., T. CUHADAROGLU, AND Y. MASATLIOGLU (2023): “Behavioral influence,” *Journal of the European Economic Association*, 21, 135–166.
- CUHADAROGLU, T. (2017): “Choosing on Influence,” *Theoretical Economics*, 12, 477–492.
- DE GIORGI, G., M. PELLIZZARI, AND S. REDAELLI (2010): “Identification of social interactions through partially overlapping peer groups,” *American Economic Journal: Applied Economics*, 2, 241–275.
- DE PAULA, A. (2017): “Econometrics of Network Models,” in *Advances in Economics and Econometrics: Theory and Applications: Eleventh World Congress*, Cambridge University Press Cambridge, vol. 1, 268–323.
- DE PAULA, A., I. RASUL, AND P. C. SOUZA (2024): “Identifying network ties from panel data: Theory and an application to tax competition,” *Review of Economic Studies*, rdae088.
- FORTIN, B., G. LACROIX, AND M.-C. VILLEVAL (2007): “Tax evasion and social interactions,” *Journal of Public Economics*, 91, 2089–2112.
- FUDENBERG, D., R. IJIMA, AND T. STRZALECKI (2015): “Stochastic choice and revealed perturbed utility,” *Econometrica*, 83, 2371–2409.
- GILBOA, I. AND D. SCHMEIDLER (1989): “Maxmin expected utility with non-unique prior,” *Journal of mathematical economics*, 18, 141–153.
- GIRSANOV, I. V. (1972): *Lectures on Mathematical Theory of Extremum Problems*, vol. 67, Springer Science & Business Media.
- GOLUB, B. AND S. MORRIS (2020): “Expectations, networks, and conventions,” *arXiv preprint arXiv:2009.13802*.
- GRAHAM, B. S. (2008): “Identifying social interactions through conditional variance restrictions,” *Econometrica*, 76, 643–660.
- HAILE, P. A., A. HORTAÇSU, AND G. KOSENOK (2008): “On the empirical content of quantal response equilibrium,” *American Economic Review*, 98, 180–200.

- HASHIDATE, Y. AND K. YOSHIHARA (2023): “Social preferences and deliberately stochastic behavior,” *arXiv preprint arXiv:2304.14977*.
- KASHAEV, N., N. LAZZATI, AND R. XIAO (2023): “Peer Effects in Consideration and Preferences,” *arXiv preprint arXiv:2310.12272*.
- KEYNES, J. M. (1937): “The general theory of employment,” *The quarterly journal of economics*, 51, 209–223.
- KLINE, B. AND E. TAMER (2020): “Econometric analysis of models with social interactions,” in *The Econometric Analysis of Network Data*, Elsevier, 149–181.
- KREMER, M. AND D. LEVY (2008): “Peer Effects and Alcohol Use among College Students,” *Journal of Economic Perspectives*, 22, 189–206.
- LAZZATI, N., J. K.-H. QUAH, AND K. SHIRAI (2023): “An ordinal approach to the empirical analysis of games with monotone best responses,” .
- LEE, S. (2012): “The Testable Implications of Zero-sum Games,” *Journal of Mathematical Economics*, 48, 39–46.
- LUCE, R. D. (1959): *Individual Choice Behavior*, John Wiley.
- MACHINA, M. J. (1985): “Stochastic Choice Functions Generated From Deterministic Preferences Over Lotteries,” *The Economic Journal*, 95, 575–594.
- MANGASARIAN, O. L. (1994): *Nonlinear programming*, SIAM.
- MANSKI, C. F. (1993): “Identification of Endogenous Social Effects: The Reflection Problem,” *The Review of Economic Studies*, 60, 531–542.
- MAS, A. AND E. MORETTI (2009): “Peers at Work,” *American Economic Review*, 99, 112–45.
- MORRIS, S. (1994): “Trade with heterogeneous prior beliefs and asymmetric information,” *Econometrica: Journal of the Econometric Society*, 1327–1347.
- REHBECK, J. (2014): “Every choice correspondence is backwards-induction rationalizable,” *Games and Economic Behavior*, 88, 207–210.
- ROCKAFELLAR, R. T. (1970): *Convex Analysis*, Princeton.
- SACERDOTE, B. (2001): “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” *The Quarterly Journal of Economics*, 116, 681–704.

- (2014): “Experimental and quasi-experimental analysis of peer effects: two steps forward?” *Annual Review of Economics*, 6, 253–272.
- SAMET, D. (1998): “Common priors and separation of convex sets,” *Games and Economic Behavior*, 24, 172–174.
- SMALE, S. (1974): “Sufficient Conditions for an Optimum,” in *Dynamical Systems, Warwick: 1974, Lecture Notes in Mathematics*, ed. by A. Manning, Springer.
- SPRUMONT, Y. (2000): “On the Testable Implications of Collective Choice Theories,” *Journal of Economic Theory*, 93, 205–232.
- USHCHEV, P. AND Y. ZENOU (2020): “Social norms in networks,” *Journal of Economic Theory*, 185, 104969.
- WAN, Y.-H. (1975): “On Local Pareto Optima,” *Journal of Mathematical Economics*, 2, 35–42.

APPENDIX

A Preliminary Results

We begin with a preliminary lemma that formalizes our discussion at the start of Section 3.1.

Lemma A.1. $p_i^N = \gamma_i v_i + \sum_{j \in N \setminus i} \gamma_j p_j^N$ for $v_i \in \Delta(X)$ with $\gamma_j \geq 0$, $\gamma_i > 0$, and $\sum_{j \in N} \gamma_j = 1$ if and only if $v_i \in \text{co}^{-1}(\Delta(p^N), p_i^N)$.

Proof. Suppose that $p_i = \gamma_i v_i + \sum_{j \in N \setminus i} \gamma_j p_j^N$ for $v_i \in \Delta(X)$ with $\gamma_j \geq 0$ and $\sum_{j \in N} \gamma_j = 1$. Since $0 < \gamma_i \leq 1$, we can recover that $v_i = \frac{1}{\gamma_i} p_i^N + \sum_{j \in N \setminus i} \frac{-\gamma_j}{\gamma_i} p_j^N$. Since $\gamma_j \geq 0$ and $\sum_{j \in N} \gamma_j = 1$, $\frac{-\gamma_j}{\gamma_i} \leq 0$ and $\frac{1}{\gamma_i} + \sum_{j \in N \setminus i} \frac{-\gamma_j}{\gamma_i} = 1$ and so $v_i \in \text{co}^{-1}(\Delta(p^N), p_i^N)$.

Now suppose that $v_i \in \text{co}^{-1}(\Delta(p^N), p_i^N)$. This tells us that $v_i = \sum_{j \in N} \gamma_j p_j^N$ with $\gamma_j \leq 0 \forall j \in N \setminus i$ and $\sum_{j \in N} \gamma_j = 1$. It follows that $\gamma_i \geq 1$. By basic algebra, we can solve for p_i and get $p_i = \frac{1}{\gamma_i} v_i + \sum_{j \in N \setminus i} \frac{-\gamma_j}{\gamma_i} p_j^N$. Since $\gamma_j \geq 0$ for $j \in N \setminus i$ and $\sum_{j \in N} \gamma_j = 1$, $\frac{-\gamma_j}{\gamma_i} \geq 0$ and $\frac{1}{\gamma_i} + \sum_{j \in N \setminus i} \frac{-\gamma_j}{\gamma_i} = 1$ and so $p_i = \pi_i v_i + \sum_{j \in N \setminus i} \pi_j p_j^N$ for $v_i \in \Delta(X)$ with $\pi_j \geq 0$, $\pi_i > 0$, and $\sum_{j \in N} \pi_j = 1$. \square

Define $X_1 = \{x \in \mathbb{R}^n : \sum_i x_i = 1\}$. So X_1 is not the simplex but rather the hyperplane containing the simplex. The following very simple result is closely related to classical results, for example Girsanov (1972), Lemma 3.11, or characterizations of Pareto optimality, relating to Negishi weights e.g. Smale (1974); Wan (1975) but is written to be close in form to Billot et al. (2000); Samet (1998). The main difference from the preceding pair of papers is that no compactness assumptions are claimed, but rather polyhedrality of the sets is required.

Theorem 4. Let A_1, \dots, A_k each be nonempty subsets of X_1 , and suppose additionally that each A_i is a polyhedron. Suppose that $\bigcap_{i=1}^k A_i = \emptyset$. Then the following are equivalent:

1. $\bigcap_{i=1}^k A_i = \emptyset$.
2. For each $i = 1, \dots, k$, there are $p_i \in \mathbb{R}^n$ for which $\sum_{i=1}^k p_i = 0$ and such that $p_i \cdot x_i > 0$ for all $x_i \in A_i$.

Proof. We show that 1 implies 2. Observe that X_1 is itself defined by a finite set of linear inequalities. We now consider $Y = \{(x, \dots, x) : x \in X_1\} \subseteq \mathbb{R}^{nk}$ and $\prod_{i=1}^k A_i$. These sets are both clearly polyhedra. Further, by hypothesis, $Y \cap \prod_{i=1}^k A_i = \emptyset$. By Corollary 19.3.3 and Theorem 11.1 of Rockafellar (1970), there exists $(p_1, \dots, p_k) \in \mathbb{R}^{nk}$ and $c \in \mathbb{R}$ for which for all $x \in X_1$, $\sum_i p_i \cdot x < c$ and for all $(x_1, \dots, x_k) \in \prod_i A_i$, $\sum_i p_i \cdot x_i > c$. In particular we may choose c so that $\inf_{(x_1, \dots, x_k) \in \prod_i A_i} \sum_i p_i \cdot x_i > c$.

Define $p_i^* = p_i - (c/k)\mathbf{1}$ and observe that, because all A_i are subsets of X_1 , the c in the inequalities gets replaced with 0 when replacing p_i with p_i^* . Specifically, this also tells us that (by considering indicator functions of the form $\mathbf{1}_j$ for $j \in \{1, \dots, n\}$), we have $\sum_i p_i^* \ll 0$. Further, it tells us that for each $x, y \in X$, $\sum_i p_i^*(x)(k+1) - \sum_i p_i^*(y)k < 0$, as $(k+1)\mathbf{1}_x - k\mathbf{1}_y \in X_1$. Thus $\sum_i \frac{(k+1)p_i^*(x) - kp_i^*(y)}{2k+1} < 0$, and by taking limits, $\sum_i p_i^*(x) \leq \sum_i p_i^*(y)$. Since x and y are arbitrary, this implies that $\sum_i p_i^*(x) = \sum_i p_i^*(y)$ for all $x, y \in X$.

Now, define for each $i = 1, \dots, k$, $c_i = \inf_{x_i \in A_i} p_i^* \cdot x_i$. Then we have $\sum_i c_i > 0$. Fix $\epsilon > 0$ so that $\sum_i (c_i - \epsilon) > 0$. For each i then let $q_i^* = p_i^* - (c_i - \epsilon)\mathbf{1}$ and observe that for all $x_i \in A_i$, $q_i^* \cdot x_i = p_i^* \cdot x_i - (c_i - \epsilon) \cdot x_i \geq \epsilon > 0$ as $x_i \in X_1$. Further, $\sum_i q_i^* \ll 0$ remains valid as $\sum_i (c_i - \epsilon) > 0$. Further, it remains true that $\sum_i q_i^*(x) = \sum_i q_i^*(y)$ for all $x, y \in X$ as we have simply subtracted a constant from each q_i^* .

We show that 2 implies 1. Suppose by means of contradiction that there is $x^* \in \bigcap_{i=1}^k A_i$. Then $p_i \cdot x^* > 0$ for all i and in particular then $(\sum_{i=1}^k p_i) \cdot x^* > 0$. But this is a contradiction as $(\sum_{i=1}^k p_i) \cdot x^* = 0 \cdot x^* = 0$. \square

B Omitted Proofs

B.1 Proof of Theorem 1

Our proof of the equivalence between conditions (1) and (2) utilizes Lemma A.1 and follows quickly from it. Our proof of the equivalence between (1) and (3) utilizes Theorem 4. For the separation, let us consider the sets $\Delta(X)$ and, for each N , $\hat{c}o^{-1}(\Delta(p^N), p_i^N)$ where

$$\hat{c}o^{-1}(\Delta(p^N), p_i^N) = \{v | v = \sum_{j \in N} \gamma_j p_j^N, \gamma_j \leq 0 \forall j \in N \setminus i, \sum_{j \in N} \gamma_j = 1\}. \quad (9)$$

Note that this differs from Equation 3 as v is not required to be in the simplex. Owing to Rockafellar (1970) Theorem 19.3, each $\hat{c}o^{-1}(\Delta(p^N), p_i^N)$ is polyhedral.

Proof. We begin with the equivalence between (1) and (2). Suppose that $\{p_i\}$ is data consistent with GLM. Then, for each $i \in \mathcal{A}$, there exist v_i and π_i^N for each $N \in \mathcal{N}_i$ such that Equation 1 holds for each (i, N) . Since v_i is fixed across each $N \in \mathcal{N}_i$, by Lemma A.1, the collection of sets $\{c o^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ must have a point of common intersection.

Now suppose that, for each $i \in \mathcal{A}$, the collection of sets $\{c o^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ has a point of common intersection. By Lemma A.1, the v in the intersection of $\{c o^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ works as

a feasible v_i for each $N \in \mathcal{N}_i$. Specifically, for each $N \in \mathcal{N}_i$, there are $\gamma_j \geq 0$, with $\gamma_i > 0$ and $\sum_{j \in N} \gamma_j = 1$, such that $p_i^N = \gamma_i v + \sum_{j \in N \setminus i} \gamma_j p_j^N$. Take these γ_j as $\pi_i^N(j)$. These v and π_i correspond to a GLM representation of the dataset.

Now we prove the equivalence between (1) and (3). By definition of no incentive compatible money pump is equivalent to the lack of existence of a bet satisfying strict feasibility, individual rationality, and incentive compatibility. This is then equivalent to the following. For every $i \in \mathcal{A}$, there does not exist a collection of vectors $\{\alpha^N\}_{N \in \mathcal{N}_i}$ with $\alpha^N \in \mathbb{R}^X$ and $\sum_{N \in \mathcal{N}_i} \alpha^N \ll 0$ such that for all $N \in \mathcal{N}_i$, $\alpha^N \cdot p_i^N > 0$ and for each $j \neq i$ and N for which $j \in N$, $\alpha^N \cdot (p_i^N - p_j^N) \geq 0$. If there are α^N as in the statement of the theorem, then the model is not valid. Suppose, by means of contradiction that there are such α^N but the model is valid. Because $\pi_i^N(i) > 0$, we know that $v_i = \frac{p_i^N}{\pi_i^N(i)} - \sum_{j \neq i} \frac{\pi_i^N(j)}{\pi_i^N(i)} p_j^N$. In particular, since $\sum_x v_i(x) = 1$, we must have $\frac{1}{\pi_i^N(i)} - \sum_{j \neq i} \frac{\pi_i^N(j)}{\pi_i^N(i)} = 1$. Let $\lambda_i = \frac{1}{\pi_i^N(i)}$ and $\lambda_j = \frac{\pi_i^N(j)}{\pi_i^N(i)}$, then we have $\lambda_i - \sum_{j \neq i} \lambda_j = 1$, where each $\lambda_k \geq 0$. Then $\lambda_i p_i^N - \sum_{j \neq i} \lambda_j p_j^N = (\lambda_i - \sum_{j \neq i} \lambda_j) p_i^N + \sum_{j \neq i} \lambda_j (p_i^N - p_j^N)$. Consequently $\alpha^N \cdot v_i = (\lambda_i - \sum_{j \neq i} \lambda_j) \alpha^N \cdot p_i^N + \sum_{j \neq i} \lambda_j \alpha^N \cdot (p_i^N - p_j^N) > 0$. So for each N , we have $\alpha^N \cdot v_i > 0$. Consequently $(\sum_N \alpha^N) \cdot v_i > 0$, but $\sum_N \alpha^N \ll 0$, a contradiction.

On the other hand, suppose the model is violated. Recall $\hat{c}o^{-1}(\Delta(p^N), p_i^N)$. Observe that this set would not change were we to weaken the inequality on α_i to $\alpha_i > 0$. $\hat{c}o^{-1}(\Delta(p^N), p_i^N)$ reflects the set of possible v_i which are compatible with the observed choice p_i^N . That is, if there were some v_i common to all $\hat{c}o^{-1}(\Delta(p^N), p_i^N)$, we could define $\pi_i^N(i) = \frac{1}{\alpha_i}$ and $\pi_i^N(j) = -\frac{\alpha_j}{\alpha_i}$. Now, by Theorem 19.3 of Rockafellar (1970), each $\hat{c}o^{-1}(\Delta(p^N), p_i^N)$ is polyhedral (owing to the fact that the set of $\alpha_j : j \in N$ satisfying the linear inequalities is a polyhedron). Since the model is violated, there is no v_i common to all $\hat{c}o^{-1}(\Delta(p^N), p_i^N)$. Consequently $\Delta(X) \cap \bigcap_{N: i \in N} \hat{c}o^{-1}(\Delta(p^N), p_i^N) = \emptyset$. By Theorem 4, there are weights α^N and q for which $q + \sum_N \alpha^N = 0$, where $q \cdot x > 0$ for all $x \in \Delta(X)$, which means $q \gg 0$, and where $\alpha^N \cdot p_i^N > 0$ (by taking a weight of one on p_i^N) and $\alpha^N \cdot ((k+1)p_i^N - kp_j^N) > 0$ for all j , which implies by taking limits with respect to k and normalizing, $\alpha^N \cdot (p_i^N - p_j^N) \geq 0$. Finally, $\sum_N \alpha^N = -q \ll 0$.

□

B.2 Proof of Corollary 1

Proof. By Theorem 1 condition (2), we know that a dataset is consistent with GLM if and only if, for each i , $\{\hat{c}o^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ have a point of common intersection. As mentioned in Section 3.1, if p_i^N lies on the interior of $\Delta(p_{-i}^N)$, then N offers no testable content for agent i . Specifically, if p_i^N lies on the interior of $\Delta(p_{-i}^N)$ then p_i^N can be written as a convex combination of p_{-i}^N and $v \in \Delta(X)$ for any choice of v with a vanishingly small weight put on v . As such, we can restrict our attention to

$co^{-1}(\Delta(p^N), p_i^N)$ for groups N where $p_i^N \notin \Delta(p_{-i}^N)$. These groups are exactly $\mathcal{N}_i^+ \cup \mathcal{N}_i^-$. We have a series of observations for three cases.

1. Suppose $N \in \mathcal{N}_i^+ \setminus \mathcal{N}_i^-$

In this case $p_i^N \geq p_j^N$ for all $j \in N \setminus i$ with a strict inequality for at least one j . Fix this j . In this case, set $\pi_i^N(k) = 0$ for each $k \in N \setminus \{i, j\}$. It then follows that every v satisfying $v \geq p_i^N$ can be rationalized by $p_i^N = \pi_i^N(i)v + \pi_i^N(j)p_j^N$ for some choice of convex weights, but any $v < p_i^N$ cannot.

2. Suppose $N \in \mathcal{N}_i^- \setminus \mathcal{N}_i^+$

In this case $p_i^N \leq p_j^N$ for all $j \in N \setminus i$ with a strict inequality for at least one j . Fix this j . In this case, set $\pi_i^N(k) = 0$ for each $k \in N \setminus \{i, j\}$. It then follows that every v satisfying $v \leq p_i^N$ can be rationalized by $p_i^N = \pi_i^N(i)v + \pi_i^N(j)p_j^N$ for some choice of convex weights but any $v > p_i^N$ cannot.

3. Suppose $N \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-$

In this case $p_i^N = p_j^N$ for each $j \in N$. In this case, any convex combination of v with p_j^N for $j \in N \setminus i$ which puts positive weight on v would be different from p_j^N if and only if $p_i^N \neq p_j^N$. Thus, in this case we have that $v_i = p_j^N = p_i^N$.

We now consider two cases.

1. Suppose that $\mathcal{N}_i^+ \cap \mathcal{N}_i^- = \emptyset$.

In this case $[0, \min_{N \in \mathcal{N}_i^-} p_i^N] = \bigcap_{N \in \mathcal{N}_i^-} \{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i^-}$ and $[\max_{N \in \mathcal{N}_i^+, 1} p_i^N, 1] = \bigcap_{N \in \mathcal{N}_i^+} \{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i^+}$. These two sets have a point of intersection if and only if $\max_{N \in \mathcal{N}_i^+} p_i^N \leq \min_{N \in \mathcal{N}_i^-} p_i^N$. The other two conditions are vacuous in this case.

2. Suppose that $\mathcal{N}_i^+ \cap \mathcal{N}_i^- \neq \emptyset$.

By observation 3 above, for each $N^* \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-$, $co^{-1}(\Delta(p^{N^*}), p_i^{N^*})$ is exactly $p_i^{N^*}$ and so it must be the case that $v_i = p_i^{N^*} = p_i^-$ if the dataset is consistent. It then follows that this value v_i must be unique among consistent data. By the arguments from case 1, we need it to be that $\max_{N \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-} p_i^N \leq \min_{N \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-} p_i^N$. Further, it must also be the case that $\max_{N \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-} p_i^N \leq p_i^- \leq \min_{N \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-} p_i^N$. However, $N^* \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-$, so $\max_{N \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-} p_i^N \geq p_i^- \geq \min_{N \in \mathcal{N}_i^+ \cap \mathcal{N}_i^-} p_i^N$. It then follows that the three conditions hold if and only if the dataset is consistent with GLM.

□

B.3 Proof of Proposition 1

Proof. This follows immediately from Lemma A.1 and Theorem 1. \square

B.4 Proof of Corollary 2

Proof. By consistency with the linear-in-means model, $co^{-1}(\Delta(p^{N_j}), p_i^{N_j})$ and $co^{-1}(\Delta(p^{N_k}), p_i^{N_k})$ intersect. By N_j and N_k being binary, $co^{-1}(\Delta(p^{N_j}), p_i^{N_j})$ and $co^{-1}(\Delta(p^{N_k}), p_i^{N_k})$ are one dimensional rays, and so they intersect once or infinitely often. By linear independence of $(p_i^{N_j} - p_j^{N_j})$ and $(p_i^{N_k} - p_k^{N_k})$ and by the definition of $co^{-1}(\Delta(p^{N_j}), p_i^{N_j})$ and $co^{-1}(\Delta(p^{N_k}), p_i^{N_k})$, these two sets intersect at a single point. By Proposition 1, this single point of intersection is exactly our identified set for v_i . \square

B.5 Proof of Proposition 2

Proof. As v_i and $\{p_k^N\}_{k \in N \setminus i}$ are linearly independent, any point in their convex hull can be written as a convex combination of these points with unique weights on each point. In the linear-in-means model, we have $p_i^N = \pi_i^N(i)v_i + \sum_{j \neq i} \pi_i^N(j)p_j^N$, and so p_i^N is in the convex hull of these points and thus each $\pi_i^N(j)$ is uniquely pinned down. \square

B.6 Proof of Theorem 2

Proof. To begin observe that no weakly incentive compatible money pump is satisfied if and only if there does not exist a collection of vectors $\{\alpha^N\}_{N \in \mathcal{N}_i}$ with $\alpha^N \in \mathbb{R}^X$ and $\sum_{N \in \mathcal{N}_i} \alpha^N \ll 0$ such that for all $N \in \mathcal{N}_i$, $\alpha^N \cdot p_i^N > 0$ and for each $j \neq i$, $\sum_{N \in \mathcal{N}_i: j \in N} \alpha^N \cdot (p_i^N - p_j^N) \geq 0$. Now suppose that LLM is satisfied, and by means of contradiction, that there exist α^N as in the prior sentence. Drop dependence of w_i on i to simplify notation. We use the notation $w(N) = \sum_{j \in N} w(j)$.

Observe that

$$\begin{aligned}
\sum_{N \in \mathcal{N}_i} w(N) \alpha^N \cdot p_i^N &= \sum_{N \in \mathcal{N}_i} \left(w(i) \alpha^N \cdot v_i + \sum_{j \in N \setminus i} w(j) \alpha^N \cdot p_j^N \right) \\
&= w(i) \left(\sum_{N \in \mathcal{N}_i} \alpha^N \right) \cdot v_i + \sum_{j \in \mathcal{A} \setminus i} w(j) \sum_{N \in \mathcal{N}_i: j \in N} \alpha^N \cdot p_j^N \\
&< \sum_{j \in \mathcal{A} \setminus i} w(j) \sum_{N \in \mathcal{N}_i: j \in N} \alpha^N \cdot p_j^N \\
&\leq \sum_{j \in \mathcal{A} \setminus i} w(j) \sum_{N \in \mathcal{N}_i: j \in N} \alpha^N \cdot p_i^N \\
&= \sum_{N \in \mathcal{N}_i} \sum_{j \in N \setminus i} w(j) \alpha^N \cdot p_i^N.
\end{aligned}$$

Here, the strict inequality follows as $v_i \in \Delta(X)$, $w(i) > 0$, and $\sum_{N \in \mathcal{N}_i} \alpha^N \ll 0$. The weak inequality follows as $\sum_{N \in \mathcal{N}_i: j \in N} \alpha^N (p_i^N - p_j^N) \geq 0$ and $w(j) \geq 0$. The equalities are by definition and algebraic manipulation. Now, subtracting the right hand side of the string of inequalities from the left hand side, we obtain:

$$\sum_{N \in \mathcal{N}_i} w(i) \alpha^N \cdot p_i^N < 0.$$

This contradicts the facts that $w(i) > 0$ and for each $N \in \mathcal{N}_i$, $\alpha^N \cdot p_i^N > 0$.

For the other direction, we want to find, for each i , numbers $f_i(x) \in \mathbb{R}$, and for each $j \neq i$, $w_i(j) \in \mathbb{R}$ and finally for each N for which $i \in N$, $\lambda_i(N)$ such that the following equations are satisfied:

1. $f_i(x) \geq 0$ for all x
2. $0 < \sum_x f_i(x)$
3. $w_i(j) \geq 0$ for all $j \neq i$
4. $f_i(x) + \lambda_i(N) p_i^N(x) + \sum_{j \in N \setminus i} w_i(j) p_j^N(x) = 0$ for all x and all N for which $i \in N$

If we find such numbers, define the Luce weights for agent i as $w(i) = \sum_x f_i(x) > 0$, $w(j) = w_i(j) \geq 0$ and define $v_i(x) = \frac{f_i(x)}{\sum_x f_i(x)}$. Observe then that by equation 4:

$$\begin{aligned}
w(i) v_i(x) + \lambda_i(N) p_i^N(x) + \sum_{j \in N \setminus i} w(j) p_j^N(x) &= 0, \text{ so that} \\
-\lambda_i(N) p_i^N(x) &= w(i) v_i(x) + \sum_{j \in N \setminus i} w(j) p_j^N(x).
\end{aligned}$$

Since we know $\sum_x v_i(x) = 1$ (by definition), $\sum_x p_i^N(x) = 1$ and $\sum_x p_j^N(x) = 1$ (by assumption), it follows automatically that $-\lambda_i(N) = w(N)$ and we have shown that this is a linear formulation of Luce.

The dual of this system, a Theorem of the Alternative, for example Motzkin's Theorem, (Mangasarian, 1994) p. 28 implies the existence of $\beta^N \in \mathbb{R}^X$ for each $N \in \mathcal{N}_i$ for which

1. For all $N \in \mathcal{N}$ for which $i \in N$, $\beta^N \cdot p_i^N = 0$.
2. For all $j \in \mathcal{A} \setminus i$, $\sum_{N \in \mathcal{N}_i: j \in N} \beta^N \cdot p_j^N \leq 0$.
3. $\sum_{N \in \mathcal{N}_i} \beta^N \ll 0$.

Now, owing to finiteness of X , we know that there exists some $b < 0$ such that for each $x \in X$, $\sum_{N \in \mathcal{N}_i} \beta^N(x) < b$. Define $\alpha^N = \beta^N - \frac{b}{|\mathcal{N}_i|} \mathbf{1}$. Observe that $\alpha^N \cdot p_i^N = \beta^N \cdot p_i^N - \frac{b}{|\mathcal{N}_i|} = \frac{-b}{|\mathcal{N}_i|} > 0$. Observe that $\sum_{N \in \mathcal{N}_i: j \in N} \alpha^N \cdot p_j^N \leq \frac{-b|\{N \in \mathcal{N}_i: j \in N\}|}{|\mathcal{N}_i|} = \sum_{N \in \mathcal{N}_i: j \in N} \alpha^N \cdot p_j^N$, so that $\sum_{N \in \mathcal{N}_i: j \in N} \alpha^N \cdot (p_i^N - p_j^N) \geq 0$. Finally, $\sum_{N \in \mathcal{N}_i} \alpha^N = (\sum_{N \in \mathcal{N}_i} \beta^N) - b \mathbf{1} \ll 0$.

□

C GLM Results when $\pi_i^N(i) \geq 0$

In this section we consider an extension of GLM where we now allow $\pi_i^N(i) \geq 0$ (allowing for equality). We call this version of the model GLM*. Our goal is to discuss and prove how our main results on GLM extend to GLM*. Recall that for an agent i , \mathcal{N}_i^{ext} corresponds to the set of groups N containing i with $p_i^N \notin \Delta(p_{-i}^N)$. Observe that in GLM*, whenever p_i^N lies in $\Delta(p_{-i}^N)$, p_i^N is consistent with GLM* as it can be written as a convex combination of p_{-i}^N . This motivates the following.

Definition 12. A set of vectors $\{b^N\}_{N \in \mathcal{N}_i^{ext}}$ with $b^N \in \mathbb{R}^X$ for each $N \in \mathcal{N}_i$ is called a **ext bet on agent i** .

Definition 13. A bet on agent i is **ext strictly feasible** if $\sum_{N \in \mathcal{N}_i^{ext}} b^N \ll 0$.

Definition 14. A bet on agent i is **ext individually rational** if $b^N \cdot p_i^N > 0$ for each $N \in \mathcal{N}_i^{ext}$.

Definition 15. A bet on agent i is **ext incentive compatible** if $b^N \cdot (p_i^N - p_j^N) \geq 0$ for each $N \in \mathcal{N}_i^{ext}$ and each $j \in N \setminus i$.

Definition 16. We say that a dataset $\{p^N\}_{N \in \mathcal{N}}$ satisfies **no ext incentive compatible money pump** if for each $i \in \mathcal{A}$ there are no ext strictly feasible, ext individually rational, and ext incentive compatible ext bets on agent i .

Theorem 5. For a dataset $\{p^N\}_{N \in \mathcal{N}}$, the following are equivalent.

1. $\{p^N\}_{N \in \mathcal{N}}$ is consistent with GLM*.
2. For every $i \in \mathcal{A}$, the collection of sets $\{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i^{ext}}$ has a point of mutual intersection.
3. $\{p^N\}_{N \in \mathcal{N}}$ satisfies no ext incentive compatible money pump.

Before proving Theorem 5, we first prove a preliminary lemma.

Lemma C.1. Suppose $p_i^N \notin \Delta(p_{-i}^N)$. Then $p_i^N = \gamma_i v_i + \sum_{j \in N \setminus i} \gamma_j p_j^N$ for $v_i \in \Delta(X)$ with $\gamma_j \geq 0$ and $\sum_{j \in N} \gamma_j = 1$ if and only if $v_i \in co^{-1}(\Delta(p^N), p_i^N)$.

Proof. Suppose that $p_i = \gamma_i v_i + \sum_{j \in N \setminus i} \gamma_j p_j^N$ for $v_i \in \Delta(X)$ with $\gamma_j \geq 0$ and $\sum_{j \in N} \gamma_j = 1$. Since $p_i^N \notin \Delta(p_{-i}^N)$, it must be the case that $\gamma_i > 0$ and so $0 < \gamma_i \leq 1$. By doing basic algebra we can recover that $v_i = \frac{1}{\gamma_i} p_i^N + \sum_{j \in N \setminus i} \frac{-\gamma_j}{\gamma_i} p_j^N$. Since $\gamma_j \geq 0$ and $\sum_{j \in N} \gamma_j = 1$, $\frac{-\gamma_j}{\gamma_i} \leq 0$ and $\frac{1}{\gamma_i} + \sum_{j \in N \setminus i} \frac{-\gamma_j}{\gamma_i} = 1$ and so $v_i \in co^{-1}(\Delta(p^N), p_i^N)$.

Now suppose that $v_i \in co^{-1}(\Delta(p^N), p_i^N)$. This tells us that $v_i = \sum_{j \in N} \gamma_j p_j^N$ with $\gamma_j \leq 0 \forall j \in N \setminus i$ and $\sum_{j \in N} \gamma_j = 1$. It follows that $\gamma_i \geq 1$. By basic algebra, we can solve for p_i and get $p_i = \frac{1}{\gamma_i} v_i + \sum_{j \in N \setminus i} \frac{-\gamma_j}{\gamma_i} p_j^N$. Since $\gamma_j \geq 0$ for $j \in N \setminus i$ and $\sum_{j \in N} \gamma_j = 1$, $\frac{-\gamma_j}{\gamma_i} \geq 0$ and $\frac{1}{\gamma_i} + \sum_{j \in N \setminus i} \frac{-\gamma_j}{\gamma_i} = 1$ and so $p_i = \pi_i v_i + \sum_{j \in N \setminus i} \pi_j p_j^N$ for $v_i \in \Delta(X)$ with $\pi_j \geq 0$ and $\sum_{j \in N} \pi_j = 1$. \square

We now proceed to prove Theorem 5.

Proof. The equivalence between (1) and (2) follows from Lemma C.1 and our discussion at the start of this section. To see the equivalence between (1) and (3), observe that, in the proof of Theorem 1, the arguments from the proof of (1) \implies (3) go through replacing \mathcal{N}_i with \mathcal{N}_i^{ext} and observing that once $\pi_i^N \notin \Delta(p_{-i}^N)$ any coefficient preceding π_i^N will be non-zero (see the first two sentences of the proof of Lemma C.1). To show (3) implies (1), observe that the arguments from (3) \implies (1) from the proof of Theorem 1 hold when we try and prove the separation of $\Delta(X)$ and $\{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i^{ext}}$. \square

Now let $\mathcal{N}_i^{-*} \subseteq \mathcal{N}_i$ denote the set of groups N satisfying $p_i^N < p_j^N$ for each $j \in N \setminus i$. Similarly, let $\mathcal{N}_i^{+*} \subseteq \mathcal{N}_i$ denote the set of groups N satisfying $p_i^N > p_j^N$ for each $j \in N \setminus i$.

Corollary 3. In the one dimension case, a dataset $\{p^N\}_{N \in \mathcal{N}}$ is consistent with GLM* if and only if, for all $i \in \mathcal{A}$, $\min_{N \in \mathcal{N}_i^{-*}} p_i^N \geq \max_{N \in \mathcal{N}_i^{+*}} p_i^N$ when both \mathcal{N}_i^{-*} and \mathcal{N}_i^{+*} are non-empty.

Proof. Observe that $\mathcal{N}_i^{-*} \cup \mathcal{N}_i^{+*} = \mathcal{N}_i^{ext}$ in the one dimension case. Further, by the arguments in the proof of Corollary 1, $co^{-1}(\Delta(p^N), p_i^N) = [0, p_i^N]$ when $N \in \mathcal{N}_i^{-*}$ and $co^{-1}(\Delta(p^N), p_i^N) = [p_i^N, 1]$ when $N \in \mathcal{N}_i^{+*}$. It then follows from Theorem 5 that our corollary holds. \square

Proposition 4. *In GLM*, the tight identified set for v_i is given by $\bigcap_{N \in \mathcal{N}_i^{ext}} co^{-1}(\Delta(p^N), p_i^N)$.*

Proof. This is an immediate consequence of Lemma C.1 and Theorem 5. \square

All of the other identification results for GLM carry through to GLM* as written.

D Relation Between Theorem 1 and Samet (1998)

In this appendix, we discuss the relation between Samet (1998) and our characterizations of GLM and GLM*. The main theorem of Samet (1998) provides a linear program that characterizes when a collection of compact convex sets fail to have a point of mutual intersection. In light of Theorems 1 and 5, the result of Samet (1998) can be used as an alternative characterization for GLM and GLM*.

Theorem 6. *The following are equivalent.*

1. $\{p^N\}_{N \in \mathcal{N}}$ is consistent with the general linear-in-means model.
2. For every $i \in \mathcal{A}$ there does not exist a collection of vectors $\{\alpha^N\}_{N \in \mathcal{N}_i}$ with $\alpha^N \in \mathbb{R}^X$ and $\sum_{N \in \mathcal{N}_i} \alpha^N = 0$ such that, for all $N \in \mathcal{N}_i$, $\alpha^N \cdot v > 0$ for each $v \in co^{-1}(\Delta(p^N), p_i^N)$.

Theorem 6 immediately follows from condition (2) of Theorem 1, the main result of Samet (1998), and the observation that each $co^{-1}(\Delta(p^N), p_i^N)$ is compact and convex. Part of the original motivation for Samet (1998) was the no interim trade result of Morris (1994). In the setting of Samet (1998) and Morris (1994), $\{\alpha^N\}_{N \in \mathcal{N}_i}$ corresponds to a trading scheme, the equality condition corresponds to a feasibility condition, and the inequality condition corresponds to an individual rationality condition.

We now interpret Theorem 6 in the context of a no trade story. Suppose we have a collection of outside observers. We have one observer for each group N in \mathcal{N}_i and we index each observer by their corresponding group N . These observers are risk neutral but ambiguity averse (in the max-min sense of Gilboa and Schmeidler (1989)) and consider state contingent trades. In this setting, a state corresponds to an alternative x and the frequency of state x corresponds to the frequency with which agent i would choose x in isolation, $v_i(x)$. In this sense, $\{\alpha^N\}_{N \in \mathcal{N}_i}$ corresponds to a trade scheme in the language of Morris (1994) where $\alpha^N(x)$ corresponds to the payout to observer N

when agent i chooses x in isolation. Each observer knows that agent i 's choices follow from Equation 1. The information available to observer N is p^N , the set of choices made in group N . From this, observer N is able to form a set of beliefs about the choices of agent i . This set of beliefs corresponds to $co^{-1}(\Delta(p^N), p_i^N)$. The condition $\sum_{N \in \mathcal{N}_i} \alpha^N = 0$ corresponds to a trading scheme being balanced and $\alpha^N \cdot v > 0$ for each $v \in co^{-1}(\Delta(p^N), p_i^N)$ corresponds to the trading scheme being profitable in expectation for every belief in observer N 's set of beliefs. With this interpretation, Theorem 6 says that there is no trade between these outside observers if and only if the dataset is consistent with GLM. In other words, there is no trade if and only if each agent's ideal point v_i is common across their groups.

Before moving on, we also note that condition (2) from Theorem 6 can be simplified somewhat. Since $co^{-1}(\Delta(p^N), p_i^N)$ is a convex set and $\alpha^N \cdot v > 0$ is a linear constraint, instead of checking $\alpha^N \cdot v > 0$ for each $v \in co^{-1}(\Delta(p^N), p_i^N)$, we can simply check $\alpha^N \cdot v > 0$ for the extreme points of $co^{-1}(\Delta(p^N), p_i^N)$. Since $co^{-1}(\Delta(p^N), p_i^N)$ is polyhedral, this is a finite process. Finally, we state the analogue of Theorem 6 for GLM*.

Theorem 7. *The following are equivalent.*

1. $\{p^N\}_{N \in \mathcal{N}}$ is consistent with GLM*.
2. For every $i \in \mathcal{A}$ there does not exist a collection of vectors $\{\alpha^N\}_{N \in \mathcal{N}_i^{ext}}$ with $\alpha^N \in \mathbb{R}^X$ and $\sum_{N \in \mathcal{N}_i^{ext}} \alpha^N = 0$ such that, for all $N \in \mathcal{N}_i^{ext}$, $\alpha^N \cdot v > 0$ for each $v \in co^{-1}(\Delta(p^N), p_i^N)$.

Theorem 7 immediately follows from condition (2) of Theorem 5, the main result of Samet (1998), and the observation that each $co^{-1}(\Delta(p^N), p_i^N)$ is compact and convex.

E Proofs and Extension of Results from Section 3.3

In this section, we consider the ULM model of Section 3.3 and allow for variation in group size. With this in mind, we now introduce the extensions of the axioms from 3.3 allowing for group size variation.

Axiom 5 (Cyclically constant*). *A dataset $\{p_N\}_{N \in \mathcal{N}}$ is cyclically constant if for every cycle $\{(i_k, j_k, N_k)\}_{k=1}^m$ we have that $\sum_{k=1}^m (1 + |N_k|)(p_{i_k}^{N_k} - p_{j_k}^{N_k}) = 0$.*

Axiom 6 (Constant Peer Effects*). *A dataset $\{p_N\}_{N \in \mathcal{N}}$ satisfies constant peer effects if for each N and M with $i, j \in N \cap M$ we have that $(1 + |N|)[p_i^N - p_j^N] = (1 + |M|)[p_i^M - p_j^M]$.*

Axiom 7 (Symmetric Peer Effects*). *For $i \in N \cap M$,*

$$p_i^N - \sum_{j \in N \setminus M} [p_j^N - p_i^N] = p_i^M - \sum_{k \in M \setminus N} [p_k^M - p_i^M] - \frac{|N| - |M|}{1 + |N|} \sum_{l \in N \cap M} [p_l^M - p_i^M]. \quad (10)$$

Observe that in the prior three axioms, each axiom reduces to the axioms from the main text when each group has the same size.

Theorem 8. *For a dataset $\{p^N\}_{N \in \mathcal{N}}$, the following are equivalent.*

1. $\{p^N\}_{N \in \mathcal{N}}$ is consistent with the uniform linear-in-means model.
2. $\{p^N\}_{N \in \mathcal{N}}$ satisfies cyclic constancy*, symmetric peer effects*, and bounded total peer effects.

Proof. We begin by proving necessity of the three axioms. We begin with cyclic constancy. By the definition of ULM we have the following for $i, j \in N$.

$$\begin{aligned}
p_i^N - p_j^N &= \frac{1}{|N|}v_i + \sum_{k \neq i} \frac{1}{|N|}p_k^N - \frac{1}{|N|}v_j - \sum_{k \neq j} \frac{1}{|N|}p_k^N \\
&= \frac{1}{|N|}v_i - \frac{1}{|N|}v_j + \frac{1}{|N|}(p_j^N - p_i^N) \\
&= \frac{1}{1 + |N|}(v_i - v_j)
\end{aligned} \tag{11}$$

The equivalence between the second and third line follows from the properties of geometric series. Now consider any cycle $\{(i_k, j_k, N_k)\}_{k=1}^m$.

$$\begin{aligned}
\sum_{k=1}^m (1 + |N_k|)(p_{i_k}^{N_k} - p_{j_k}^{N_k}) &= \sum_{k=1}^m (1 + |N_k|) \left(\frac{1}{1 + |N_k|} (v_{i_k} - v_{j_k}) \right) \\
&= \sum_{k=1}^m (v_{i_k} - v_{j_k}) \\
&= 0
\end{aligned} \tag{12}$$

The first line follows from substitution. The second line follows from canceling like terms. The third line follows from the fact that for cycles $j_k = i_{k+1}$, and $j_m = i_1$. To see that bounded total peer effects is necessary, observe the following.

$$\begin{aligned}
p_i^N &= \frac{1}{|N|}v_i + \frac{1}{|N|} \sum_{j \in N \setminus i} p_j^N \\
v_i &= |N|p_i^N - \sum_{j \in N \setminus i} p_j^N \\
&= p_i^N - \sum_{j \in N \setminus i} [p_j^N - p_i^N]
\end{aligned} \tag{13}$$

The first equality holds by definition of our model. The second and third equalities hold by rearrangement. Since we require our bliss point v_i to be in the simplex, bounded total peer effect then holds by

the last line. We now prove necessity of symmetric peer effects. By Equation 13 we have the following.

$$\begin{aligned}
p_i^N - \sum_{j \in N \setminus i} [p_j^N - p_i^N] &= p_i^M - \sum_{k \in M \setminus i} [p_k^M - p_i^M] \\
p_i^N - \sum_{j \in N \setminus M} [p_j^N - p_i^N] + \sum_{l \in (N \cap M) \setminus i} \frac{1 + |M|}{1 + |N|} [p_l^M - p_i^M] & \\
= p_i^M - \sum_{k \in M \setminus N} [p_k^M - p_i^M] - \sum_{l \in (N \cap M) \setminus i} [p_l^M - p_i^M] & \tag{14} \\
p_i^N - \sum_{j \in N \setminus M} [p_j^N - p_i^N] = p_i^M - \sum_{k \in M \setminus N} [p_k^M - p_i^M] - \frac{|N| - |M|}{1 + |N|} \sum_{l \in N \cap M} [p_l^M - p_i^M] &
\end{aligned}$$

The first equality comes from Equation 13. The second equality comes from cyclic constancy (which we have proved holds).⁸ The third equality follows from gathering like terms. Thus all three of our axioms are necessary.

We now proceed to show sufficiency of our three axioms. To do so we need to construct a $v_i \in \Delta(X)$ for each i such that these v_i are consistent with the data we observe. To begin, for each i , choose some N such that $i \in N$. Define \hat{v}_i as follows.

$$\hat{v}_i = p_i^N - \sum_{j \in N \setminus i} (p_j^N - p_i^N)$$

By bounded total peer effect, \hat{v}_i is non-negative. Since p_i^N and p_j^N are probabilities, $\sum_{x \in X} \hat{v}_i(x) = 1$ and thus \hat{v}_i lies in the simplex. We now verify if our construction induces our observed data. Consider an arbitrary M such that $i \in M$. We start with \hat{p}_i^M as the choice probabilities implied by our

⁸Note that here we only need to use binary cycles (i.e. $[i, j] - > [j, i]$).

construction.

$$\begin{aligned}
\hat{p}_i^M &= \frac{1}{|M|} \hat{v}_i + \frac{1}{|M|} \sum_{j \in M \setminus i} p_j^M \\
&= \frac{1}{|M|} (p_i^N - \sum_{j \in N \setminus i} [p_j^N - p_i^N] + \sum_{j \in M \setminus i} p_j^M) \\
\hat{p}_i^M - p_i^M &= \frac{1}{|M|} (p_i^N - \sum_{j \in N \setminus i} [p_j^N - p_i^N] - (p_i^M - \sum_{j \in M \setminus i} [p_j^M - p_i^M])) \\
&= \frac{1}{|M|} (p_i^N - \sum_{j \in N \setminus M} [p_j^N - p_i^N] + \\
&\quad \sum_{l \in (N \cap M) \setminus i} \frac{1 + |M|}{1 + |N|} [p_l^M - p_i^M] - (p_i^M - \sum_{j \in M \setminus i} [p_j^M - p_i^M])) \\
&= \frac{1}{|M|} (p_i^N - \sum_{j \in N \setminus M} [p_j^N - p_i^N] - (p_i^M - \sum_{j \in M \setminus N} [p_j^M - p_i^M]) \\
&\quad - \frac{|N| - |M|}{1 + |N|} \sum_{l \in N \cap M} [p_l^M - p_i^M]) \\
&= 0
\end{aligned} \tag{15}$$

The first equality follows from the definition of our model. The second equality follows from our definition of \hat{v}_i . The third equality follows by subtracting p_i^M from both sides. The fourth equality follows from cyclic constancy.⁹ The fifth equality follows from collecting like terms. The sixth equality follows from symmetric peer effects. We have constructed \hat{v}_i which is in the simplex and induces our observed data. Thus cyclic constancy, symmetric peer effects, and bounded total peer effects are sufficient. \square

Proposition 5. *Cyclic constancy* holds iff for each $i \in \mathcal{A}$, there is v_i for which $\sum_x v_i(x) = 1$ and for each $N \in \mathcal{N}$, $O^N \in \mathbb{R}^X$ such that $\sum_x O^N(x) = 0$ such that for all N and $i \in N$, $p_i^N(x) = \frac{1}{|N|} v_i(x) + \frac{1}{|N|} \sum_{j \in N \setminus i} p_j^N(x) + O^N(x)$.*

Remark 4. *In Proposition 5, we have not specified conditions on v_i and O^N ensuring that p_i^N is actually a probability measure.*

Proof. Fix any $i^* \in \mathcal{A}$ and define v_{i^*} arbitrarily so that $\sum_x v_{i^*} = 1$; we do not require that $v_{i^*} \geq 0$. Define $v_j = 3(p_j^{\{i^*, j\}} - p_{i^*}^{\{i^*, j\}})$. Cyclic constancy implies that for any $N \in \mathcal{N}$ and $\{i, j\} \subseteq N$, $(1 + |N|)(p_i^N - p_j^N) = v_i - v_j$. That is, $(1 + |N|)(p_i^N - p_j^N) = 3(p_i^{\{i, i^*\}} - p_{i^*}^{\{i, i^*\}}) - 3(p_j^{\{j, i^*\}} - p_{i^*}^{\{j, i^*\}}) = v_i - v_j$. Define the uniform Luce rule according to Luce weights v_i (this may not necessarily have probabilities as frequencies here). Call this rule U_i^N , so that for each $N \in \mathcal{N}$ and each $i \in N$, $U_i^N = \frac{1}{|N|} v_i +$

⁹Again, here we only use binary cycles.

$\sum_{j \in N \setminus i} U_j^N$. By definition, for all N and all $i, j \in N$ with $i \neq j$, we have: $(p_i^N - p_j^N) = (U_i^N - U_j^N)$ (this equality obviously holds when $i = j$). Consequently for all such i, N :, $p_i^N = U_i^N + (p_j^N - U_j^N)$.

By taking averages $p_i^N = U_i^N + \frac{1}{|N|} \sum_{j \in N} (p_j^N - U_j^N)$. Now let $\bar{O}^N = \frac{1}{|N|} \sum_{j \in N} (p_j^N - U_j^N)$. Observe that $\sum_x \bar{O}^N(x) = 0$. Therefore, for all $N \in \mathcal{N}$ and all $i \in N$, $p_i^N = U_i^N + \bar{O}^N$. Finally, using the representation of U_i^N , we obtain $p_i^N = \frac{1}{|N|} v_i + \frac{1}{|N|} \sum_{j \in N \setminus i} U_j^N + \bar{O}^N$, and using the fact that $U_j^N = p_j^N - \bar{O}^N$, we get

$$p_i^N = \frac{1}{|N|} v_i + \frac{1}{|N|} \sum_{j \in N \setminus i} (p_j^N - \bar{O}^N) + \bar{O}^N$$

or $p_i^N = \frac{1}{|N|} v_i + \frac{1}{|N|} \sum_{j \in N \setminus i} p_j^N + \frac{1}{|N|} \bar{O}^N$. Setting $O^N = \frac{1}{|N|} \bar{O}^N$ establishes the result.

Conversely if we start with v_i for each $i \in \mathcal{A}$ for which $\sum_x v_i(x) = 1$ (not necessarily a probability), and for each $N \in \mathcal{N}$, a vector O^N for which $\sum_x O^N(x) = 0$, then if $p_i^N = U_i^N + O^N$ is a probability measure for all $N \in \mathcal{N}$ and $i \in N$, then simple algebra establishes that for any $N \in \mathcal{N}$ and $i, j \in N$, $(1 + |N|)(p_i^N - p_j^N) = v_i - v_j$. As a telescoping series, a sum across a cycle results in 0. \square