

1 DISCRETE CHOICE WITH GENERALIZED SOCIAL INTERACTIONS 1

2

3 OSCAR VOLPE 3

4 Department of Economics, University of Chicago 4

5

6 This paper explores how identity affects behavior through social interactions. 6

7 I study a discrete choice model where people wish to conform to the actions of 7

8 some members of their network, while deviating from the actions of others. Under 8

9 this generalized framework, I explore what aggregate outcomes arise from nonco- 9

10 operative decisionmaking. I characterize the uniqueness and stability of equilibria, 10

11 and I discuss implications of negative spillovers for welfare and inefficiency. Ad- 11

12 ditionally, I demonstrate how the model may be taken to data. I introduce a novel 12

13 identification strategy that accounts for unobserved network effects by leveraging 13

14 within-network variation in individual characteristics. I also construct internal in- 14

15 struments to overcome the issue of measurement error, which is a primary source 15

16 of endogeneity in network-based models with incomplete information. Finally, 16

17 I apply this methodology using data from the large-scale education experiment 17

18 Project STAR, where I find robust evidence of gender differences in peer effects. 18

19

20

21 KEYWORDS: identity, social interactions, strategic complements & substitutes, 21

22 contextual effects, panel, measurement error, partial identification, Project STAR. 22

23

24

25 1. INTRODUCTION 25

26

27 The role of social interactions in individual decisionmaking has received widespread 27

28 attention in economics. This work has largely focused on settings with uniform strategic 28

29 complementarities, where agents seek to conform to the average action in the rest of their 29

30 network. Meanwhile, less attention has been given to settings with nonuniform interactions, 30

31 where agents conform to the behavior of some people, while deviating from others. These 31

32

33

34

35 Oscar Volpe: ovolpe@uchicago.edu 35

36 I give thanks to Steven Durlauf, whose feedback and encouragement has made this project possible. I am also 36

37 grateful for invaluable comments and advice from Stéphane Bonhomme, Ozan Candogan, Arun Chandrasekhar, 37

38 Michael Dinerstein, Ali Hortaçsu, Robert Moffitt, Magne Mogstad, Philip Reny, and Alex Torgovitsky, as well 38

39 as participants in the North American Urban Economics Association Meeting, the Applied Micro working group, 39

40 the IO working group, and the third year research seminar at the University of Chicago. Any errors are my own. 40

41

1 types of interactions are more challenging to study, particularly when agents make discrete 1
2 choices, since the nonlinear dependence between outcomes and the potential for multiple 2
3 equilibria may complicate the theoretical understanding of behavioral outcomes and also 3
4 pose barriers to identification and estimation. Nevertheless, nonuniform interactions appear 4
5 in many real-world contexts. Consider, for example, the impact of gender on peer effects 5
6 in schools. Male and female students often face pressure to conform to members of their 6
7 own gender, while there may be less pressure to conform across genders. In some cases, 7
8 students might even wish to distinguish themselves from the opposite gender—a preference 8
9 that can be modeled by a negative interaction effect. As [Hoxby \(2000\)](#) and others discuss, 9
10 these types of peer effects may contribute to gender differences in behavior and academic 10
11 achievement. More generally, as [Akerlof & Kranton \(2000\)](#) argue, social identity (e.g., 11
12 gender, race, class, or politics) helps to determine a wide range of phenomena, and it is 12
13 relevant for most everyday decisions that people make. Therefore, it is worth understanding 13
14 the role of nonuniform, and even negative, social interaction effects on economic outcomes. 14

15 In this paper, I develop a framework for studying generalized social interactions, where 15
16 agents are affected differently by different people in their network. I focus on a model with 16
17 binary choices where agents experience positive and negative spillover effects. I character- 17
18 ize what aggregate outcomes occur in equilibrium, and I show how the model is empirically 18
19 tractable. This work is especially useful for understanding the impact of identity on social 19
20 and economic behavior, as well as the emergence of polarized or segregated networks. 20

21 I begin my analysis by modeling a network of agents who make binary choices sub- 21
22 ject to social influences and private, idiosyncratic preferences. Following [Brock & Durlauf](#) 22
23 [\(2001\)](#), I assume that agents act with incomplete information, such that they are influenced 23
24 by the expected average outcomes in the network, rather than by every person’s realized 24
25 outcome. However, unlike previous work, I assume that the social interaction effects differ, 25
26 and might even be negative, across agents. To achieve this objective, I categorize the agents 26
27 into distinct social groups (or “identities”). These classifications play a role in determining 27
28 private payoffs, and they also form the basis for nonuniform social interactions. In partic- 28
29 ular, the interaction effects can be expressed in terms of a matrix \mathbf{J} , where each entry $J_{k\ell}$ 29
30 indicates how an agent in group k is influenced by the expected average action in group ℓ . 30

31 I study equilibrium properties of the model under noncooperative decisionmaking. These 31
32 properties are well-studied in settings with uniform complementarities, where the spillover 32

1 effects are all positive; e.g., consider [Glaeser, Sacerdote, & Scheinkman \(1996, 2003\)](#),
2 [Brock & Durlauf \(2001\)](#), [Calvo-Armengol, Patacchini, & Zenou \(2009\)](#), and [Cabrales,](#)
3 [Calvo-Armengol, & Zenou \(2011\)](#), among others. In these settings, interactions produce
4 social multipliers, which can lead to large differences in aggregate outcomes across similar
5 populations. Also, models with uniformly positive interactions can exhibit multiple, locally
6 stable equilibria, which are well-ordered in the sense that they form a complete lattice.¹

7 Overall, less is known about the equilibrium properties of models with both positive and
8 negative interaction effects. One issue is that the tools economists use to analyze settings
9 with complementarities do not readily transfer to settings with strategic substitutabilities,
10 i.e., negative spillovers. For example, models with negative interaction effects may fail to
11 possess a pure strategy Nash equilibrium. Moreover, even when equilibria do exist, it may
12 be that none of them are dynamically stable, which means they are unlikely to be observed
13 by researchers.² Throughout the networks literature, e.g., [Ballester et al. \(2006\)](#), [Galeotti](#)
14 [et al. \(2010\)](#), [Bramoullé et al. \(2014\)](#), contraction mapping arguments are often used to
15 prove that there exists a unique, stable equilibrium if the spillover effects are sufficiently
16 weak. I derive a similar, albeit slightly more general, result for my model using an index
17 theorem. However, the bulk of my theoretical analysis focuses on settings with strong in-
18 teraction effects where unstable equilibria arise and outcomes become harder to interpret.

19 I identify two conditions—one strictly weaker than the other—each guaranteeing the
20 existence of a locally stable equilibrium in contexts with strong interaction effects. Both
21 conditions have meaningful economic interpretations. In either a strict or weak sense, they
22 require that agents are not repelled by their own social group. Through these conditions, I
23 show how key results in the complementarities literature extend to a broader class of models

25 ¹The microeconomic foundations for models of complementarity are largely developed by Donald Topkis (see
26 [Topkis \(1998\)](#) for a summary), [Vives \(1990\)](#), and [Milgrom & Roberts \(1990\)](#). This work leverages key results
27 from lattice theory, e.g., Tarski's fixed point theorem, to study equilibrium behavior in settings with uniform
28 complementarities. [Cooper & John \(1988\)](#) also provide an early discussion about the role of complementarity in
29 economics. Additional contributions have been made by [Milgrom & Shannon \(1994\)](#) and [Athey \(2001, 2002\)](#).

30 ²[Jackson & Zenou \(2015\)](#) describe several unresolved questions regarding how and why network-based mod-
31 els with substitutabilities behave differently from models with uniform complementarities. Of course, strategic
32 substitutabilities appear frequently throughout economics, albeit outside of the social interactions literature; e.g.,
see [Bramoullé & Kranton \(2007\)](#), [Bramoullé et al. \(2014\)](#), and [Elliott & Golub \(2019\)](#) for recent contributions.

1 that involve substitutabilities. I start by establishing necessary and sufficient conditions for 1
2 multiple equilibria to exist. I show that multiplicity depends on a single parameter: the 2
3 spectral radius of the Jacobian matrix of the equilibrium condition. This number measures 3
4 the intensity of the cycles in a network, thereby quantifying the cumulative strength of 4
5 the interaction effects. I find that a unique, stable equilibrium exists if the spectral radius is 5
6 below a certain threshold, while multiple, locally stable equilibria exist if the spectral radius 6
7 is above this threshold. Next, I consider the social welfare of agents at different equilibria. 7
8 I find that negative interaction effects introduce welfare trade-offs, such that it impossible 8
9 to jointly maximize the expected utility of agents in different social groups. Specifically, if 9
10 two agents who otherwise prefer the same action are negatively influenced by one another, 10
11 then the equilibrium that maximizes welfare for one does not do so for the other. This result 11
12 may be instructive for policy analysis since it carries implications about social inefficiency. 12

13 In the second half of the paper, I turn to the question of identification. That is, given data 13
14 on individual choices across multiple networks, what can be said about the role of social 14
15 interactions? The identification of network-based discrete choice models with incomplete 15
16 information is carefully studied by [Brock & Durlauf \(2001, 2007\)](#), and more recently by 16
17 [Bhattacharya et al. \(2023\)](#). [Paula \(2017\)](#) and [Kline & Tamer \(2020\)](#) give recent reviews of 17
18 this literature.³ My analysis builds on this work by tackling two key identification problems. 18

19 First, I address the issue of unobservable contextual effects. These factors may prevent 19
20 researchers from uncovering the role of social interactions. For example, when comparing 20
21 student achievement across different classrooms, it is hard to distinguish the impact of 21
22 peer effects from unobserved teacher quality or unknown class characteristics. Much of the 22
23 applied literature tends to rule out these unknown network factors.⁴ Alternatively, some 23
24 papers have used a panel to difference-out the fixed effects between time periods; e.g., see 24
25 [Hoxby \(2000\)](#) and [Brock & Durlauf \(2007\)](#). However, this approach relies on having access 25
26 to panel data, as well as the assumption that the model parameters do not change over time. 26

27 To handle the issue of unobservable network effects, I introduce a novel approach that 27
28 allows for the partial identification of social interactions. Importantly, this method places 28

29

30 ³[Blume et al. \(2011\)](#) also review issues related to identification of linear and nonlinear network-based models. 30

31 ⁴For example, [Brock & Durlauf \(2001, 2007\)](#) assume that the network effect is a constant linear function of 31
32 observed variables, whereas [Bhattacharya et al. \(2023\)](#) assume that it exhibits a specific linear factor structure. 32

1 no restrictions on the network-level determinants of people’s actions. Rather, it leverages a 1
2 panel structure that is already inherent in the model, whereby members of different social 2
3 groups interact in the same network. Since these agents face the same contextual factors, I 3
4 can control for network effects by comparing the outcomes of different social groups in a 4
5 given network. Using this approach, I recover the differences between any two interaction 5
6 effects, which I can then use to measure the amount of polarization in a network. This 6
7 technique of exploiting within-network variation in individual-level characteristics is new 7
8 to the social interactions literature. However, similar approaches have been proposed in 8
9 other areas of economics. For example, [Berry & Haile \(2022\)](#) use data about heterogeneity 9
10 within markets to reduce the number of instruments needed to estimate systems of demand. 10

11 Next, I address an issue that is common to most network-based models with incomplete 11
12 information, which is that the expected average outcomes in a network are never directly 12
13 observed. Instead, a researcher only sees the average actions among finitely-many agents. 13
14 Since these observed averages converge to the true expectations as the networks grow large, 14
15 previous research has treated this issue as an estimation problem rather than as a barrier to 15
16 identification; e.g., see [Bhattacharya et al. \(2023\)](#). These estimation strategies rely on dou- 16
17 ble asymptotic arguments, such that the number of networks and the size of each network 17
18 must both tend to infinity. I argue that these approaches are not advisable given that replac- 18
19 ing expectations with observed averages always leads to measurement error. Indeed, I show 19
20 that, in any setting with finite networks, the observed average outcomes are endogenous in 20
21 the model. Moreover, by failing to account for the endogeneity, most common estimation 21
22 methods would produce biased estimates—even as the number of networks tends to infinity. 22

23 I approach this issue as a classical measurement error problem ([Wooldridge, 2013](#), sec. 23
24 15.4). Specifically, I note that the observed average outcomes are noisy approximations of 24
25 the true expectations and are therefore endogenous. To correct for this endogeneity, I use 25
26 internal instruments. This procedure involves randomly partitioning the network into two 26
27 parts and then computing the sample average action within each part. I then treat one of 27
28 these sample averages as the endogenous variable and the other as an instrument. Since 28
29 each network is partitioned randomly, this method ensures that the model parameters are 29
30 identified even for small networks. I then define an IV estimator and prove it is consistent. 30
31 I demonstrate the efficacy of the estimation strategy by running Monte Carlo simulations. 31

1 Finally, I provide an empirical application of the model and identification strategy using 1
 2 the class size reduction experiment Project STAR. Prior research has used this data to eval- 2
 3 uate peer effects in classrooms, e.g., see [Boozer & Cacciola \(2001\)](#) and [Graham \(2008\)](#). 3
 4 However, I leverage the generalized interactions framework to study a new question: How 4
 5 do peer effects differ by gender? In the data, I find strong evidence of gender differences 5
 6 in peer effects. Specifically, I find that male and female students are both more likely to 6
 7 conform to members of their own gender than to members of the opposite gender. This 7
 8 empirical finding speaks to a longstanding literature about the role of gender and social 8
 9 pressure in academic achievement, e.g., see [Hoxby \(2000\)](#), [Lavy & Schlosser \(2011\)](#), and 9
 10 [Bostwick & Weinberg \(2022\)](#). While previous work has primarily focused on the effects of 10
 11 gender composition, I explore how students are directly affected by the expected achieve- 11
 12 ment of their male and female peers. Moreover, I perform this analysis while accounting 12
 13 for unobserved classroom-level determinants of student outcomes. My empirical findings 13
 14 illustrate the advantages of using a generalized interactions framework to learn about sys- 14
 15 tematic heterogeneity in social interactions. They also contribute to a growing literature 15
 16 about identity in economics; see [Charness & Chen \(2020\)](#) and [Shayo \(2020\)](#) for reviews. 16

17 This paper proceeds as follows. Section II describes the binary choice model. Section III 17
 18 explores the equilibrium properties of the model, such as existence, uniqueness, dynamic 18
 19 stability, and social welfare. Section IV considers alternative specifications and explores 19
 20 how the equilibrium properties generalize to a broader class of models. Section V explains 20
 21 the identification strategy, along with the estimation procedure and simulation results. Sec- 21
 22 tion VI presents the empirical application and key findings. Lastly, Section VII concludes. 22

24 2. A MODEL WITH GENERALIZED SOCIAL INTERACTION EFFECTS 24

25
 26 I study a binary choice model with interaction effects that vary on the basis of group iden- 26
 27 tity. This model extends Brock & Durlauf's (2001) binary choice framework by assuming 27
 28 that individuals are influenced differently—perhaps even negatively—by different people. 28
 29 For example, a person might seek to conform to the average behavior in certain groups, 29
 30 while distinguishing herself from others. Unlike previous work, I allow the distribution of 30
 31 idiosyncratic preferences to be nonparametric. This feature ensures that the equilibrium 31
 32 properties of the model will be robust under relatively loose functional form assumptions. 32

2.1. *Individual Preferences*

Consider a network of I agents, where each agent i belongs to one of K social groups. Every agent chooses a binary action $\omega_i \in \{0, 1\}$ at a common time. Let $\bar{\omega}^k$ be the average action in group k , and let $\bar{\omega}_{-i}^k$ be the average action for members of group k excluding i .

When forming decisions, people are influenced by the expected behavior of everyone else in their network. For any person i in group k , the utility from choosing an action ω_i is:

$$U_i(\omega_i|k) = v_k(\omega_i) + J_{kk}\omega_i E_i(\bar{\omega}_{-i}^k) + \sum_{\ell \neq k} J_{k\ell}\omega_i E_i(\bar{\omega}^\ell) + \epsilon_i(\omega_i). \quad (1)$$

Here, $E_i(\bar{\omega}_{-i}^k)$ and $E_i(\bar{\omega}^\ell)$ are i 's subjective expectations about $\bar{\omega}_{-i}^k$ and $\bar{\omega}^\ell$, respectively. The term $v_k(\omega_i)$ is the private utility associated with a choice, which may vary by group membership. Finally, $\epsilon_i(\omega_i)$ is an idiosyncratic preference that is independent across agents.

Under this framework, utility exhibits proportional spillovers, so there is a multiplicative interaction between an agent's choice and the expected average choice in every group.⁵

Each $J_{k\ell}$ captures how much members of group k seek to conform to the mean choice in ℓ .

Since the action is binary, I can write $v_k(\cdot)$ as an affine function: $v_k(\omega_i) = h_k\omega_i + \eta_k$. Also, I can write $\epsilon_i(\omega_i) = \varepsilon_i\omega_i + \xi_i$ without loss of generality, where ε_i and ξ_i are random coefficients. Note that h_k parameterizes the deterministic private utility bias toward $\omega_i = 1$ for an agent in group k , while ε_i captures the agent's idiosyncratic payoff from this action.

I assume that the idiosyncratic payoffs ε_i may be distributed differently in every group:

$$P(\varepsilon_i \leq z|k) = F_{\varepsilon|k}(z), \quad \text{for } k = 1, \dots, K, \quad (2)$$

where $F_{\varepsilon|k}(\cdot)$ is continuously differentiable, symmetric about zero, and has positive density everywhere. I make no further parametric assumptions about these distributions. So, this framework applies for a variety of empirical specifications, e.g., logistic or Gaussian errors.

Three quantities are especially important for characterizing the model. First, there is a vector $h = (h_1, \dots, h_K)'$ of private utility terms, which specifies each group's intrinsic preference over the two actions. Second, there is a vector of distribution functions $\{F_{\varepsilon|k}\}_{k=1}^K$, which determine how likely it is that any idiosyncratic payoff is realized in each group.

⁵The spillover $J_{k\ell}\omega_i E_i(\bar{\omega}^\ell)$ may also be generated from a quadratic conformity effect $-\frac{1}{2}J_{k\ell}[\omega_i - E_i(\bar{\omega}^\ell)]^2$ as studied by [Bernheim \(1994\)](#). See [Brock & Durlauf \(2001\)](#) and [Blume et al. \(2015\)](#) for additional discussion.

Third, there is a matrix $\mathbf{J} \in \mathbb{R}^{K \times K}$ containing all the social interaction effects:

$$\mathbf{J} = \begin{bmatrix} J_{11} & J_{12} & \cdots & J_{1K} \\ J_{21} & J_{22} & \cdots & J_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ J_{K1} & J_{K2} & \cdots & J_{KK} \end{bmatrix}. \quad (3)$$

Throughout this paper, I refer to \mathbf{J} as the *interaction matrix*. It may also be interpreted as the adjacency matrix of a directed graph $(\mathcal{K}, \mathbf{J})$, where the nodes $\mathcal{K} = \{1, \dots, K\}$ represent different groups of individuals.⁶ The entries of \mathbf{J} specify the nature and intensity of the relationships between groups. Note that these effects may not be symmetric, so $J_{k\ell}$ need not equal $J_{\ell k}$ for $k, \ell \in \mathcal{K}$. In addition, any of the interaction effects could be negative.

2.2. Equilibrium under Noncooperative Decisionmaking

When analyzing this model, I focus on (pure strategy) Bayesian Nash equilibria where agents act noncooperatively. In other words, agents do not coordinate with one another when making decisions. Each agent i in group k chooses the action $\omega_i = 1$ with probability:

$$P(\omega_i = 1|k) = F_{\varepsilon|k} \left(h_k + J_{kk} E_i(\bar{\omega}_{-i}^k) + \sum_{\ell \neq k} J_{k\ell} E_i(\bar{\omega}^\ell) \right). \quad (4)$$

Since ω_i takes values in the set $\{0, 1\}$, the expected action $E(\omega_i|k)$ also equals $P(\omega_i = 1|k)$.

I assume that everyone has rational expectations about other people's choices. So, while agents cannot directly observe the actions of others, they do correctly infer these actions in expectation, i.e., $E_i(\omega_j|k) = E(\omega_j|k)$ for all agents i and j , and all groups k . By symmetry of the expected choice equations, it also follows that $E(\omega_i|k) = E(\omega_j|k)$ for all i, j , and k .

An equilibrium is given by the expected average choices $\{E(\bar{\omega}^k)\}_{k=1}^K$ that are consistent with individually optimal decisions. Letting m^{k*} denote $E(\bar{\omega}^k)$, I write this condition as:

$$m^{k*} = F_{\varepsilon|k} \left(h_k + \sum_{\ell=1}^K J_{k\ell} m^{\ell*} \right), \quad \text{for } k = 1, \dots, K. \quad (5)$$

A fixed point solution $m^* = (m^{1*}, \dots, m^{K*})$ to this system of equations is an equilibrium.

⁶For linear-in-means models, this matrix is sometimes called a ‘‘sociomatrix’’ (e.g., see [Blume et al., 2015](#)).

3. EQUILIBRIUM PROPERTIES OF THE MODEL

I now examine various properties of an equilibrium, such as existence, dynamic stability, and uniqueness. I also explore the implications of equilibrium outcomes for social welfare.

3.1. Existence

Since the functions $\{F_{\varepsilon|k}\}_{k=1}^K$ are continuous and the support of m^* is $[0, 1]^K$, Brouwer's fixed point theorem implies that there is at least one solution to the equilibrium condition.

Property 1. There exists at least one equilibrium m^* in the binary choice model.

3.2. Dynamic Stability

Another key equilibrium property is dynamic stability. This property asserts that any iteration on best response dynamics converges to an equilibrium. Formally, consider a dynamic analogue of the equilibrium system: $m_t^k = F_{\varepsilon|k}(h_k + \sum_{\ell=1}^K J_{k\ell} m_{t-1}^\ell)$ for $k = 1, \dots, K$. An equilibrium m^* is *locally stable* if it is a limiting solution to this dynamical system, where the initial iterate m_0 lies within some sufficiently small neighborhood of m^* . Alternatively, an equilibrium is *unstable* if there exists a neighborhood of m^* such that, for any m_0 arbitrarily close to m^* , there is some eventual iterate m_t that lies outside of the neighborhood.

The question of local stability is fundamental to comparative statics. If a locally stable equilibrium is slightly perturbed, then the average outcomes in a network would return to that equilibrium. Meanwhile, if an equilibrium is unstable, then nearby outcomes would diverge from it. In practice, locally stable equilibria are the ones that a researcher observes, while unstable equilibria represent tipping points between different equilibrium outcomes.

To assess when an equilibrium is locally stable, I need to introduce some terminology. First, I define the *Jacobian matrix* of the right-hand-side of the equilibrium system (5) as:

$$\mathbf{D}(m^*) = \boldsymbol{\beta}(m^*)\mathbf{J}. \quad (6)$$

Here, $\boldsymbol{\beta}(m^*) = \text{diag}[f_{\varepsilon|1}(h_1 + \sum_{\ell=1}^K J_{1\ell} m^{\ell*}), \dots, f_{\varepsilon|K}(h_K + \sum_{\ell=1}^K J_{K\ell} m^{\ell*})]'$ equals a diagonal matrix of densities $f_{\varepsilon|k}(h_k + \sum_{\ell=1}^K J_{k\ell} m^{\ell*})$, each representing the likelihood that someone in group k is close to indifferent between the two choices at an equilibrium m^* . So, $\mathbf{D}(m^*)$ equals the interaction matrix \mathbf{J} where each row k is weighted by the expected

fraction of group k that is near indifferent at m^* . Since the densities $f_{\varepsilon|k}$ are strictly positive, the entries $D_{k\ell}(m^*)$ of the Jacobian matrix have the same signs as the spillover effects $J_{k\ell}$.

Next, I define the *spectral radius* of a Jacobian matrix $\rho(\mathbf{D}(m^*))$ as the largest eigenvalue of this matrix in absolute value. Formally, for any square matrix \mathbf{A} , this quantity equals:

$$\rho(\mathbf{A}) = \max \{ |\lambda| : \lambda \text{ is an eigenvalue of } \mathbf{A} \}. \quad (7)$$

A matrix \mathbf{A} is convergent, in the sense that $\lim_{t \rightarrow \infty} \mathbf{A}^t = \mathbf{0}$, if and only if $\rho(\mathbf{A}) < 1$. So, a larger spectral radius corresponds to a more expansive matrix. In this model, $\rho(\mathbf{D}(m^*))$ will measure the collective strength of the social interaction effects within and across groups. As seen through the next property, this quantity also governs the local stability of equilibria.

Property 2. Any m^* is locally stable if $\rho(\mathbf{D}(m^*)) < 1$ and is unstable if $\rho(\mathbf{D}(m^*)) > 1$.

This property almost gives a necessary and sufficient condition for local stability. It falls short only when $\rho(\mathbf{D}(m^*))$ equals one, which takes place with probability measure zero.⁷

3.3. Uniqueness

This model has the potential to exhibit multiple equilibria. In other words, for some fixed parameter values, there may be multiple solutions to the equilibrium equations. In this subsection, I examine what social environments lead to uniqueness versus multiplicity.

3.3.1. Social Environments with a Unique Equilibrium

When is there only one equilibrium? This situation arises whenever the social interaction effects are not strong enough to generate any unstable equilibria. To see why, consider the next property, which follows from Sard's Theorem and the Poincaré-Hopf Index Theorem.

Property 3. For almost all $\{F_{\varepsilon|k}\}_{k=1}^K$, the number of equilibria is finite and odd. Also, if there are d_s locally stable equilibria, then there are at least $d_s - 1$ unstable equilibria.

By Property 3, I obtain a sufficient condition for uniqueness. Namely, if all equilibria are locally stable, then there can only be one. This reasoning leads to the following corollary.

⁷When $\rho(\mathbf{D}(m^*)) = 1$, the stability of m^* is governed by the Hessian matrices of the equilibrium equations. However, for almost all $\{F_{\varepsilon|k}\}_{k=1}^K$, the Jacobian $\mathbf{D}(m^*)$ evaluated at any equilibrium m^* has no unit eigenvalues.

Corollary. There is a unique, locally stable equilibrium if $\rho(\mathbf{D}(m^*)) < 1$ at all equilibria.

Recall that $\rho(\mathbf{D}(m^*))$ measures the intensity of social interactions, weighted by the likelihood that agents are indifferent between the two choices. So, a unique equilibrium exists if the social interactions are relatively weak and/or if agents have strong private preferences.⁸

3.3.2. *Social Environments with Multiple Equilibria*

When is there more than one equilibrium? By Property 3, multiplicity can only occur in settings with strong interaction effects where unstable equilibria exist. However, strong interactions do not always imply multiplicity. In some cases, there is a unique equilibrium that is unstable.⁹ To interpret equilibrium behavior when there are strong interaction effects, I must first examine the role of strategic complementarity and substitutability in networks.

Negative Spillovers and Global Instability of Equilibria

If all the entries of the interaction matrix \mathbf{J} are non-negative, then the model exhibits strategic complementarity between all agents. In this case, each person's utility is a supermodular function of individual choices, which means that the marginal payoff from one's action (weakly) increases when anyone else chooses that same action. This type of model has important properties. Most notably, it almost always has a locally stable equilibrium.¹⁰

If the interaction matrix has negative entries, then these same equilibrium properties do not generally apply. In particular, certain interaction effects are incompatible with locally stable equilibria. This instability has an economic interpretation. It occurs when agents are strongly repelled by the choices in their own group. In these cases, the group can never settle on one average action because its members are always discontented with the outcome.

To better understand this point, assume there is one group in the network, i.e. $K = 1$, and suppose, for simplicity, that there is no private utility bias, i.e. $h = 0$. In this case, an

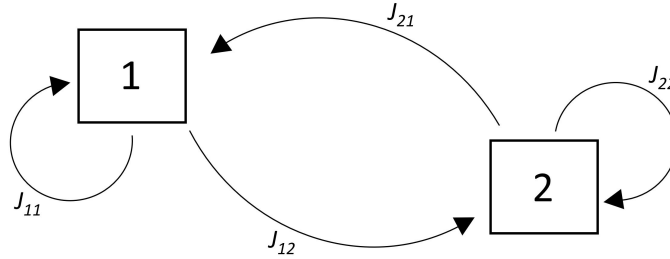
⁸By using the Poincaré-Hopf Index Theorem, I obtain a sufficient condition for uniqueness that is weaker than the conditions that are adopted in much of the literature, i.e., those that are implied by the Banach contraction mapping theorem. Note that Property 3 is consistent with a longstanding literature about the oddness of Nash equilibria in finite games; for example, see [Wilson \(1971\)](#), [Harsanyi \(1973\)](#), and [Kohlberg & Mertens \(1986\)](#).

⁹One example of this phenomenon occurs when the interaction matrix \mathbf{J} is symmetric and only has eigenvalues with non-positive real parts. For such interactions, the model always has a unique equilibrium m^* , which is locally stable when $\rho(\mathbf{D}(m^*)) < 1$ and is unstable when $\rho(\mathbf{D}(m^*)) > 1$. See the Online Appendix for justification.

¹⁰See [Milgrom & Roberts \(1990, 1994\)](#) and [Milgrom & Shannon \(1994\)](#) for properties of supermodular games.

equilibrium m^* is defined as the fixed point solution to $m^* = F_\varepsilon(Jm^*)$ (group subscripts are removed to ease notation). When $J < 0$, agents are repelled by the expected average action in the population. Therefore, if $E(\bar{\omega})$ is high, then people tend to prefer the low action. Conversely, if $E(\bar{\omega})$ is low, then most people prefer the high action. When $J < 0$ is sufficiently large in magnitude, the equilibrium behavior in the model becomes unstable.¹¹

FIGURE 1.—Social Interactions with Two Groups



When there are multiple groups, the role of negative interactions is more complicated. To illustrate why, suppose that there are now two groups, i.e. $K = 2$. The interactions in this case are depicted in Figure 1, where arrows indicate the direction of each effect. By the same reasoning as before, a stable equilibrium cannot exist if the within-group effects J_{11} and J_{22} are very negative. Additionally, consider what happens when $J_{12} > 0$ and $J_{21} < 0$. Members of group 1 seek to conform to the mean behavior in group 2, while members of group 2 want to distinguish themselves from group 1. These social influences, if they are strong enough, lead to self-contradictory behavior: if $E(\bar{\omega}^2)$ is high, then $E(\bar{\omega}^1)$ is high, which means that $E(\bar{\omega}^2)$ is low, and so forth. In this setting, all equilibria become unstable.

Social Interactions that Maintain Stable Equilibria

Going forward, it will be useful to classify what types of social interactions are conducive to stable behavioral outcomes. The discussion above suggests that global instability arises when preferences are self-contradictory in the sense that a population responds negatively to its own average choice. So, it makes sense to rule out these cases. The following condition captures the idea that all groups respond positively to their own average choices.

Assumption A. There is an invertible matrix \mathbf{B} where $\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$ has all non-negative entries.

¹¹Specifically, global dynamic instability occurs if $J < -f_\varepsilon^{-1}(0)$. For an explanation, see the Online Appendix.

This assumption states that the interaction matrix \mathbf{J} is similar to a non-negative matrix. When two matrices are similar, they represent the same linear operator under different bases. In this way, Assumption A implies that the model behaves similarly to one where all the interactions are non-negative, i.e. where there is uniform strategic complementarity.

On its face, Assumption A is hard to interpret. Nevertheless, it covers a range of contexts where agents do not wish to deviate from their own group. I highlight two special cases.

Example 1. The way that agents are affected by their own group depends on the cycles in a network. These cycles dictate how an agent's choice is reflected back onto itself via social interactions. Suppose that the product of interaction effects along any cycle is non-negative.

Assumption A.1. For any social group k , $J_{kj_1}J_{j_1j_2}\cdots J_{j_Mk} \geq 0$ for every $j_1, \dots, j_M \in \mathcal{K}$.

Under this condition, agents are never repelled by their own group. For example, with two groups, A.1 implies: $J_{11} \geq 0$, $J_{22} \geq 0$, and $J_{12}J_{21} \geq 0$. This restriction does not rule out negative interactions, but it requires that J_{12} and J_{21} have the same sign. In other words, between-group relations are mutual: agents either “agree to agree” or “agree to disagree”.

More broadly, A.1 holds if the indirect spillover effect $J_{km}J_{m\ell}$ has the same sign as the direct spillover effect $J_{k\ell}$, i.e. if $\text{sgn}(J_{k\ell}) = \text{sgn}(J_{km}J_{m\ell})$ for all $k, \ell, m \in \mathcal{K}$.¹² In this way, A.1 invokes the phrases “the friend of my friend is my friend” and “the enemy of my friend is my enemy”. This condition is rooted in balance theory (Cartwright & Harary, 1956).

To see how Assumptions A and A.1 are related, consider the following lemma. It reveals that A.1 is a special case of Assumption A where the change-of-basis matrix \mathbf{B} is diagonal.

Lemma 1. A.1 holds if and only if $\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$ is non-negative for some diagonal matrix \mathbf{B} .

Example 2. Suppose that the interaction matrix is symmetric, positive semi-definite. In this case, \mathbf{J} satisfies Assumption A since it can be diagonalized so that $\mathbf{J} = \mathbf{B}^{-1}\mathbf{\Lambda}\mathbf{B}$, where \mathbf{B} is orthogonal and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues of \mathbf{J} , which are all non-negative. One example in this class of matrices is the *diagonally dominant* matrix, which satisfies

¹²To motivate this restriction, consider a random effects version of the model where the interaction effects are unweighted. Specifically, let $J_{ij} \in \{-1, 1\}$, where J_{ij} varies across all agents i and j . Interpret \mathbf{J} as a matrix of average interactions $\mathbf{J}_{k\ell} = \mathbb{E}(J_{ij} | i \in k, j \in \ell)$. Then, the restriction need only hold in an average sense, where $\mathbb{P}(J_{i_0i_1}J_{i_1i_2} = J_{i_0i_2} | i_0 \in k, i_1 \in m, i_2 \in \ell) \geq 0.5$ for every $k, \ell, m \in \mathcal{K}$. See the Online Appendix for a proof.

$J_{kk} \geq \sum_{\ell \neq k} |J_{k\ell}|$ for $k = 1, \dots, K$. This condition implies that the cohesion in a group is strong relative to the between-group effects. So, agents tend to conform to their own group.

Deriving a Sufficient Condition for Multiple Equilibria

In settings with strong interaction effects, Assumption A is instrumental in characterizing equilibrium behavior. This assumption ensures that there is almost always one locally stable equilibrium. In addition, the locally stable equilibria will outnumber the unstable equilibria.

Property 4. Let Assumption A be satisfied. Then, for almost all distributions $\{F_{\varepsilon|k}\}_{k=1}^K$, there is exactly one more locally stable equilibrium than there are unstable equilibria.

To justify Property 4, I first give a proof in the case where \mathbf{J} is non-negative. This proof uses the Perron-Frobenius theorem, as well as other mathematical results related to non-negative matrices. I then show how the proof extends to settings where Assumption A holds (where \mathbf{J} is similar to a non-negative matrix). These arguments are laid out in the Appendix.

With Assumption A, I obtain a sufficient condition for multiplicity. Specifically, multiple equilibria arise when agents have a strong desire to conform to the average action in their own group. In these social environments, aggregate behaviors become self-reinforcing, which leads to multiple locally stable equilibrium outcomes from the same fundamentals.

Corollary. Suppose that Assumption A holds. If $\rho(\mathbf{D}(m^*)) > 1$ at some equilibrium m^* , then the model has multiple equilibria, and at least two equilibria are locally stable.

3.4. Comparisons of Aggregate Welfare across Social Groups

In this model, there is no strict Pareto ranking across equilibria since extreme realizations of the random payoff ε_i can dominate an agent's utility. So, to learn about social welfare, I examine the expected utility of agents in a group. In doing so, I can assess what equilibrium makes agents better off on average. In group k , the expected utility at equilibrium m^* is:

$$\mathbb{E} \left(\max_{\omega_i} U_i(\omega_i|k) | m^* \right) = \mathbb{E} \left(\max_{\omega_i} \left\{ h_k \omega_i + \eta_k + \sum_{\ell=1}^K J_{k\ell} \omega_i m^{\ell*} + \varepsilon_i \omega_i + \xi_i \right\} \right). \quad (8)$$

To evaluate expected utility, it helps to rescale the choices so that $\omega_i \in \{-1, 1\}$. This modification has no impact on agents' behavior; however, it makes welfare calculations easier to interpret. For example, it implies that—in absence of private preferences, i.e., if

1 $h_k = \varepsilon_i = 0$ for all i, k —an agent’s realized payoff if everyone selects the high action is 1
 2 the same as it would be if everyone selects the low action. In this way, there is no negative 2
 3 externality inherent to an equilibrium when all agents are ambivalent between their choices. 3

4 In general, welfare analysis in models of social interactions is highly sensitive to the 4
 5 way that utility is specified, even when different specifications yield the same expected 5
 6 choice functions. The reason is that welfare depends on the exact mechanisms that give 6
 7 rise to spillover effects (e.g., social learning, pressure to conform, or free-riding), not just 7
 8 the spillover effects themselves; see [Bhattacharya et al. \(2023\)](#) for a detailed explanation. 8

9 When the action takes values in the set $\{-1, 1\}$, the expected utility in group k equals: 9

$$10 \quad E\left(\max_{\omega_i} U_i(\omega_i|k)|m^*\right) = E\left(\left|h_k + \sum_{\ell=1}^K J_{k\ell}m^{\ell*} + \varepsilon_i\right|\right) + \eta_k + E(\xi_i|k). \quad (9) \quad 11$$

12
 13 As seen through the next result, the equilibrium that generates the highest expected utility is 13
 14 the one where most group members choose the same action, i.e., where $E(\bar{\omega}^k)$ is largest in 14
 15 magnitude. In addition, if agents are privately biased toward the high (low) action, then they 15
 16 tend to maximize their expected utility at the equilibrium where $E(\bar{\omega}^k)$ is highest (lowest). 16

17 **Property 5.** Let \mathcal{M}^* be the set of equilibria. For any k , $\operatorname{argmax}_{m^* \in \mathcal{M}^*} E(\max_{\omega_i} U_i(\omega_i|k)|m^*)$ 17

18 equals $\operatorname{argmax}_{m^* \in \mathcal{M}^*} |E(\bar{\omega}^k)|$. Also, there always exists some threshold T_k for which: 18

19 (i) If $h_k > T_k$, then $\operatorname{argmax}_{m^* \in \mathcal{M}^*} E(\max_{\omega_i} U_i(\omega_i|k)|m^*) = \operatorname{argmax}_{m^* \in \mathcal{M}^*} E(\bar{\omega}^k)$.¹³ 19

20 (ii) If $h_k < T_k$, then $\operatorname{argmax}_{m^* \in \mathcal{M}^*} E(\max_{\omega_i} U_i(\omega_i|k)|m^*) = \operatorname{argmin}_{m^* \in \mathcal{M}^*} E(\bar{\omega}^k)$. 20

21 When do different social groups favor the same equilibrium? To answer this question, 21
 22 it is useful to characterize when two groups are positively or negatively influenced by one 22
 23 another. These notions are generally hard to define because social influences reflect the 23
 24 sum of direct spillover effects (e.g., $J_{k\ell}$) and indirect spillover effects (e.g., $J_{km}J_{m\ell}$), which 24
 25 arise through interactions with other groups. Fortunately, using Assumption A.1, I can more 25
 26 easily interpret the nature of relations between groups. Consider the following definitions. 26
 27

28 **Definition 1.** k is *connected* to ℓ if $J_{kj_1}J_{j_1j_2} \dots J_{j_M\ell} \neq 0$ for some $j_1, j_2, \dots, j_M \in \mathcal{K}$. 28

29 ¹³Why would $T_k \neq 0$? In any group k , the payoff from choosing ω_i depends on h_k , as well as on h_ℓ for each 29
 30 group ℓ that influences k . So, even if $h_k > 0$, most agents in group k may still prefer the low action if (1) $h_\ell < 0$ 30
 31 for some ℓ that attracts k or if (2) $h_\ell > 0$ for some ℓ that repels k . Only if h_k is strong enough to overcome these 31
 32 external influences, i.e., if h_k is above some T_k , is $E(\max_{\omega_i} U_i(\omega_i|k)|m^*)$ maximized where $E(\bar{\omega}^k)$ is highest. 32

Definition 2. k is positively (negatively) influenced by ℓ if, for any positive integer M ,
 $J_{kj_1}J_{j_1j_2}\dots J_{j_M\ell} \geq (\leq) 0$ for all $j_1, j_2, \dots, j_M \in \mathcal{K}$, with at least one inequality strict.

Whenever group k is positively (negatively) influenced by group ℓ , its members seek to conform to (deviate from) the average behavior in ℓ . If all groups are connected, i.e., if \mathbf{J} is irreducible, then Assumption A.1 ensures that: (1) every group is positively influenced by itself and (2) any two groups are either positively or negatively influenced by one another.

When Assumption A.1 holds, the equilibria in the model are ordered in a distinctive way. This ordering determines how the relative welfare of equilibria varies across social groups.

Property 6. Suppose that Assumption A.1 holds, and consider any two groups k and ℓ .

(i) Suppose that k and ℓ are positively influenced by one another. Then the equilibrium at which $E(\bar{\omega}^k)$ is highest (lowest) is the same one at which $E(\bar{\omega}^\ell)$ is highest (lowest).

(ii) Suppose that k and ℓ are negatively influenced by one another. Then the equilibrium where $E(\bar{\omega}^k)$ is highest is the same equilibrium where $E(\bar{\omega}^\ell)$ is lowest, and vice versa.

To understand the implications of Property 6, consider any social environment where A.1 applies and there are multiple equilibria. In this setting, there are two *extremal equilibria* (call them \underline{m}^* and \bar{m}^*), where $E(\bar{\omega}^k)$ is either maximized or minimized for all groups k . As I show in the Online Appendix, both \underline{m}^* and \bar{m}^* are always locally stable. So, for appropriate initial conditions, a fixed-point iteration on (5) converges to an extremal equilibrium.

Taken together, Properties 5 and 6 reveal that social interactions can introduce welfare trade-offs, such that it is impossible to maximize aggregate welfare jointly in every group. In particular, if two groups are biased toward the same action, i.e., if $h_k > T_k$ and $h_\ell > T_\ell$, and if they are negatively influenced by one another, then they will maximize their expected utility at different equilibria. A similar trade-off arises when two groups are biased toward different actions, i.e., if $h_k > T_k$ and $h_\ell < T_\ell$, and are positively influenced by one another.

4. EXTENSIONS AND ALTERNATIVE NETWORK-BASED MODELS

4.1. Games on Networks

The properties in Section 3 are not exclusive to the binary choice framework. They have implications for a much broader class of models where agents interact in a network. To see how, consider a game with K players. Each player k chooses a_k from a compact action

1 space $A_k \in \mathbb{R}$. Given a profile of actions $a \in A_1 \times A_2 \times \dots \times A_K$, the player's best response 1
 2 is $a_k^* = q_k\left(\sum_{\ell=1}^K J_{k\ell} a_\ell\right)$, where $q_k(\cdot)$ is a non-decreasing function that maps from \mathbb{R} to A_k . 2

3 This game encompasses an abundance of economic models. For example, an action a_k 3
 4 could represent a person's investment in a public good, where everyone benefits from how 4
 5 much their neighbors contribute.¹⁴ Alternatively, a_k could be the output of a firm that com- 5
 6 petes in an oligopoly, where each firm's action influences the market price. These types of 6
 7 models are well-studied, and they both involve strategic substitutabilities between agents. 7

8 Another interpretation of this game is that each player is a community of individuals. 8
 9 In the binary choice model, the players are social groups, where the members of each 9
 10 group make one of two choices subject to social influences and idiosyncratic biases. Agents 10
 11 act noncooperatively, and a_k refers to the average choice within group k . This framework 11
 12 would also apply to a different type of model, in which the residents of a country or local 12
 13 institution make a collective decision. For example, consider modeling spillover effects in 13
 14 US state policy, where voters support more liberal or conservative agendas based on the 14
 15 laws enacted in other states. Here, a_k would represent the collective action taken in state k . 15

16 I focus on pure strategy Nash equilibria, which are defined by the action profiles a^* where 16
 17 no player k wishes to deviate from a_k^* . It is well known that an equilibrium exists if there are 17
 18 continuous best responses and compact, convex action spaces. However, even without these 18
 19 restrictions, an equilibrium would still exist if the interaction matrix \mathbf{J} satisfies Assumption 19
 20 A.1. To prove this result, I use Tarski's fixed point theorem, which ensures existence if 20
 21 the best responses are increasing, i.e., if \mathbf{J} is a non-negative matrix. I then show that this 21
 22 property extends to settings where A.1 holds.¹⁵ Crucially, this approach does not rely on 22
 23 continuous best responses. So, in the binary choice model, Assumption A.1 is enough to 23
 24 ensure existence even when the random utility component ε_i is not continuously distributed. 24

25
 26
 27 ¹⁴Bramoullé et al. (2014) study a public goods game that is nested within this framework. In their paper, players 27
 28 choose from an interval $[0, 1]$ and the best responses are $a_k^* = \max\{0, 1 - \delta \sum_{\ell \neq k} g_{k\ell} a_\ell\}$, where $\delta \in [0, 1]$ and 28
 29 $g_{k\ell} \in \{0, 1\}$ indicates whether two players k and ℓ are linked. Notice that their paper focuses exclusively on 29
 contexts with pure strategic substitutes, whereas my analysis allows for both positive and negative spillovers.

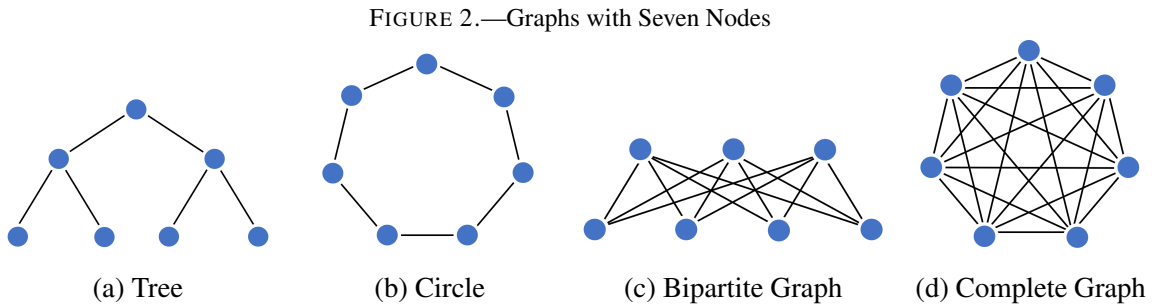
30 ¹⁵Specifically, under A.1, the best response functions map to an alternate system of equations with a non- 30
 31 negative interaction matrix. Any equilibrium in the original model corresponds to an equilibrium in a different 31
 32 model that has supermodular payoffs. Given this mapping, I use Tarski's theorem to prove that equilibria exist. 32

If the best responses are differentiable at each equilibrium a^* , then the spectral radius of the Jacobian matrix is all that is needed to determine uniqueness. In particular, let \tilde{a}^* be the equilibrium at which $\rho(\mathbf{D}(a^*))$ is greatest. If $\rho(\mathbf{D}(\tilde{a}^*)) < 1$, then \tilde{a}^* is the unique equilibrium. Also, whenever Assumption A holds, there are multiple equilibria if $\rho(\mathbf{D}(\tilde{a}^*)) > 1$.¹⁶

The equilibria of this game are generally not well-ordered. For example, they often do not form a lattice structure, which is useful for drawing conclusions about social welfare. As seen through Property 6, ensuring well-ordered equilibria can be achieved by imposing Assumption A.1. This condition ensures that there exist extremal equilibria. It also implies that there are trade-offs between different players in the network. Namely, if k and ℓ are negative influenced by one another, then a_k^* is highest wherever a_ℓ^* is lowest, and vice versa.

4.2. Stable Network Structures

Given the implications of Assumptions A and A.1 for equilibrium behavior, it is worth understanding what network structures satisfy these conditions. Namely, what types of exclusion restrictions (entries of 0 in the interaction matrix) are associated with stable equilibrium outcomes? I explore this question by examining four canonical types of networks.



Example 1 (Tree). If the links between agents are sparse, then Assumptions A and A.1 are likely to hold even in very large networks. To see why, consider a tree with K nodes. This

¹⁶One convenient feature of this game is that each player's best response either (weakly) increases or decreases monotonically in another player's action. So, if the best responses are differentiable, then the entries of \mathbf{D} retain the same sign for all values of a . This assumption may be weakened. Namely, suppose that a player's current action determines whether they conform to or deviate from others. Then the characterization of equilibrium properties requires checking whether Assumptions A and A.1 are satisfied locally in certain regions of the support of a .

network is used to study peer effects in social hierarchies or firm interactions in vertical production networks. Trees also encompass two common types of network structures: lines and stars. Since a tree has no cycles, any walk to and from the same node requires retracing the same edges. So, if the interactions are symmetric or even *weakly mutual*, in the sense that $J_{k\ell}, J_{\ell k} \geq 0$ or $J_{k\ell}, J_{\ell k} \leq 0$ for all $k, \ell \in \mathcal{K}$, then Assumption A.1 is always satisfied.¹⁷

Example 2 (Circle). Consider a circle with K nodes. This network is used to study domino effects that arise when agents only interact with close contacts, rather than with the entire population; e.g., see Ellison (1993). Suppose that the interactions are weakly mutual, and let d_{edge} be the number of edges involving negative interactions. If d_{edge} is even, then A.1 always holds. If d_{edge} is odd, then each agent is negatively influenced by herself. In this case, Assumptions A and A.1 fail. Therefore, the stability of equilibria for this type of network depends on the number of agents, which determines the parity of negative interactions.

Example 3 (Bipartite Graphs). Consider an environment with pure strategic substitutes, i.e., let $J_{k\ell} \leq 0$ for any $k, \ell \in \mathcal{K}$. In this setting, A.1 holds if and only if the corresponding graph is bipartite. More broadly, A.1 holds if and only if agents can be partitioned into two teams, such that negative interactions only exist between members of different teams. Note that, even if A.1 fails, Assumption A may still be used to ensure equilibrium stability.¹⁸

Example 4 (Complete Graphs). Consider a network where all agents are linked. As there are no exclusion restrictions, A.1 is unlikely to hold if there are many agents in the network. For example, if the interactions are weakly mutual, then the fraction of complete graphs for which A.1 holds is $1/2^{\gamma_K}$, where $\gamma_K = (K-1)(K-2)/2$ for $K \geq 2$. With three agents, this fraction is $1/2$. With seven agents (depicted in Figure 2d), it is $1/32,768$. So, Assumption A.1 is more likely to apply in settings with very few agents and/or where the

¹⁷If the interactions are not weakly mutual, then $J_{k\ell}J_{\ell k} < 0$ for some $k, \ell \in \mathcal{K}$, which means that A.1 fails. This scenario is strategically similar to a matching pennies game, which has no pure strategy Nash equilibrium.

¹⁸Assumption A typically holds if the within-group spillovers J_{kk} are positive and large relative to the between-group spillovers $J_{k\ell}$, $\ell \neq k$. For example, let the interaction matrix \mathbf{J} and the change-of-basis matrix \mathbf{B} equal:

$$\mathbf{J} = \begin{bmatrix} \delta & 1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

If $\delta \geq 3$, then $\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$ is a non-negative matrix, which means that the model will have a locally stable equilibrium.

1 interaction effects are fairly uniform. In the binary choice model, I take the latter approach 1
 2 by partitioning the network into a few subgroups, in which the spillover effects are constant. 2

4.3. Preferences over Network Composition 4

5
 6 Until now, I have assumed that agents are influenced by the expected average action in 6
 7 a group, regardless of which group comprises a larger share of the total population. I now 7
 8 consider an alternative setup, where utility depends on the expected composition of people 8
 9 who choose an action. Namely, suppose agents care about $E(k|\omega_i)$ instead of $E(\omega_i|k)$. 9

10 This reformulation may be used to study how social pressure affects network selection. 10
 11 For example, suppose that agents are choosing whether to enter a new environment, such as 11
 12 a school, and they care about what types of people they are likely to encounter there. This 12
 13 scenario invariably leads to negative interaction effects, since a preference that one group 13
 14 is in the majority is equivalent to a preference that other groups are in the minority. These 14
 15 preferences can be analyzed using the same techniques outlined in the previous sections. 15

16 Consider a model with two social groups: a and b .¹⁹ Let λ_a be the share of people in 16
 17 group a , and let $s_a(\omega_i)$ be the probability of being in group a given that one chooses ω_i . 17
 18 The utility from a choice ω_i depends on the composition of agents who make that choice. 18

$$19 \quad U_i(\omega_i|k) = v_k(\omega_i) + J_k s_a(\omega_i) + \epsilon_i(\omega_i), \quad \text{for } k \in \{a, b\}. \quad (10) \quad 20$$

21
 22 Under this framework, the parameter J_k indicates how much people in group k benefit from 22
 23 associating with members of group a . Both $v_k(\cdot)$ and $\epsilon_i(\cdot)$ are specified exactly as before. 23

24 In equilibrium, the expected composition of agents making a choice must be consistent 24
 25 with individually optimal decisionmaking. By Bayes' rule, any equilibrium should satisfy: 25

$$26 \quad s_a(\omega_i) = \frac{\lambda_a P(\omega_i|a)}{\lambda_a P(\omega_i|a) + (1 - \lambda_a) P(\omega_i|b)}, \quad \text{for } \omega_i \in \{0, 1\}, \quad (11) \quad 27$$

28
 29 where $P(\omega_i = 1|k) = F_{\epsilon|k}(h_k + J_k(s_a(1) - s_a(0)))$. By Brouwer's FPT, a solution exists. 29
 30

31
 32 ¹⁹I study the two-group case for simplicity, but this framework may be generalized to cases with many groups. 32

Uniqueness and dynamic stability of equilibria depend on the Jacobian matrix $\mathbf{D} \in \mathbb{R}^{2 \times 2}$ of the system (11). Letting $\beta_k = f_{\varepsilon|k}(h_k + J_k(s_a(1) - s_a(0)))$, I define this matrix so that:

$$\mathbf{D} = \begin{bmatrix} D_{11} & -D_{11} \\ -D_{22} & D_{22} \end{bmatrix}, \quad \text{where: } \begin{cases} D_{11} = s_a(0)[1 - s_a(0)] \left(\frac{J_a \beta_a}{\mathbb{P}(0|a)} - \frac{J_b \beta_b}{\mathbb{P}(0|b)} \right) \\ D_{22} = s_a(1)[1 - s_a(1)] \left(\frac{J_a \beta_a}{\mathbb{P}(1|a)} - \frac{J_b \beta_b}{\mathbb{P}(1|b)} \right) \end{cases} \quad (12)$$

Suppose that agents prefer to associate with their own group, i.e., $J_a \geq 0$ and $J_b \leq 0$. In this case, the matrix \mathbf{D} satisfies Assumption A.1. Hence, there is almost always a locally stable equilibrium—even if the spillover effects are strong enough to generate multiplicity.

If agents prefer not to associate with their own group, i.e., if $J_a < 0$ or $J_b > 0$, then the model may not have any locally stable equilibria. Global instability occurs when agents' behavior is self-undermining, such that a group always seeks to deviate from its own action. As an example, consider Allende (2019), who studies how peer effects drive student sorting in Peru. In her model, all parents prefer to send their children to schools that are made up of wealthy, high-achieving peers. If the peer effects are sufficiently weak, then a unique, stable equilibrium would exist. However, if low-performing students overwhelmingly seek to associate with their high-achieving peers, then a locally stable equilibrium may not exist.

In this model, the amount of diversity in a network is tied to stability of an equilibrium. To see why, note that the rows of the Jacobian matrix \mathbf{D} are scaled by $s_a(\omega_i)[1 - s_a(\omega_i)]$. This term grows larger as $s_a(\omega_i)$ approaches $1/2$, and it tends to zero as $s_a(\omega_i)$ approaches 0 or 1. So, if the agents who choose an action are more diverse, then the Jacobian is more expansive, and the realized equilibrium is more likely to be unstable. Conversely, if one group makes up the majority of people choosing an action, then the equilibrium is more likely to be locally stable. Of course, this property only holds for static models, where group membership is fixed. In the school choice example, some traits like student achievement are likely to change over time, while other traits, e.g., racial identity, may be fixed. One area for future work is to explore how the dynamics of identity affect the formation of networks.

5. IDENTIFICATION AND ESTIMATION OF ENDOGENOUS INTERACTION EFFECTS

I now discuss how the model can be brought to data. Specifically, I show how to use data on individual choices to learn about the social interaction effects $\{J_{kl}\}_{k,\ell}$. These effects are of primary interest to applied researchers as they have major implications for public policy.

5.1. Empirical Setting

Suppose that a researcher has data pertaining to many networks. In practice, a network could represent a household, a school, a neighborhood, or a workplace. Agents only interact with other people in their network, and these interactions may differ on the basis of social identity, e.g., by gender or occupation. In the data, the researcher either observes the entire network or a random sample that is drawn from each network. In addition, the researcher knows which network n an agent belongs to, as well as an agent's social identity k . Let $I_{n,k}$ denote the total number of agents in the sample belonging to network n and social group k .

To ensure a broad scope for application of the model, I assume that an agent's private utility is heterogeneous across networks. So, for any agent i with social identity k who resides in a network n , the private payoff from an action ω_i depends on: (1) individual-level factors ε_i , (2) identity-specific factors h_k , and (3) contextual network effects α_n . Crucially, each of these features is unobserved by the researcher. Hence, an agent i 's action ω_i equals:

$$\omega_i = \mathbb{1} \left\{ h_k + \alpha_n + J_{kk} E_i(\bar{\omega}_{n,-i}^k) + \sum_{\ell \neq k} J_{k\ell} E_i(\bar{\omega}_n^\ell) + \varepsilon_i \geq 0 \right\}, \quad (13)$$

where h_k , α_n , ε_i , and $\{J_{k\ell}\}_{k,\ell}$ are unknown coefficients. The model can also be adapted to incorporate observed individual-level covariates. However, since including covariates does not meaningfully change my analysis, I omit them for now and give details in the Appendix.

Throughout this section, I make two assumptions about the error structure in the model. First, I assume that the idiosyncratic payoffs ε_i and ε_j are independent for any two agents i and j within and across networks. Hence, there is no covariation between the error terms. Second, I assume that ε_i , conditional on group membership k , is independent of the network effect α_n . This condition implies that there is no self-selection into networks based on agents' unobserved idiosyncratic payoffs.²⁰ Both these assumptions are standard in the social interactions literature, and they can be relaxed to allow for selection on observables.²¹

Assumption B.1. (i) errors $\{\varepsilon_i\}_i$ are pairwise independent; (ii) $P(\varepsilon_i \leq z | k, \alpha_n) = F_{\varepsilon|k}(z)$.

²⁰This restriction is stronger than needed to prove identification. As shown by Manski (1988) and elaborated on by Horowitz (2009), full conditional independence could be replaced by a quantile independence restriction.

²¹Assume that $P(\varepsilon_i \leq z | W_n, \alpha_n, k) = F_{\varepsilon|W_n,k}(z)$ for an observed contextual factor W_n . Even if $F_{\varepsilon|W_n,k}$ does not equal $F_{\varepsilon|k}$, identification still follows by comparing networks with the same observable characteristics.

As before, agents act with incomplete information and consistent beliefs. While they do not see everyone's idiosyncratic payoff ε_i , they know about the generic parameters h_k , α_n , and $\{J_{k\ell}\}_{k,\ell}$, as well as the distributions $\{F_{\varepsilon|k}\}_{k=1}^K$. Therefore, they would accurately infer the expected average choices $\{E(\bar{\omega}_n^k)\}_{k=1}^K$. In equilibrium, the expected outcomes equal:

$$m_n^{k*} = F_{\varepsilon|k} \left(h_k + \alpha_n + \sum_{\ell=1}^K J_{k\ell} m_n^{\ell*} \right), \quad \text{for } k = 1, \dots, K \text{ and } n = 1, \dots, N. \quad (14)$$

As shown in Section 3, multiple equilibria can arise if the social interaction effects are sufficiently large in magnitude. When multiple equilibria exist, I assume that agents know which equilibrium is realized. So, there is no coordination involved in selecting equilibria.

5.2. Overview of the Identification Strategy

The model is point identified if there is only one set of parameter values $\{h_k\}_k$, $\{\alpha_n\}_n$, $\{J_{k\ell}\}_{k,\ell}$, and $\{F_{\varepsilon|k}\}_k$ that is consistent with the data under the equilibrium equations. Even if there is no self-selection into networks based on ε_i , two obstacles to identification remain. In the rest of this section, I describe both these obstacles and show how to overcome them.

Network-Level Unobservables

The first barrier to identification is the presence of unobservable network effects α_n , which impede my ability to learn about social interaction effects. To account for this issue, I propose a new method that allows for partial identification of social interactions while imposing no added assumptions on the network-level determinants of agent's choices. Specifically, I am able to difference-out the network fixed effects for members of two social groups residing in the same network. I outline this procedure in detail within the next subsection.

Unobserved Expected Average Choices

The second barrier to identification is that the expected averages $\{m_n^{k*}\}_{k,n}$ are never observed. Instead, a researcher only sees the average choices among a finite number of agents in a network. This consideration is an inherent feature of network-based models with incomplete information. However, in the literature, identification proofs often assume that $\{m_n^{k*}\}_{k,n}$ are already known; e.g., see Brock & Durlauf (2001) and Blume et al. (2015). The basis for this assumption is that the expectation m_n^{k*} may be consistently estimated by the

1 observed average choice $\bar{\omega}_n^k$ as the size of each network tends to infinity. So, a researcher's 1
 2 inability to see m_n^{k*} is treated as an estimation problem that is separate from identification. 2

3 I claim that it is not desirable to treat the expected average actions as known quantities. 3
 4 In many applications, the network sizes are small, which makes the observed averages poor 4
 5 approximations for the true expectations. Even in large networks, $\bar{\omega}_n^k$ is a noisy measure of 5
 6 m_n^{k*} . So, by replacing m_n^{k*} with $\bar{\omega}_n^k$ when estimating the model, the estimates are biased 6
 7 due to measurement error. This bias exists even as the number of networks tends to infinity. 7

8 I explain how to correct for the bias with internal instruments. In particular, I randomly 8
 9 split each network into two subsamples, and I use the average choice in one subsample as 9
 10 an instrument for the (endogenous) average choice in the other subsample. By construction, 10
 11 these averages are both noisy measures of m_n^{k*} , and the measurement errors do not depend 11
 12 on one another. So, this IV procedure is valid. I demonstrate that this approach leads to 12
 13 consistent estimates of the social interaction effects, even as the network sizes remain small. 13

14 Remarks on Reflection and Multiple Equilibria 14

15 In linear simultaneous equation models, a primary threat to identification is the *reflection* 15
 16 *problem* (Manski, 1993). This issue arises whenever the expectation m_n^{k*} is linearly depen- 16
 17 dent on α_n such that—even if α_n were a deterministic function of observed variables—it 17
 18 would be impossible to disentangle the role of social interactions from contextual network 18
 19 effects. To overcome reflection, researchers typically rely on exclusion restrictions, which 19
 20 are variables that only affect some agents in a network, while leaving others unaffected. 20

21 In my case, the reflection problem does not arise. Indeed, Brock & Durlauf (2001, 2007) 21
 22 show that this issue is not a threat to identification in the binary choice framework because 22
 23 the data always uncovers a nonlinear relationship between m_n^{k*} and α_n . This nonlinearity 23
 24 is inherent to the model and it acts like an exclusion restriction, which allows me to distin- 24
 25 guish between the social and contextual effects. So, after accounting for the other two issues 25
 26 raised above, the parameters in the model will be identified without further restrictions. 26

27 Finally, the possibility of multiple equilibria would not interfere with identification. The 27
 28 model is identified as long as there is a many-to-one mapping from the data to the pref- 28
 29 erence parameters. Moreover, when estimating the model, I do not need to solve the equi- 29
 30 librium equations (14) directly. So, my analysis is robust to any issues that stem from 30
 31 31

nonuniqueness; see [Bhattacharya et al. \(2023\)](#), sec. 5.5, for more detail.²² In most cases, it makes sense to assume that a realized equilibrium is locally stable; otherwise, it is unlikely to be observed. However, no part of my identification strategy relies on dynamic stability.

5.3. Identification with Known Expected Average Choices

As a first step, suppose that the expected average choices $\{m_n^{k*}\}_{k,n}$ are observed. In this case, the main obstacle to identification is the unobserved network effect α_n . To handle this issue, I propose to difference-out the fixed effects by contrasting the outcomes of two social groups in the same network. The intuition for this strategy is contained in the lemma below. I include the proof along with the result so that it can be more clearly interpreted.

Lemma 2. (*Sufficiency.*) For any network $n \in \{1, \dots, N\}$ and any social group $k \in \mathcal{K}$:

$$E(\omega_i | k, \alpha_n, \{m_n^{\ell*}\}_{\ell=1}^K) = E(\omega_i | k, \{m_n^{\ell*}\}_{\ell=1}^K). \quad (15)$$

Proof. Choose some social group \tilde{k} such that $\tilde{k} \neq k$. The expected average choice $m_n^{\tilde{k}*}$ is defined according to equation (14). Also, as $F_{\varepsilon|\tilde{k}}$ is strictly increasing, it is invertible. So:

$$\alpha_n = F_{\varepsilon|\tilde{k}}^{-1}(m_n^{\tilde{k}*}) - h_{\tilde{k}} - \sum_{\ell=1}^K J_{\tilde{k}\ell} m_n^{\ell*}. \quad (16)$$

By plugging this expression for α_n into the definition of $E(\omega_i | k, \alpha_n, \{m_n^{\ell*}\}_{\ell=1}^K)$, I find that:

$$E(\omega_i | k, \alpha_n, \{m_n^{\ell*}\}_{\ell=1}^K) = F_{\varepsilon|k} \left(h_k - h_{\tilde{k}} + \sum_{\ell=1}^K (J_{k\ell} - J_{\tilde{k}\ell}) m_n^{\ell*} + F_{\varepsilon|\tilde{k}}^{-1}(m_n^{\tilde{k}*}) \right). \quad (17)$$

The network effect cancels and $E(\omega_i | k, \alpha_n, \{m_n^{\ell*}\}_{\ell=1}^K)$ is a constant function of $\{m_n^{\ell*}\}_{\ell=1}^K$.

Q.E.D.

This lemma shows how the observed differences between individuals in a network can be used to control for unknown contextual effects. Consider any two agents i and j with different social identities (k and \tilde{k} , respectively) who both reside in the same network n . Since these agents share the same context, all network-level determinants of i 's choice

²²In discrete choice models with full information, nonuniqueness can affect identification (see [Tamer, 2003](#)).

are captured by j 's decision. Any difference between ω_i and ω_j is driven by idiosyncratic preferences (ε_i versus ε_j), as well as factors that relate to social identity (k versus \tilde{k}).

This framework offers a natural panel structure, allowing me to control for contextual effects by comparing the outcomes of different types of agents in the same network. As a result, I can recover the differences between two interaction effects, i.e., $\{J_{k_1\ell} - J_{k_2\ell}\}_{\ell=1}^K$ for all groups k_1 and k_2 . These quantities are economically meaningful as they specify how the members of two groups k_1 and k_2 differ in their desire to conform to another group ℓ .

I give two versions of my identification result. First, I provide conditions for semiparametric identification, where the error distributions $\{F_{\varepsilon|k}\}_{k=1}^K$ are known by the researcher. Then, I give conditions for nonparametric identification, where $\{F_{\varepsilon|k}\}_{k=1}^K$ are unknown. While the nonparametric version allows for more flexibility, it also requires that there is a lot of variation in the data. In both versions, identification is achieved for unobserved α_n .

Conditions for Semiparametric Identification

Before stating the identification result, I first write down the following assumption:

Assumption B.2. The contextual network effect α_n is a continuous random variable on \mathbb{R} .

This assumption corresponds to condition A.4 in [Brock & Durlauf \(2007\)](#), which is based on [Manski \(1988\)](#). It requires that α_n varies across networks and takes infinitely-many values. By implication, the equilibrium outcomes $\{m_n^{k*}\}_{k=1}^K$ would differ across networks. This heterogeneity is crucial for uncovering a nonlinear relationship between the expected average choices. It ensures there is ample variation to isolate the role of social interactions.

Theorem 1. Suppose that Assumptions B.1 & B.2 hold and that m_n^{k*} is observed for all networks n and all social groups k . If all the distributions $\{F_{\varepsilon|k}\}_{k=1}^K$ are known, then:

- (i) Without further assumptions, $\{h_{k_1} - h_{k_2}\}_{k_1, k_2}$ and $\{J_{k_1\ell} - J_{k_2\ell}\}_{k_1, k_2, \ell}$ are identified.
- (ii) If $\alpha_n = W_n' d$ for an observed vector W_n , then d , $\{h_k\}_k$, and $\{J_{k\ell}\}_{k, \ell}$ are identified.

This theorem has two parts. Part (i) leverages Lemma 2 by showing that the model is partially identified even if researchers have no prior knowledge about the network effects. In particular, I can recover the differences in identity fixed effects $h_{k_1} - h_{k_2}$, as well as the differences between the interaction effects $\{J_{k_1\ell} - J_{k_2\ell}\}_{\ell=1}^K$ for any social groups $k_1, k_2 \in \mathcal{K}$. This finding is new to the literature, and it is the main contribution of the theorem. Part (ii)

considers a special case—studied by Brock & Durlauf (2001, 2007) and others—where α_n is a constant linear function of observed variables. In this case, the model is fully identified.

Remark 1. Subtracting $J_{k_1 k_2} - J_{k_2 k_2}$ from $J_{k_1 k_1} - J_{k_2 k_1}$ gives a measure of polarization:

$$\delta_{k_1 k_2} = J_{k_1 k_1} + J_{k_2 k_2} - J_{k_1 k_2} - J_{k_2 k_1}. \quad (18)$$

This term $\delta_{k_1 k_2}$ quantifies how much agents seek to resemble their own group plus how much they seek to distinguish from the other group. By Theorem 1, it is identified in the data for any k_1 and k_2 even in settings where the contextual network effect α_n is unobserved.

This notion of polarization contributes to a burgeoning literature documenting how people’s choices are divided along cultural or political lines; e.g., see Boxell et al. (2022) and Bertrand & Kamenica (2023). In this literature, there is still no clear consensus on how to measure polarization as it is a relatively abstract concept. One benefit of my approach is that I am able to measure polarization in a way that is motivated by an economic model.²³

Remark 2. If the error distributions are known, e.g., if $\{F_{\varepsilon|k}\}_{k=1}^K$ are standard Gaussian or logistic, then—under appropriate normalizations—all the interaction effects $\{J_{k\ell}\}_{k,\ell}$ are identified up to scale. To see how, let $J_{\ell\ell} = 1$ for some group ℓ . Since $\{J_{k\ell} - J_{\ell\ell}\}_{k=1}^K$ are point identified, every coefficient $J_{k\ell}$ is known relative to $J_{\ell\ell}$. So, by setting the diagonal entries of the interaction matrix \mathbf{J} to one, each element $J_{k\ell}$ is recovered from the data.²⁴

Conditions for Nonparametric Identification

For nonparametric identification, I need an extra assumption. Specifically, there must be an exogenous, individual-level covariate that varies continuously on an unbounded support. Using this variation, I can recover each of the error distributions $\{F_{\varepsilon|k}\}_{k=1}^K$, which I then use to identify the rest of the model. This strategy closely follows Brock & Durlauf (2007).

To set ideas, I modify the choice equation to allow for exogenous covariates $X_i \in \mathbb{R}^r$. Let $\omega_i = \mathbb{1}\{X_i'c + h_k + \alpha_n + \sum_{\ell=1}^K J_{k\ell}m_n^{\ell*} + \varepsilon_i \geq 0\}$ where $P(\varepsilon_i \leq z|X_i, k, \alpha_n) = F_{\varepsilon|k}(z)$ and $P(X_i \leq x|k, \alpha_n) = P(X_i \leq x|k)$. Following Manski (1988), I impose the next assumption:

²³Under my framework, polarization is defined with respect to a particular choice ω_i . So, if social identity is more salient for some choices than others, then $\{\delta_{k_1 k_2}\}_{k_1, k_2}$ will depend on the decision that agents are making.

²⁴In many practical contexts, $\{F_{\varepsilon|k}\}_{k=1}^K$ are only known up to a scale parameter, e.g., the variance. In these cases, $\{J_{k_1 \ell} - J_{k_2 \ell}\}_{k_1, k_2, \ell}$ is already identified up to scale, and additional normalizations would not be advisable.

Assumption B.3. For all k , $\text{supp}(X|k)$ is not contained in a proper linear subspace of \mathbb{R}^r ; also, there is a component x_j of X —with a nonzero coefficient c_j —such that, for almost all values of $x_{-j|k}$, the distribution of $x_{j|k}$ given $x_{-j|k}$ has positive density everywhere.

Assumption B.3 ensures that there is enough variation in the data to recover the model parameters even if the distributions $\{F_{\varepsilon|k}\}_{k=1}^K$ are unknown. Consider the following result.

Theorem 2. Suppose that Assumptions B.1, B.2, & B.3 hold and let m_n^{k*} be observed for all networks n and social groups k . Then $(\{F_{\varepsilon|k}\}_{k=1}^K, c)$ is identified up to scale and:

- (i) The parameter vector $(\{h_{k_1} - h_{k_2}\}_{k_1, k_2}, \{J_{k_1 \ell} - J_{k_2 \ell}\}_{k_1, k_2, \ell})$ is identified up to scale.
- (ii) If $\alpha_n = W_n' d$ for an observed W_n , then $(d, \{h_k\}_k, \{J_{k\ell}\}_{k, \ell})$ is identified up to scale.

5.4. Identification with Unknown Expected Average Choices

In practice, researchers do not see the expected average choice m_n^{k*} . Instead, they only see the average outcome $\bar{\omega}_n^k$ among finitely-many individuals. This quantity may be interpreted as a noisy measure of the true expectation, such that $\bar{\omega}_n^k = m_n^{k*} + u_{nk}$ where u_{nk} has mean zero and $\sqrt{I_{n,k}} \times u_{nk}$ converges in distribution to $N(0, m_n^{k*}(1 - m_n^{k*}))$ as $I_{n,k} \rightarrow \infty$.

If a researcher simply replaces m_n^{k*} with $\bar{\omega}_n^k$ without accounting for measurement error, then the previous identification arguments break down. To understand this point, it helps to re-write an agent's choice equation in terms of the quantities that are observed in the data.

$$\omega_i = \mathbb{1} \left\{ h_k + \alpha_n + \sum_{\ell=1}^K J_{k\ell} \bar{\omega}_n^\ell + \tilde{\varepsilon}_i \geq 0 \right\}, \quad \text{where } \tilde{\varepsilon}_i = \varepsilon_i - \sum_{\ell=1}^K J_{k\ell} u_{n\ell}. \quad (19)$$

In this equation, the sample averages $\{\bar{\omega}_n^\ell\}_{\ell=1}^K$ are correlated with the random payoff $\tilde{\varepsilon}_i$. Indeed, even if agent i 's choice ω_i were excluded from the sample mean $\bar{\omega}_n^k$, it is still the case that $\text{Cov}(\bar{\omega}_n^\ell, \tilde{\varepsilon}_i) = -J_{k\ell} \times \text{Var}(u_{n\ell})$ for every $\ell \in \{1, \dots, K\}$. Hence, $\{\bar{\omega}_n^\ell\}_{\ell=1}^K$ is endogenous in the model, which implies that conditions (i) and (ii) of Assumption B.1 fail.

To correct for this endogeneity, I propose an IV strategy that uses internal instruments. This strategy involves carrying out two steps. First, in any network n and social group k , I randomly split the sample into two subsets: a and b . In practice, there are many ways to form these subsets (e.g., drawing Bernoulli random variables), and their sizes will not matter for identification. Second, I compute the average actions $\bar{\omega}_{n,a}^k$ and $\bar{\omega}_{n,b}^k$ in each subset.

By construction, the two sample averages $\bar{\omega}_{n,a}^k$ and $\bar{\omega}_{n,b}^k$ are both noisy measures of m_n^{k*} , where the measurement errors ($u_{nk,a}$ and $u_{nk,b}$, respectively) are independent of one another. This independence is a consequence of the incomplete information setting, where agents only respond to expectations rather than to the realized choices of others. Given this property, I can account for endogeneity in the model via IV estimation, where $\bar{\omega}_{n,b}^k$ is used as an instrument for $\bar{\omega}_{n,a}^k$. Before formalizing my result, I first motivate it with an example.

Example (Brock & Durlauf, 2001). In their original paper, the authors study the model:

$$m_n^* = \tanh(h + W_n' d + J m_n^*), \quad \text{for } n = 1, \dots, N, \quad (20)$$

where W_n is a vector of observed contextual factors and (h, d', J) are unknown parameters.²⁵ In the case where m_n^* is known, the authors prove that the model is identified if there is sufficient variation in W_n for uncovering the nonlinear dependence between m_n^* and W_n .

However, the expectation m_n^* is generally unknown. So, any researcher who wishes to apply the model to data would instead be relying on the observed average choice $\bar{\omega}_n$. Unless one accounts for measurement error, many common estimation strategies (e.g., maximum likelihood estimation or OLS regression) will lead to biased estimates. To illustrate this point, suppose that one wants to estimate the model via OLS using observed means $\{\bar{\omega}_n\}_n$. Defining $u_n = \bar{\omega}_n - m_n^*$ and $v_n = \tanh^{-1}(\bar{\omega}_n) - \tanh^{-1}(m_n^*)$, I re-write equation (20) as:

$$\tanh^{-1}(\bar{\omega}_n) = \tilde{h} + W_n' d + J \bar{\omega}_n + \tilde{\xi}_n, \quad \text{where: } \begin{cases} \tilde{h} &= h + E(v_n) \\ \tilde{\xi}_n &= -J u_n + v_n - E(v_n) \end{cases} \quad (21)$$

To ease notation, I define \tilde{m}_n^* , $\tilde{\omega}_n$, and \tilde{Y} as the residuals from a least squares regression of m_n^* , $\bar{\omega}_n$, and $\tanh^{-1}(\bar{\omega}_n)$, respectively, on the vector $(1, W_n')$. The OLS estimand for J is:

$$J^{\text{OLS}} = \frac{\text{Cov}(\tilde{\omega}_n, \tilde{Y})}{\text{Var}(\tilde{\omega}_n)} = J \times \frac{\text{Var}(\tilde{m}_n^*)}{\text{Var}(\tilde{m}_n^*) + \text{Var}(u_n)} + \frac{\text{Cov}(u_n, v_n)}{\text{Var}(\tilde{m}_n^*) + \text{Var}(u_n)}. \quad (22)$$

Observe that this estimand is “doubly-biased” because both the explanatory variable and the outcome variable are measured with error. As a result, $\text{Var}(u_n) \neq 0$ and $\text{Cov}(u_n, v_n) \neq 0$ in any setting with finite network sizes. So, the estimand J^{OLS} generally does not equal J .

²⁵Brock & Durlauf (2001) allow for individual-level covariates X_i in their model, which I omit to ease notation.

To overcome this issue, I use two-stage least squares. To do so, I start by partitioning each network into two parts (a and b), and then I compute the average choices $\bar{\omega}_{n,a}$ and $\bar{\omega}_{n,b}$. By construction, $\bar{\omega}_{n,b}$ is a valid instrument for $\bar{\omega}_{n,a}$ since $\bar{\omega}_{n,b} \perp u_{n,a}$. In addition, $\bar{\omega}_{n,b}$ satisfies instrument relevance because $\text{Cov}(\bar{\omega}_{n,a}, \bar{\omega}_{n,b}) = \text{Var}(m_n^*) \neq 0$. Given these properties, I can recover J via TSLS. This strategy involves: (1) regressing $\bar{\omega}_{n,a}$ on $(1, W_n', \bar{\omega}_{n,b})$ in the first stage to estimate all the fitted values $\mathbf{L}(\bar{\omega}_{n,a}|1, W_n', \bar{\omega}_{n,b})$, and then (2) regressing $\tanh^{-1}(\bar{\omega}_{n,a})$ on $(1, W_n', \mathbf{L}(\bar{\omega}_{n,a}|1, \bar{\omega}_{n,b}))$ in the second stage. By standard IV arguments, this approach yields consistent estimates of the parameters d and J even for small networks.

An IV Estimator to Recover Endogenous Interaction Effects

To formalize this identification strategy, I begin by defining $\bar{\omega}_{n,a}^k$ and $\bar{\omega}_{n,b}^k$, which are the average outcomes in each randomly-generated subset of a network n and social group k .²⁶ I also define $u_{nk,a} = \bar{\omega}_{n,a}^k - m_n^{k*}$ and $v_{nk,a} = F_{\varepsilon|k}^{-1}(\bar{\omega}_{n,a}^k) - F_{\varepsilon|k}^{-1}(m_n^{k*})$ as the measurement error that enters the model when a researcher tries to approximate m_n^{k*} and $F_{\varepsilon|k}^{-1}(m_n^{k*})$, respectively, using the average outcome $\bar{\omega}_{n,a}^k$. In the following theorem, I show that the interaction effects are recovered via linear IV, where $\bar{\omega}_{n,b}^k$ is used as an instrument for $\bar{\omega}_{n,a}^k$.

Theorem 3. Suppose that Assumptions B.1 & B.2 are satisfied. Then:

(i) For any two social groups $k_1, k_2 \in \{1, \dots, K\}$, define the IV estimand β_{k_1, k_2}^{IV} so that:

$$\beta_{k_1, k_2}^{IV} = \mathbf{E}(Z_n X_n')^{-1} \mathbf{E}(Z_n Y_n), \quad \text{where: } \begin{cases} Y_n &= F_{\varepsilon|k_1}^{-1}(\bar{\omega}_{n,a}^{k_1}) - F_{\varepsilon|k_2}^{-1}(\bar{\omega}_{n,a}^{k_2}) \\ X_n &= (1, \bar{\omega}_{n,a}^1, \dots, \bar{\omega}_{n,a}^K)' \\ Z_n &= (1, \bar{\omega}_{n,b}^1, \dots, \bar{\omega}_{n,b}^K)' \end{cases} \quad (23)$$

Then $\beta_{k_1, k_2}^{IV} = [h_{k_1} - h_{k_2} + \mathbf{E}(v_{nk_1, a}) - \mathbf{E}(v_{nk_2, a}), J_{k_1 1} - J_{k_2 1}, \dots, J_{k_1 K} - J_{k_2 K}]'$.

(ii) Let $\alpha_n = W_n' d$ for an observed vector W_n , and define the IV estimand β_k^{IV} so that:

$$\beta_k^{IV} = \mathbf{E}(Z_n X_n')^{-1} \mathbf{E}(Z_n Y_n), \quad \text{where: } \begin{cases} Y_n &= F_{\varepsilon|k}^{-1}(\bar{\omega}_{n,a}^k) \\ X_n &= (1, W_n, \bar{\omega}_{n,a}^1, \dots, \bar{\omega}_{n,a}^K)' \\ Z_n &= (1, W_n, \bar{\omega}_{n,b}^1, \dots, \bar{\omega}_{n,b}^K)' \end{cases} \quad (24)$$

For any social group k , the estimand β_k^{IV} is equal to $[h_k + \mathbf{E}(v_{nk, a}), d', J_{k 1}, \dots, J_{k K}]'$.

²⁶Here, an implicit assumption is that there are at least two agents within any network n and social group k . Otherwise, it will be impossible to further partition the sample, which means that $\bar{\omega}_{n,a}^k$ and $\bar{\omega}_{n,b}^k$ are undefined.

By the sample analogue principle, I can construct estimators for β_{k_1, k_2}^{IV} and β_k^{IV} that converge in probability to the desired parameters as the number of networks grows large.²⁷ By standard arguments, these estimators are capable of inference, and they may be used for hypothesis testing. Importantly, this estimation strategy yields consistent estimates even when the size of each network remains small. Therefore, I do not rely on double asymptotics.

TABLE I
PERFORMANCE OF IV ESTIMATORS FOR DIFFERENT NETWORK SIZES

Number of Networks	Agents per Network	Mean Squared Error					
		IV Estimator (i)		IV Estimator (ii)			
		$\widehat{J_{11}} - \widehat{J_{12}}$	$\widehat{J_{22}} - \widehat{J_{21}}$	\hat{J}_{11}	\hat{J}_{12}	\hat{J}_{21}	\hat{J}_{22}
N = 50	100	0.249	0.922	2.740	4.264	1.858	2.226
	500	0.041	0.102	0.153	0.610	0.154	0.377
	1000	0.017	0.044	0.070	0.254	0.067	0.155
	2000	0.008	0.020	0.032	0.112	0.032	0.068
N = 500	100	0.066	0.122	0.188	0.989	0.155	0.859
	500	0.006	0.013	0.017	0.076	0.016	0.060
	1000	0.002	0.005	0.007	0.028	0.007	0.020
	2000	0.001	0.002	0.003	0.012	0.003	0.008
N = 5000	100	0.049	0.042	0.023	0.152	0.014	0.064
	500	0.003	0.003	0.003	0.022	0.002	0.025
	1000	0.001	0.001	0.001	0.006	0.001	0.006
	2000	0.000	0.000	0.000	0.002	0.000	0.002

Notes. MSE's are computed across $M = 5000$ simulation draws. For each specification, I assume equal network sizes, and I construct X_n using 2/3 of the network sample, while using the remaining 1/3 to define the instrument Z_n . I assume that agents have logistic preferences, i.e., $\varepsilon_i|k \stackrel{i.i.d.}{\sim} \text{Logistic}(0, 1)$ for $k \in \{1, 2\}$. For more details about the data generating process, see the Online Appendix and replication code.

To evaluate the efficacy of the estimation method, I conduct Monte Carlo simulations. In Table 1, I compare the performance of the IV estimates as I vary the size of each network,

²⁷The IV estimator $(\frac{1}{N} \sum_{n=1}^N Z_n X_n')^{-1} (\frac{1}{N} \sum_{n=1}^N Z_n Y_n)$ converges to $E(Z_n X_n')^{-1} E(Z_n Y_n)$ as $N \rightarrow \infty$.

as well as the number of networks, in simulated data. I perform this analysis with both IV estimators: (i) computing $\{\hat{\beta}_{k_1, k_2}^{IV}\}_{k_1, k_2}$ when α_n is unknown, and then (ii) computing $\{\hat{\beta}_k^{IV}\}_k$ when $\alpha_n = W_n' d$ for some known vector W_n . For all these simulations, I fix $K = 2$.

The simulation results in Table 1 illustrate two key properties of the IV estimators. First, the estimators perform better in settings with larger networks. Indeed, as the network sizes grow, the sample average choices $\{\bar{\omega}_{n,a}^k\}_{k,n}$ will better approximate the expectations $\{m_n^{k*}\}_{k,n}$, reducing the amount of noise in the model. Second, the estimators become more precise as the number of networks N grows. This pattern is implied by Theorem 3, which shows that, for any fixed network sizes, I can consistently estimate the interaction effects.²⁸

6. APPLICATION: DIFFERENCES IN CLASSROOM PEER EFFECTS BY GENDER

In this section, I present an empirical application of my model and identification strategy. This application uses data from Project STAR, a large-scale education experiment that randomly assigned students to classrooms of different sizes and subsequently measured their test scores. In this context, I study the role of peer effects on academic performance, and I use the generalized interactions framework to assess how these spillovers differ by gender.

6.1. Description of the Data and Specification Choices

The Project STAR experiment took place among students in 79 Tennessee public schools who entered kindergarten in 1985. In each school, students and teachers were randomly assigned to three types of classrooms: small (13 to 17 students), large (22 to 25 students), and large with a teacher's aide. At the end of the school year, students took the *Stanford Achievement Tests in Math and Reading*. The scores from these exams were recorded as part

²⁸When implementing this IV strategy, the researcher has flexibility in choosing how to partition the networks. Namely, one can select how much of the sample to use to define $\{\bar{\omega}_{n,a}^k\}_{k,n}$ and $\{\bar{\omega}_{n,b}^k\}_{k,n}$, as long as this selection process does not depend on the individual-level determinants of choices. Given this freedom, there may be a way to split the sample in an “optimal way” in order to improve estimator performance. Also, if data on individual choices is available, then the researcher may obtain more precise estimates via maximum likelihood estimation. These issues are worth considering when estimating the model, and I leave them as a subject for future work.

1 of the study. In total, the experiment encompassed 6,325 students across 325 classrooms.²⁹ 1

2 For more details about the program design and implementation, see [Word et al. \(1990\)](#). 2

3 The data reports each student's raw test score, which I transform into a binary outcome 3
4 to make the setting suitable for my model. Since there was no pass/fail threshold for these 4
5 tests, I study a variety of outcomes, which include: scoring in the top 25%, 50%, and 75% 5
6 among Tennessee kindergarten students on both the math and reading exams. I estimate the 6
7 model separately for each of these outcomes to assess the role of peer effects on each one. 7

8 Under the protocols of the experiment, students within a school are unable to self-select 8
9 into classrooms. So, Assumption B.1 holds conditional on a school. I account for this con- 9
10 ditional independence by including school fixed effects when estimating the model. Fol- 10
11 lowing [Graham \(2008\)](#), I also give a version of the estimates where I restrict the sample to 11
12 schools that have three classrooms (one of each type). This exercise addresses the unlikely 12
13 possibility that students were nonrandomly assigned to classrooms of the same type. 13

14 To evaluate how peer effects differ by gender, I focus on a version of the model with two 14
15 social identities: male (m) and female (f). The Project STAR data is well-suited to study 15
16 these gender differences because male and female students are well-represented within and 16
17 across classrooms. Among participating students, there were 3,250 boys and 3,075 girls. 17

18 Student outcomes are specified according to equation (13). This framework allows for 18
19 unknown gender effects h_f and h_m , which may reflect prior socialization or developmental 19
20 differences between girls and boys. It also incorporates peer effects ($J_{ff}, J_{fm}, J_{mf}, J_{mm}$) 20
21 that vary based on gender identity. Furthermore, the model allows for unobserved class- 21
22 room characteristics α_n , which includes anything from teacher quality to the furnishings 22
23 and layout of the room. Throughout my analysis, I make no assumptions about α_n —only 23
24 that it is a continuous random variable. Therefore, my approach allows for a wide range of 24
25 classroom-level determinants that are typically not accounted for in applied work. 25

26 When estimating the model, I assume that the idiosyncratic payoff terms follow logistic 26
27 distributions, i.e., $\varepsilon_i|k \stackrel{\text{i.i.d.}}{\sim} \text{Logistic}(0, 1)$ for $k \in \{f, m\}$. This assumption relieves the burden 27
28
29

30 ²⁹The public-use data does not contain classroom identifiers. However, following [Boozer & Cacciola \(2001\)](#) 30
31 and [Graham \(2008\)](#), I can uniquely assign students to classrooms by matching on observed classroom variables. I 31
32 recover a sample of 6,248 students—among which 5,801 have non-missing test scores—across 321 classrooms. 32

of recovering these distributions nonparametrically, a task that can be quite challenging in practice. In my setting, the data is simply not rich enough for nonparametric estimation.³⁰

6.2. Empirical Results

Table 2 reports the IV estimates for $J_{ff} - J_{mf}$ and $J_{mm} - J_{fm}$ using the Project STAR data. To interpret these results, I first consider the descriptive evidence for gender differences in test scores. Averaging across all Project STAR classrooms, girls performed 1.01% better than boys on their combined math and reading scores. However, this difference does not reflect student outcomes at the classroom-level, where the gender gap favored boys nearly as many times as it favored girls. On average, the gender gap within a classroom was 2.83%—nearly three times higher than it was across classrooms. This disparity may suggest that factors at the classroom-level contributed to gender differences in student achievement. By analyzing these patterns under the generalized interactions framework, I can isolate the role of differential peer effects from other classroom-level factors that affect test scores.

The estimates in Table 2 indicate the presence of classroom peer effects that differ by gender. Both $J_{ff} - J_{mf}$ and $J_{mm} - J_{fm}$ are estimated to be positive, which implies that the pressure to conform is higher among peers of the same gender than it is for peers of opposite genders. In particular, a 1% increase in the expected fraction of girls scoring in the top 50% in math is estimated to raise the log-odds of achieving this outcome by 0.047 more for a female student than for a male student, on average. These estimates are roughly consistent across all six outcome variables. Moreover, there is no evidence to reject the hypothesis that $J_{ff} - J_{mf}$ equals $J_{mm} - J_{fm}$, i.e., that peer effect differences are symmetric.

To defend the validity of my approach, I perform two types of robustness checks. First, I test whether the gender-specific parameters $(h_f, h_m, J_{ff}, J_{fm}, J_{mf}, J_{mm})$ depend on observed classroom characteristics, such as the share of minority students, the poverty rate, the school location (urban versus rural), as well as a teacher’s education and experience. Such dependence would suggest that the model is misspecified, which could interfere with

³⁰Recall from Theorem 2 that recovering $\{F_{\varepsilon|k}\}_k$ requires significant variation of individual-level covariates in each network. For small to moderate sample sizes, these functions may be poorly approximated, which causes the rest of the estimates to be imprecise. If $\{F_{\varepsilon|k}\}_k$ are known a priori, then this step is entirely avoided. Hence, there is always a trade-off between making fewer parametric assumptions and achieving more precise estimates.

TABLE II
IV ESTIMATES FOR MATH AND READING TEST SCORES

	<i>Outcome Variable:</i>					
	Math			Reading		
	Top 25%	Top 50%	Top 75%	Top 25%	Top 50%	Top 75%
$J_{ff} - J_{mf}$	4.219 (3.109)	4.710*** (0.866)	5.116 (3.530)	4.659*** (0.230)	4.886*** (0.920)	4.444*** (0.151)
$J_{mm} - J_{fm}$	4.310 (2.747)	4.845*** (0.879)	5.509 (4.958)	4.556*** (0.302)	4.924*** (0.957)	4.674*** (0.175)
<i>Intercept</i>	0.002 (0.605)	0.220 (0.313)	0.024 (0.280)	-0.008 (0.085)	0.059 (0.301)	0.166 (0.108)
Number of Classrooms	321	321	321	321	321	321
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
$F_{(df1,df2)}$ 1st-Stage ($\bar{\omega}_n^m$)	8.17 _(2,60)	9.09 _(2,83)	4.71 _(2,55)	7.27 _(2,50)	6.43 _(2,86)	5.73 _(2,44)
$F_{(df1,df2)}$ 1st-Stage ($\bar{\omega}_n^f$)	5.41 _(2,60)	11.25 _(2,83)	9.53 _(2,55)	14.12 _(2,50)	9.22 _(2,86)	5.06 _(2,44)

Notes. Estimates are obtained by computing $\hat{\beta}_{f,m}^{IV}$, which corresponds to the estimand in equation (23). For implementation, I randomly split each classroom so that half the sample is used to form endogenous variables X_n , and the remaining half is used to form instruments Z_n . *p<0.1; **p<0.05; ***p<0.01.

identification. Although I cannot test for variation along unobserved dimensions, I find no evidence to indicate that gender-specific effects vary by observed classroom characteristics.

Second, I investigate how sensitive the IV estimates are to the way that the classrooms are partitioned. Specifically, I produce histograms that show how each of the estimates differ across a variety of random classroom partitions. I show that, under alternative partitions, the parameter estimates would not deviate too much from the results reported in Table 2. So, the estimates reliably demonstrate strong evidence of differential peer effects by gender.

7. CONCLUSION

The goal of this paper is to extend the theory of discrete choice and social interactions to more general settings, where agents are affected differently by different people. I analyze

1 how aggregate outcomes depend on the features of a network, and I consider externalities 1
 2 that arise under both positive and negative spillovers. Finally, I show how data may be 2
 3 used to recover the social interaction effects in the model, and I propose a new estimation 3
 4 strategy that can be used in settings with finite networks and unobserved contextual effects. 4

5 One contribution of this work is to classify when models with negative interactions have 5
 6 the same types of equilibrium properties as models with uniformly positive interactions. 6
 7 I present two conditions (Assumptions A and A.1), which both imply that agents are not 7
 8 repelled by their own actions. These conditions ensure that there is almost always a locally 8
 9 stable equilibrium, and they allow me to derive a sufficient condition for multiple equilibria. 9
 10 Under these conditions, I also draw conclusions about welfare that extend over a broad class 10
 11 of network-based models. To my knowledge, both these conditions are new to the literature. 11

12 In the second part of the paper, I explain how to tackle two key obstacles to identification. 12
 13 First, I show that the generalized interactions framework offers a panel structure, which 13
 14 may be leveraged to isolate the role of social interactions from unobserved network effects. 14
 15 Then, I demonstrate how to use internal instruments to correct for measurement error that 15
 16 enters the model when the expected average choices are not observed. This method assures 16
 17 that the model is empirically tractable in settings with small to moderate network sizes. 17

18 I implement my framework using data from Project STAR, and I find evidence that peer 18
 19 effects differ systematically on the basis of gender. These differences would not be captured 19
 20 by a model that assumes uniform interaction effects. Hence, this application highlights the 20
 21 advantages of the generalized interactions framework. Note that my analysis only scratches 21
 22 the surface on the economic implications of heterogeneous interactions. One could take the 22
 23 model further by allowing for full heterogeneity in spillover effects between all individuals. 23
 24 Therefore, while this paper contributes to the discussion about heterogeneous interactions 24
 25 in discrete choice environments, there is still much room for future research on this topic. 25

26 27 28 APPENDIX 28

29
30 *This appendix contains proofs of the main results in the paper. Additional discussion and 30
 31 findings are provided in a supplementary appendix that is intended for online publication. 31*

32 **Proof of Property 1** 32

1 *Proof.* Define $\mathcal{Q} : [0, 1]^K \rightarrow [0, 1]^K$ so that $\mathcal{Q}_k(m) = F_{\varepsilon|k}(h_k + \sum_{\ell=1}^K J_{k\ell}m^\ell)$ for $k \in \mathcal{K}$. 1
 2 As \mathcal{Q} is a continuous, self-mapping function on a non-empty, compact, convex set in \mathbb{R}^K , 2
 3 Brouwer's fixed point theorem ensures that there is a vector m^* that solves $m^* = \mathcal{Q}(m^*)$. 3

4 **Proof of Property 2**

Q.E.D. 4

5 *Proof.* Suppose $\rho(\mathbf{D}(m^*)) < 1$. For any $\epsilon > 0$, there is a matrix norm $\|\cdot\|$ for which 5
 6 $\|\mathbf{D}(m^*)\| \leq \rho(\mathbf{D}(m^*)) + \epsilon$. Define ϵ so that $\epsilon < 1 - \rho(\mathbf{D}(m^*))$. Then $\|\mathbf{D}(m^*)\| < 1$ for 6
 7 some matrix norm. The system (5) is a contraction at m^* under this norm, which implies: 7

$$8 \quad \|m_t - m^*\| = \|\mathcal{Q}(m_{t-1}) - \mathcal{Q}(m^*)\| \leq \kappa \|m_{t-1} - m^*\|, \quad 9$$

10 where $\kappa = [0, 1)$ for any m_{t-1} that lies in some sufficiently small neighborhood of m^* . 11
 12 Iterating on the inequality above implies $\|m_t - m^*\| \leq \kappa^t \|m_0 - m^*\|$ for $\lim_{t \rightarrow \infty} \kappa^t = 0$. 12

13 Suppose $\rho(\mathbf{D}(m^*)) > 1$. Since $\|\mathbf{D}(m^*)\| \geq \rho(\mathbf{D}(m^*))$ for any matrix norm $\|\cdot\|$, it must 13
 14 be that $\|\mathbf{D}(m^*)\| > 1$. By Henry (1981), Theorem 5.1.5., there is some $u > 0$ such that, for 14
 15 any $\delta > 0$, there is an initial iterate m_0 where $\|m_0 - m^*\| < \delta$ for which some future iterate 15
 16 m_t , where $t \geq 1$, satisfies $\|m_t - m^*\| \geq u$. So, by definition, m^* is an unstable equilibrium. 16

17 **Proof of Property 3**

Q.E.D. 17

18 *Proof.* Define the mapping $\mathcal{H} : [0, 1]^K \rightarrow \mathbb{R}^K$ so that $\mathcal{H}_k(m) = m^k - \mathcal{Q}_k(m)$ for $k \in \mathcal{K}$. 18
 19 By construction, m^* is an equilibrium if and only if $\mathcal{H}(m^*) = \mathbf{0}_K$. Since no equilibrium lies 19
 20 on the boundary of $[0, 1]^K$, I restrict attention to the interior $(0, 1)^K$.³¹ Let $\mathbf{D}_{\mathcal{H}}$ denote the 20
 21 Jacobian matrix of \mathcal{H} , and define a set $\mathcal{C} = \{m \in (0, 1)^K : \det(\mathbf{D}_{\mathcal{H}}(m)) = 0\}$. By Sard's 21
 22 Theorem, $\mathcal{H}(\mathcal{C})$ has Lebesgue measure zero. So, for almost all $\mathcal{H}(m)$ evaluated on the 22
 23 domain $(0, 1)^K$, the matrix $\mathbf{D}_{\mathcal{H}}(m)$ is invertible. Also, for any fixed $y = \mathcal{H}(m)$, almost 23
 24 all distributions $\{F_{\varepsilon|k}\}_{k=1}^K$ satisfy $y \notin \mathcal{H}(\mathcal{C})$. When applied to the case where $\mathcal{H}(m) = \mathbf{0}_K$, 24
 25 Sard's Theorem implies that $\mathbf{D}_{\mathcal{H}}(m^*)$ is almost always invertible on the full set of equilibria 25
 26 m^* . By the inverse function theorem, it is further guaranteed that each equilibrium is locally 26
 27 unique. Finally, since the set of equilibria is compact, it must have a finitely-many elements. 27

28 To prove that there is an odd number of equilibria, I rely on a version of the Poincaré- 28
 29 Hopf Index Theorem that is proven in Milnor (1965), Ch. 6. This theorem is stated below. 29
 30

31
 32 ³¹Since each $F_{\varepsilon|k}$ has positive density everywhere, $\mathcal{H}_k(m) < 0$ when $m_k = 0$ and $\mathcal{H}_k(m) > 0$ when $m_k = 1$. 32

Poincaré-Hopf Theorem (Milnor, 1965). Let v be a smooth vector field on the disk D^K that points outward on the boundary and has a finitely many isolated zeros $\{x_{(j)}^*\}_{j=1}^M$ where $\det(\mathbf{D}_v(x_{(j)}^*)) \neq 0$. Then $\sum_{j=1}^M \text{index}_{x_{(j)}^*}(v) = 1$, where $\text{index}_{x_{(j)}^*}(v)$ equals 1 if $\det(\mathbf{D}_v(x_{(j)}^*)) > 0$ and -1 otherwise.

This theorem applies in this context because the set of equilibria is defined on $(0, 1)^K$, which is diffeomorphic to an open disk. In addition, \mathcal{H} is a smooth vector field pointing outward at all boundary points, since $\lim_{m_k \rightarrow 0} \mathcal{H}(m) < 0$ and $\lim_{m_k \rightarrow 1} \mathcal{H}(m) > 0$ for all k . Finally, for almost all $\{F_{\varepsilon|k}\}_{k=1}^K$, it is the case that \mathcal{H} has finitely-many isolated zeros and that, at each of these zeros, the Jacobian has a nonzero determinant. So, the index theorem implies that $\sum_{j=1}^M \text{index}_{m_{(j)}^*}(\mathcal{H}) = 1$, which means that there is an odd number of equilibria.

Finally, suppose there are d_s locally stable equilibria. At each one of these equilibria m^* , the eigenvalues of $\mathbf{D}(m^*)$ all lie below unity in absolute value with probability one. Hence:

$$\det(\mathbf{D}_{\mathcal{H}}(m^*)) = \det(I - \mathbf{D}(m^*)) = \prod_{k=1}^K (1 - \lambda_k(m^*)) > 0,$$

where $\{\lambda_k(m^*)\}_{k=1}^K$ are the eigenvalues of $\mathbf{D}(m^*)$. By the index theorem, there must also be at least $d_s - 1$ equilibria m^* at which $\det(\mathbf{D}_{\mathcal{H}}(m^*)) < 0$. Each of these equilibria has at least one eigenvalue that exceeds unity. Hence, there are at least $d_s - 1$ unstable equilibria.

Proof of Lemma 1

Q.E.D.

Proof. “ \Rightarrow ” Suppose that \mathbf{J} is similar to a non-negative matrix \mathbf{A} by way of a diagonal change-of-basis matrix \mathbf{B} . Thus, there exists some diagonal \mathbf{B} for which $\mathbf{A} = \mathbf{B}\mathbf{J}\mathbf{B}^{-1} \geq \mathbf{0}$. Since \mathbf{B} is diagonal, the elements of \mathbf{A} can be expressed as $A_{k\ell} = B_{kk}J_{k\ell}/B_{\ell\ell}$, for all k, ℓ . So, for any selection of indices k and ℓ_1, \dots, ℓ_M in the set $\mathcal{K} = \{1, \dots, K\}$, it must be that:

$$0 \leq A_{k\ell_1}A_{\ell_1\ell_2} \cdots A_{\ell_M k} = \left(\frac{B_{kk}}{B_{\ell_1\ell_1}}J_{k\ell_1}\right) \left(\frac{B_{\ell_1\ell_1}}{B_{\ell_2\ell_2}}J_{\ell_1\ell_2}\right) \cdots \left(\frac{B_{\ell_M\ell_M}}{B_{kk}}J_{\ell_M k}\right) = J_{k\ell_1}J_{\ell_1\ell_2} \cdots J_{\ell_M k}$$

“ \Leftarrow ” Suppose that A.I holds. I first restrict attention to the case where \mathbf{J} is irreducible. For any k , I define $\{\gamma_{\ell}^k\}_{\ell=1}^K$ so that $\gamma_{\ell}^k = 1$ if k is positively influenced (see Definition 2) by ℓ , and $\gamma_{\ell}^k = -1$ otherwise. Fixing some index $k_0 \in \mathcal{K}$, I construct the matrix \mathbf{B} such that:

$$\mathbf{B} = \text{diag} \left[\gamma_1^{k_0} \cdots \gamma_K^{k_0} \right]'$$

Notice that \mathbf{B} is involutory, i.e. $\mathbf{B}^{-1} = \mathbf{B}$. Thus, for all (g, ℓ) , Assumption A.1 ensures that:

$$[\mathbf{B}\mathbf{J}\mathbf{B}^{-1}]_{k,\ell} = [\mathbf{B}\mathbf{J}\mathbf{B}]_{k,\ell} = \gamma_k^{k_0} \gamma_{\ell}^{k_0} J_{k\ell} = \gamma_{k_0}^k \gamma_{\ell}^{k_0} J_{k\ell} = \gamma_{\ell}^k J_{k\ell} = |J_{k\ell}|$$

1 Observe that $\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$ equals the absolute value of \mathbf{J} , which means $\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$ is non-negative. 1
 2 If \mathbf{J} is not irreducible, then this same reasoning applies to all irreducible blocks of \mathbf{J} . Thus, 2
 3 it must always be true that $\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$ is non-negative for some diagonal matrix $\mathbf{B} \in \mathbb{R}^{K \times K}$. 3

4 **Proof of Property 4**

Q.E.D. 4

5 *Proof.* First, let \mathbf{J} —and therefore $\mathbf{D}(m)$ for every $m \in [0, 1]^K$ —be a non-negative ma- 5
 6 trix. In addition, assume there exists an equilibrium $m^* \in (0, 1)^K$ satisfying $\rho(\mathbf{D}(m^*)) > 1$. 6

7 If $\mathbf{D}(m^*)$ is an irreducible matrix, then the Perron-Frobenius theorem guarantees that: 7

$$8 \quad \mathbf{D}(m^*)x = \rho(\mathbf{D}(m^*))x > x, \quad 8$$

9
 10 for some strictly positive $x \in \mathbb{R}_{++}^K$. It follows that $\mathbf{D}(m^*)\delta x > \delta x$ for any scalar $\delta > 0$. By 10
 11 taking the first-order Taylor approximation of $\mathcal{Q}(m^* + \delta x)$ about m^* , I obtain the following: 11

$$12 \quad \mathcal{Q}(m^* + \delta x) = \underbrace{\mathcal{Q}(x^*)}_{=x^*} + \underbrace{\mathbf{D}(x^*)\delta x}_{>\delta x} + h_1(m^* + \delta x)\delta x, \quad \text{where } \lim_{\delta \rightarrow 0} h_1(m^* + \delta x) = 0 \quad 12$$

13
 14 For sufficiently small δ , the vector $a = m^* + \delta x$, for $a \in (0, 1)^K$, satisfies $\mathcal{Q}(a) > a > m^*$. 14
 15 By an analogous argument, there is also some $b = m^* - \delta x$, which satisfies $\mathcal{Q}(b) < b < m^*$. 15
 16 Since no equilibrium lies on the boundary of $[0, 1]^K$, the following inequalities are satisfied: 16
 17

$$18 \quad \mathbf{0}_K < \mathcal{Q}(\mathbf{0}_K) < \mathcal{Q}(b) < b < m^* < a < \mathcal{Q}(a) < \mathcal{Q}(\mathbf{1}_K) < \mathbf{1}_K \quad 18$$

19
 20 Brouwer's fixed point theorem ensures that \mathcal{Q} has two more fixed points \underline{m}^* and \overline{m}^* , such 20
 21 that $\mathbf{0}_K < \underline{m}^* < m^*$ and $m^* < \overline{m}^* < \mathbf{1}_K$. If either of these equilibria is unstable, i.e., if 21
 22 $\rho(\mathbf{D}(\underline{m}^*)) > 1$ or $\rho(\mathbf{D}(\overline{m}^*)) > 1$, then these same arguments can be used to show that there 22
 23 exist two more equilibria: one that lies between the two unstable equilibria and another that 23
 24 lies between the unstable equilibrium and the boundary. Since there is a finite number of 24
 25 equilibria with probability one, I may conclude that there are almost always more stable 25
 26 equilibria than unstable equilibria. Together with Property 3, this result implies that there 26
 27 almost always exists one more locally stable equilibrium than there are unstable equilibria. 27

28 Next, consider the case where $\mathbf{D}(m^*)$ is a reducible matrix. If $\rho(\mathbf{D}(m^*)) > 1$, then the 28
 29 same must hold for some irreducible block of $\mathbf{D}(m^*)$. Let \mathcal{B} denote the set of indices within 29
 30 this block. Applying the Perron-Frobenius theorem to that block, there exists some vector x , 30
 31 satisfying $x_\ell > 0$ for $\ell \in \mathcal{B}$ and $x_\ell = 0$ otherwise, so that $\mathcal{Q}(m^* + \delta x) > m^* + \delta x$ for $\delta > 0$ 31
 32 sufficiently small. Setting $a = m^* + \delta x$, it follows that $m^* < a < \mathcal{Q}(a) < \mathcal{Q}(\mathbf{1}_K) < \mathbf{1}_K$. 32

By Brouwer's theorem, there must exist a fixed point of \mathcal{Q} between m^* and $\mathbf{1}_K$, and (by analogous arguments) a fixed point of \mathcal{Q} between $\mathbf{0}_K$ and m^* . So, just as in the previous case, there is always exactly one more stable equilibrium than there are unstable equilibria.

Finally, I show how this result, which is specific to monotone mappings, i.e., where \mathbf{J} is non-negative, can be extended to a certain class of non-monotone mappings. Suppose that Assumption A holds, i.e., let there be an invertible matrix \mathbf{B} such that $\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$ is non-negative. For $\mathcal{I} = [0, 1]^K$, define the mapping $\hat{\mathcal{Q}} : \mathbf{B}\mathcal{I} \rightarrow \mathbf{B}\mathcal{I}$ so that $\hat{\mathcal{Q}}(m) = \mathbf{B}\mathcal{Q}(\mathbf{B}^{-1}m)$. This mapping has a Jacobian matrix $\mathbf{D}_{\hat{\mathcal{Q}}}(m) = \mathbf{B}\mathbf{D}(\mathbf{B}^{-1}m)\mathbf{B}^{-1}$. Note that $\phi : \mathbb{R}^K \rightarrow \mathbb{R}^K$, where $\phi(m) = \mathbf{B}m$, is a bijective linear map. Therefore, $\phi(\cdot)$ is a *homeomorphism*, and it preserves all interior points.³² It follows that $\mathbf{B}\mathbf{D}(m)\mathbf{B}^{-1}$ is non-negative on $\text{int}(\mathcal{I})$ if and only if $\mathbf{D}_{\hat{\mathcal{Q}}}(m) = \mathbf{B}\mathbf{D}(\mathbf{B}^{-1}m)\mathbf{B}^{-1}$ is non-negative on $\text{Bint}(\mathcal{I})$, which equals $\text{int}(\mathbf{B}\mathcal{I})$. So, if Assumption A holds, then the matrix $\mathbf{D}_{\hat{\mathcal{Q}}}(m)$ is non-negative for every $m \in \text{int}(\mathbf{B}\mathcal{I})$.

To summarize, Assumption A implies that $\hat{\mathcal{Q}}(x) = \mathbf{B}\mathcal{Q}(\mathbf{B}^{-1}x)$ is monotonic on $\text{int}(\mathbf{B}\mathcal{I})$. Also, m^* is a fixed point of \mathcal{Q} if and only if $\mathbf{B}m^*$ is a fixed point of $\hat{\mathcal{Q}}$. To see why, write $\hat{\mathcal{Q}}(\mathbf{B}m^*) = \mathbf{B}\mathcal{Q}(m^*) = \mathbf{B}m^*$. Moreover, note that $\mathbf{D}(m)$ and $\mathbf{D}_{\hat{\mathcal{Q}}}(\mathbf{B}m)$ are similar matrices, since $\mathbf{D}_{\hat{\mathcal{Q}}}(\mathbf{B}m) = \mathbf{B}\mathbf{D}(m)\mathbf{B}^{-1}$, which implies that they share the same eigenvalues. In particular, their spectral radii are equivalent: $\rho(\mathbf{D}_{\mathcal{Q}}(m)) = \rho(\mathbf{D}_{\hat{\mathcal{Q}}}(\mathbf{B}m))$ for all $m \in \text{int}(\mathcal{I})$.

By this relationship, I can assess the existence, uniqueness, and local stability of fixed points of the mapping \mathcal{Q} by focusing on the (non-decreasing) mapping $\hat{\mathcal{Q}}$. While the fixed points of \mathcal{Q} and $\hat{\mathcal{Q}}$ may be ordered differently on their respective domains, their number and stabilities are the same. So, this property extends to models where Assumption A holds.

Proof of Property 5

Q.E.D.

Proof. Since ε_i is distributed symmetrically about zero, $\mathbb{E}(\max_{\omega_i} U_i(\omega_i|k)|m^*)$, which is defined in equation (9), strictly increases in $|h_k + \sum_{\ell=1}^K J_{k\ell}m^{\ell*}|$. In addition, since $|\mathbb{E}(\bar{\omega}^k)|$ is a strictly increasing function of $|h_k + \sum_{\ell=1}^K J_{k\ell}m^{\ell*}|$ in equilibrium, it must be that $\mathbb{E}(\max_{\omega_i} U_i(\omega_i|k)|m^*)$ is highest at the equilibrium where $|\mathbb{E}(\bar{\omega}^k)|$ is highest. Parts (i) and (ii) follow directly from the fact that $\lim_{h_k \rightarrow \infty} \mathbb{E}(\bar{\omega}^k) = 1$ and $\lim_{h_k \rightarrow -\infty} \mathbb{E}(\bar{\omega}^k) = -1$.

Proof of Property 6

Q.E.D.

³²A *homeomorphism* is a continuous bijection between two topological spaces that has a continuous inverse.

Proof. Suppose \mathbf{J} satisfies A.1. By Lemma 1, there exists some diagonal matrix \mathbf{B} for which $\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$ is non-negative. Without loss of generality, choose \mathbf{B} to be the matrix that is defined in the “ \Leftarrow ” part of the proof of Lemma 1. That is, fix some group k , and let \mathbf{B} be $\text{diag}[\gamma_1^k, \dots, \gamma_K^k]'$, where $\gamma_\ell^k = 1$ if k is positively influenced by ℓ , and $\gamma_\ell^k = -1$ otherwise.

As shown in the proof of Property 4, $\hat{\mathcal{Q}}(m) = \mathbf{B}\mathcal{Q}(\mathbf{B}^{-1}m)$ is monotonic on $\text{int}(\mathbf{B}\mathcal{I})$. Also, because $(\mathbf{B}\mathcal{I}, \leq)$ is a complete lattice, Tarski’s fixed point theorem ensures that the set of fixed points of $\hat{\mathcal{Q}}$ forms a complete lattice. So, $\hat{\mathcal{Q}}$ has greatest and least fixed points. Define $\mathbf{B}\bar{m}^*$ to be the greatest fixed point of $\hat{\mathcal{Q}}$. Then $E(\bar{\omega}^\ell)$ is maximal at \bar{m}^* whenever k is positively influenced by ℓ , and $E(\bar{\omega}^\ell)$ is minimal at \bar{m}^* otherwise. Let $\mathbf{B}\underline{m}^*$ be the lowest fixed point of $\hat{\mathcal{Q}}$. Then $E(\underline{\omega}^\ell)$ is minimal at \underline{m}^* whenever k is positively influenced by ℓ , and $E(\underline{\omega}^\ell)$ is maximal at \underline{m}^* otherwise. This argument applies for any choices of $k, \ell \in \mathcal{K}$.

Proof of Theorem 1

Q.E.D.

Proof. Choose any k_1 and k_2 in \mathcal{K} , and define the function $\nu_{k_1, k_2} : \mathbb{R}^K \rightarrow \mathbb{R}$ such that $\nu_{k_1, k_2}(m) = (h_{k_1} - h_{k_2}) + \sum_{\ell=1}^K (J_{k_1\ell} - J_{k_2\ell})m^\ell + F_{\varepsilon|k_2}^{-1}(m^{k_2})$. By equation (17), the expected average action $m_n^{k_1^*}$ equals $F_{\varepsilon|k_1}(\nu_{k_1, k_2}(m_n^*))$. Also, since the cumulative distribution function $F_{\varepsilon|k_1} : \mathbb{R} \rightarrow [0, 1]$ is strictly increasing on \mathbb{R} , the following equality is satisfied:

$$\begin{aligned} m_n^{k_1^*} &= F_{\varepsilon|k_1}(h_{k_1} - h_{k_2} + \sum_{\ell=1}^K (J_{k_1\ell} - J_{k_2\ell})m_n^{\ell*} + F_{\varepsilon|k_2}^{-1}(m_n^{k_2*})) \\ &= F_{\varepsilon|k_1}(\widehat{h_{k_1} - h_{k_2}} + \sum_{\ell=1}^K (\widehat{J_{k_1\ell} - J_{k_2\ell}})m_n^{\ell*} + F_{\varepsilon|k_2}^{-1}(m_n^{k_2*})) \end{aligned}$$

if and only if $(h_{k_1} - h_{k_2}) - (\widehat{h_{k_1} - h_{k_2}}) = \sum_{\ell=1}^K [(\widehat{J_{k_1\ell} - J_{k_2\ell}}) - (J_{k_1\ell} - J_{k_2\ell})]m_n^{\ell*}$. This property holds for all networks n . Also, since $\{F_{\varepsilon|k}\}_{k=1}^K$ are nonlinear, $\{m_n^{k*}\}_{k=1}^K$ are nonlinear functions of one another. Sufficient variation in $\{m_n^{k*}\}_{k=1}^K$ across networks ensures:

$$h_{k_1} - h_{k_2} = \widehat{h_{k_1} - h_{k_2}} \quad \text{and} \quad J_{k_1\ell} - J_{k_2\ell} = \widehat{J_{k_1\ell} - J_{k_2\ell}},$$

for all $\ell \in \mathcal{K}$. Also, since k_1 and k_2 are chosen arbitrarily, this result holds for all $k_1, k_2 \in \mathcal{K}$. The proof in the case where $\alpha_n = W'_n d$ relies on an analogous argument, and it also closely follows the proof of Proposition 1 in Brock & Durlauf (2007). I exclude it for this reason.

Proof of Theorem 2

Q.E.D.

Proof. To show that c , $\{h_{k_1} - h_{k_2}\}_{k_1, k_2}$, and $\{J_{k_1\ell} - J_{k_2\ell}\}_{k_1, k_2, \ell}$ are identified, I apply Corollary 5 of Proposition 2 in Manski (1988). This result not only allows me to recover the

parameters of interest, but it also guarantees that the error distributions $\{F_{\varepsilon|k}\}_{k=1}^K$ are identified up to scale. Ordinarily, recovering the error distributions would be unnecessary. However, in this setting, the presence of endogenous interaction effects means that $\{m_n^{\ell*}\}_{\ell=1}^K$ is functionally dependent on $\{F_{\varepsilon|k}\}_{k=1}^K$. To handle this issue, I take a two-step approach to identification: first I recover c and $\{F_{\varepsilon|k}\}_{k=1}^K$, then I use these quantities to recover the rest.

To start, I show that c and $\{F_{\varepsilon|k}\}_{k=1}^K$ are identified up to scale. Consider any group $k \in \mathcal{K}$. By Assumption B.3, there is some element x_j of X that varies continuously across \mathbb{R} . Without loss of generality, let $x_j = x_1$, and normalize the coefficient c_1 to one. In addition, fix some network n , and define the quantity $\zeta_n^k = h_k + \alpha_n + \sum_{\ell=1}^K J_{k\ell} m_n^{\ell*}$. For any agent in group k who resides in network n and has covariates X_i , the expected choice ω_i equals:

$$\mathbb{E}(\omega_i | X_i, k, \alpha_n, \{m_n^{\ell*}\}_{\ell=1}^K) = F_{\varepsilon|k}(\zeta_n^k + X_i'c)$$

For identification of $\{c, F_{\varepsilon|k}\}$, I must show: $F_{\varepsilon|k}(\zeta_n^k + X_i'c) = \hat{F}_{\varepsilon|k}(\hat{\zeta}_n^k + X_i'\hat{c})$ implies that $(c, F_{\varepsilon|k}) = (\hat{c}, \hat{F}_{\varepsilon|k})$ for all $X_i \in \text{supp}(X|k)$. This result holds by Manski's (1988) corollary. Also, since it holds for all $k \in \mathcal{K}$, I conclude that c and $\{F_{\varepsilon|k}\}_{k=1}^K$ are identified up to scale.

Having shown that c and $\{F_{\varepsilon|k}\}_{k=1}^K$ can be recovered, the rest of the proof follows by the same arguments used to prove Theorem 1. So, the rest of the parameters are also identified.

Proof of Theorem 3

Q.E.D.

Proof. For part (i), I begin by re-writing equation (17) in terms of the observed quantities:

$$F_{\varepsilon|k_1}^{-1}(\bar{\omega}_{n,a}^{k_1}) - F_{\varepsilon|k_2}^{-1}(\bar{\omega}_{n,a}^{k_2}) = h_{k_1} - h_{k_2} + \mathbb{E}(v_{nk_1,a}) - \mathbb{E}(v_{nk_2,a}) + \sum_{\ell=1}^K (J_{k_1\ell} - J_{k_2\ell}) \bar{\omega}_{n,a}^{\ell} + \tilde{\xi}_n,$$

where $\tilde{\xi}_n = -\sum_{\ell=1}^K (J_{k_1\ell} - J_{k_2\ell}) u_{n\ell,a} + v_{nk_1,a} - v_{nk_2,a} - \mathbb{E}(v_{nk_1,a}) + \mathbb{E}(v_{nk_2,a})$ and $\mathbb{E}(\tilde{\xi}_n) = 0$. Note: $(\bar{\omega}_{n,b}^1, \dots, \bar{\omega}_{n,b}^K) \perp (u_{n1,a}, \dots, u_{nK,a})$ and $\text{Cov}(\bar{\omega}_{n,a}^{\ell}, \bar{\omega}_{n,b}^{\ell}) = \text{Var}(m_n^{\ell*})$ for all $\ell \in \mathcal{K}$. Therefore, both exogeneity and instrument relevance are satisfied. By standard IV arguments, $\beta_{k_1, k_2}^{IV} = [h_{k_1} - h_{k_2} + \mathbb{E}(v_{nk_1,a}) - \mathbb{E}(v_{nk_2,a}), J_{k_11} - J_{k_21}, \dots, J_{k_1K} - J_{k_2K}]'$.

For part (ii), I again re-write the equilibrium equation in terms of the observed quantities:

$$F_{\varepsilon|k}^{-1}(\bar{\omega}_{n,a}^k) = h_k + \mathbb{E}(v_{nk,a}) + W_n' d + \sum_{\ell=1}^K J_{k\ell} \bar{\omega}_{n,a}^{\ell} + \tilde{\xi}_n,$$

1 where $\tilde{\xi}_n = -\sum_{\ell=1}^K J_{k\ell} u_{n\ell,a} + v_{nk,a} - E(v_{nk,a})$. Just as before, the instruments satisfy 1
 2 exogeneity and relevance, which guarantees that $\beta_k^{IV} = [h_k + E(v_{nk,a}), d', J_{k1}, \dots, J_{kK}]'$. 2

3 *Q.E.D.* 3

4 REFERENCES 4

- 5 AKERLOF, GEORGE A. & RACHEL E. KRANTON (2000): "Economics and Identity," *The Quarterly Journal of* 5
 6 *Economics*, 115 (3), 715–753. [2] 6
- 7 ALLENDE, CLAUDIA (2019): "Competition Under Social Interactions and the Design of Education Policies," 7
 8 Working Paper. [21] 8
- 9 ATHEY, SUSAN (2001): "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of 9
 10 Incomplete Information," *Econometrica*, 69 (4), 861–889. [3] 10
- 11 ——— (2002): "Monotone Comparative Statics under Uncertainty," *The Quarterly Journal of Economics*, 117 11
 (1), 187–223. [3]
- 12 BALLESTER, CORALIO, ANTONI CALVÓ-ARMENGOL, & YVES ZENOU (2006): "Who's Who in Networks. 12
 13 Wanted: The Key Player," *Econometrica*, 74 (5), 1403–1417. [3] 13
- 14 BERNHEIM, B. DOUGLAS (1994): "A Theory of Conformity," *Journal of Political Economy*, 102, 841–77. [7] 14
- 15 BERRY, STEVEN T. & PHILIP A. HAILE (2022): "Nonparametric Identification of Differentiated Products De- 15
 16 mand Using Micro Data," Working paper. [5] 16
- 17 BERTRAND, MARIANNE & EMIR KAMENICA (2023): "Coming Apart? Cultural Distances in the United States 17
 over Time," *American Economic Journal: Applied Economics*, 15 (4), 100–141. [27] 17
- 18 BHATTACHARYA, DEBOPAM, PASCALINE DUPAS, & SHIN KANAYA (2023): "Demand and Welfare Analysis in 18
 19 Discrete Choice Models with Social Interactions," *The Review of Economic Studies*, rdad053. [4, 5, 15, 25] 19
- 20 BLUME, LAWRENCE, WILLIAM BROCK, STEVEN DURLAUF, & YANNIS IOANNIDES (2011): "Identification of 20
 Social Interactions," North-Holland, vol. 1 of *Handbook of Social Economics*, 853–964. [4]
- 21 BLUME, LAWRENCE, WILLIAM BROCK, STEVEN DURLAUF, & RAJSHRI JAYARAMAN (2015): "Linear Social 21
 22 Interactions Models," *Journal of Political Economy*, 123 (2), 444–496. [7, 8, 23] 22
- 23 BOOZER, MICHAEL A. & STEPHEN E. CACCIOLA (2001): "Inside the 'Black Box' of Project STAR: Estimation 23
 24 of Peer Effects Using Experimental Data," Center Discussion Paper 832. [6, 33] 24
- 25 BOSTWICK, VALERIE K. & BRUCE A. WEINBERG (2022): "Nevertheless She Persisted? Gender Peer Effects in 25
 Doctoral STEM Programs," *Journal of Labor Economics*, 40 (2), 397–436. [6]
- 26 BOXELL, LEVI, MATTHEW GENTZKOW, & JESSE M. SHAPIRO (2022): "Cross-Country Trends in Affective 26
 27 Polarization," *The Review of Economics and Statistics*, 1–60. [27] 27
- 28 BRAMOULLÉ, YANN & RACHEL KRANTON (2007): "Risk Sharing Across Communities," *American Economic* 28
 29 *Review*, 97 (2), 70–74. [3] 29
- 30 BRAMOULLÉ, YANN, RACHEL KRANTON, & MARTIN D'AMOURS (2014): "Strategic Interaction and Net- 30
 works," *American Economic Review*, 104 (3), 898–930. [3, 17]
- 31 BROCK, WILLIAM & STEVEN DURLAUF (2001): "Discrete Choice with Social Interactions," *The Review of* 31
 32 *Economic Studies*, 68 (2), 235–260. [2, 3, 4, 6, 7, 23, 24, 27, 29] 32

- 1 ——— (2007): “Identification of Binary Choice Models with Social Interactions,” *Journal of Econometrics*, 140 1
2 (1), 52–75. [4, 24, 26, 27, 41] 2
- 3 CABRALES, ANTONIO, ANTONI CALVO-ARMENGOL, & YVES ZENOU (2011): “Social Interactions and 3
4 Spillovers,” *Games and Economic Behavior*, 72 (2), 339–360. [3] 4
- 5 CALVO-ARMENGOL, ANTONI, ELEONORA PATAACCHINI, & YVES ZENOU (2009): “Peer Effects and Social 5
6 Networks in Education,” *The Review of Economic Studies*, 76 (4), 1239–1267. [3] 6
- 7 CARTWRIGHT, DORWIN & FRANK HARARY (1956): “Structural balance: A Generalization of Heider’s Theory,” 7
8 *Psychological Review*, 63 (5), 277–293. [13] 7
- 9 CHARNESS, GARY & YAN CHEN (2020): “Social Identity, Group Behavior, and Teams,” *Annual Review of* 8
10 *Economics*, 12 (1), 691–713. [6] 9
- 11 COOPER, RUSSELL & ANDREW JOHN (1988): “Coordinating Coordination Failures in Keynesian Models,” *The* 10
12 *Quarterly Journal of Economics*, 103 (3), 441–463. [3] 11
- 13 ELLIOTT, MATTHEW & BENJAMIN GOLUB (2019): “A Network Approach to Public Goods,” *Journal of Political* 11
14 *Economy*, 127 (2), 730–776. [3] 12
- 15 ELLISON, GLENN (1993): “Learning, Local Interaction, and Coordination,” *Econometrica*, 61, 1047–1071. [19] 13
- 16 GALEOTTI, ANDREA, SANJEEV GOYAL, MATTHEW O. JACKSON, FERNANDO VEGA-REDONDO, & LEFAT 14
17 YARIV (2010): “Network Games,” *The Review of Economic Studies*, 77 (1), 218–244. [3] 15
- 18 GLAESER, EDWARD L., BRUCE SACERDOTE, & JOSÉ A. SCHEINKMAN (1996): “Crime and Social Interac- 16
19 tions,” *The Quarterly Journal of Economics*, 111 (2), 507–548. [3] 16
- 20 ——— (2003): “The Social Multiplier,” *Journal of the European Economic Association*, 345–353. [3] 17
- 21 GRAHAM, BRYAN S. (2008): “Identifying Social Interactions through Conditional Variance Restrictions,” *Econo-* 18
22 *metrica*, 76 (3), 643–660. [6, 33] 19
- 23 HARSANYI, J. C. (1973): “Oddness of the number of equilibrium points: A new proof,” *International Journal of* 20
24 *Game Theory*, 2 (1), 235–250. [11] 21
- 25 HENRY, DANIEL (1981): *Geometric Theory of Semilinear Parabolic Equations.*, vol. 840 of *Lecture Notes in* 22
26 *Mathematics*, Berlin: Springer-Verlag. [37] 22
- 27 HOROWITZ, JOEL (2009): *Semiparametric and Nonparametric Methods in Econometrics.*, vol. 692. [22] 23
- 28 HOXBY, CAROLINE M. (2000): “The Effects of Class Size on Student Achievement: New Evidence from Popu- 24
29 lation Variation,” *The Quarterly Journal of Economics*, 115 (4), 1239–1285. [2, 4, 6] 25
- 30 JACKSON, MATTHEW O. & YVES ZENOU (2015): “Games on Networks,” Elsevier, vol. 4 of *Handbook of Game* 26
31 *Theory with Economic Applications*, 95–163. [3] 27
- 32 KLINE, BRENDAN & ELIE TAMER (2020): “Econometric Analysis of Models with Social Interactions,” in *Econo-* 28
33 *metric Analysis of Network Data*, ed. by Bryan Graham & Aureo de Paula, Academic Press, 149–181. [4] 28
- 34 KOHLBERG, ELON & JEAN-FRANCOIS MERTENS (1986): “On the Strategic Stability of Equilibria,” *Economet-* 29
35 *rica*, 54 (5), 1003–1037. [11] 30
- 36 LAVY, VICTOR & ANALIA SCHLOSSER (2011): “Mechanisms and Impacts of Gender Peer Effects at School,” 31
37 *American Economic Journal: Applied Economics*, 3 (2), 1–33. [6] 32

- 1 MANSKI, CHARLES F. (1988): “Identification of Binary Response Models,” *Journal of the American Statistical* 1
2 *Association*, 83 (403), 729–738. [22, 26, 27, 41, 42] 2
- 3 ——— (1993): “Identification of Endogenous Social Effects: The Reflection Problem,” *The Review of Economic* 3
4 *Studies*, 60 (3), 531–542. [24] 3
- 5 MILGROM, PAUL & JOHN ROBERTS (1990): “The Economics of Modern Manufacturing: Technology, Strategy, 4
6 and Organization,” *The American Economic Review*, 80 (3), 511–528. [3, 11] 5
- 7 ——— (1994): “Comparing Equilibria,” *The American Economic Review*, 84 (3), 441–459. [11] 6
- 8 MILGROM, PAUL & CHRIS SHANNON (1994): “Monotone Comparative Statics,” *Econometrica*, 62 (1), 157–180. 7
9 [3, 11] 8
- 10 MILNOR, JOHN (1965): *Topology from the Differentiable Viewpoint*, The University Press of Virginia. [37, 38] 9
- 11 PAULA, AUREO DE (2017): *Econometrics of Network Models*, Cambridge University Press, vol. 1 of *Econometric* 10
12 *Society Monographs*, 268–323. [4] 10
- 13 SHAYO, MOSES (2020): “Social Identity and Economic Policy,” *Annual Review of Economics*, 12, 355–389. [6] 11
- 14 TAMER, ELIE (2003): “Incomplete Simultaneous Discrete Response Model with Multiple Equilibria,” *The Review* 12
15 *of Economic Studies*, 70 (1), 147–165. [25] 12
- 16 TOPKIS, D.M. (1998): *Supermodularity and Complementarity*, Frontiers of Economic Research, Princeton Uni- 13
17 versity Press. [3] 14
- 18 VIVES, XAVIER (1990): “Nash equilibrium with strategic complementarities,” *Journal of Mathematical Eco-* 15
19 *nomics*, 19 (3), 305–321. [3] 16
- 20 WILSON, ROBERT (1971): “Computing Equilibria of N-Person Games,” *SIAM Journal on Applied Mathematics*, 17
21 21 (1), 80–87. [11] 17
- 22 WOOLDRIDGE, J.M. (2013): *Introductory Econometrics: A Modern Approach*, Cengage Learning. [5] 18
- 23 WORD, ELIZABETH R., JOHN JOHNSTON, HELEN PATE BAIN, B. DEWAYNE FULTON, JAYNE B. ZAHARIAS, 19
24 CHARLES M. ACHILLES, MARTHA NANNETTE LINTZ, JOHN FOLGER, & CAROLYN BRED A (1990): “The 20
25 State of Tennessee’s Student/Teacher Achievement Ratio (STAR) Project: Technical Report,” . [33] 21
26 22
27 23
28 24
29 25
30 26
31 27
32 28
33 29
34 30
35 31
36 32