A Social Interaction Model with Ordered Choices*

Xiaodong Liu†and Jiannan Zhou

Department of Economics, University of Colorado, Boulder, CO 80309, U.S.A.

September 20, 2017

Abstract

We introduce a social interaction model with ordered choices. We provide a micro-foundation for the econometric model based on an incomplete information network game, and characterize the sufficient condition for the existence of a unique equilibrium of the game. We discuss the identification of the model, and propose to estimate the model by the NFXP and NPL algorithms. We conduct Monte Carlo simulations to investigate the finite sample performance of these two estimation methods.

JEL classification: C31, C35

Key words: ordered probit and logit models, rational expectations, social networks

*We would like to thank the editor Badi H. Baltagi and an anonymous referee for helpful comments. All remaining errors are our own.

†Corresponding Author. Department of Economics, University of Colorado Boulder, UCB 256, Boulder, CO 80309, USA. Email: xiaodong.liu@colorado.edu.

1 Introduction

Estimation and inference methods have been proposed for social interaction models, with binary choices (e.g., Brock and Durlauf, 2001; Lee et al., 2014; Lin and Xu, 2016), unordered multinomial choices (e.g., Brock and Durlauf, 2002; Brock and Durlauf, 2006; Xu, 2016), and continuous choice variables (e.g., Bramoullé et al., 2009; Lee et al., 2010; Liu and Lee, 2010). However, little work has been done on the analysis of network models with ordered multinomial choices. Network data are usually from surveys, and thus the response variable of an empirical study is often ordinal in nature. For example, in the widely used National Longitudinal Study of Adolescent Health (Add Health) data,[1] academic performance in a certain subject ("A", "B", "C", "D or lower"), study effort ("I try very hard to do my best", "I try hard enough, but not as hard as I could", "I don't try very hard", "I never try at all"), and smoking and drinking frequency ("never", "once or twice", "once a month or less", "2 or 3 days a month", "once or twice a week", "3 to 5 days a week", "nearly everyday") are all coded as ordinal variables.

In this paper, we introduce a social interaction model with ordered choices. We provide a micro-foundation for the econometric model based on an incomplete information network game. We characterize the sufficient condition for the existence of a unique rational expectation equilibrium of the game, which in turn guarantees the coherency and completeness of the econometric model. We discuss the identification and estimation of the econometric model. As the econometric model involves the equilibrium rational expectation that needs to be solved from a fixed point mapping, it can be estimated by the nested fixed point (NFXP) algorithm (Rust, 1987) or the nested pseudo likelihood (NPL) algorithm (Aguirregabiria and Mira, 2007). We investigate the finite sample performance of these two estimation methods in Monte Carlo simulations. We find the NFXP estimator has a smaller standard deviation, while the NPL estimator is computationally more efficient.

The rest of the paper is organized as follows. Section 2 presents the econometric model and provides the sufficient condition for the existence of a unique equilibrium of the underlying network

[1] For more information about the Add Health data, see http://www.cpc.unc.edu/projects/addhealth.

game. Section 3 discusses the identification of the model. Section 4 explains how to estimate the model parameters by the NFXP and NPL algorithms and how to interpret parameter estimates in terms of marginal effects. Section 5 provides simulation results on the finite sample performance of the proposed NFXP and NPL estimators. Section 6 concludes. The proofs are collected in the online appendix.

2 Model

A set of individuals $\mathcal{N} = \{1, \cdots, n\}$ interacts within a network. Let $\mathbf{W} = [w_{ij}]$ be an $n \times n$ predetermined adjacency matrix, where the (i, j)th element w_{ij} ($w_{ij} \geq 0$) captures the proximity of individuals i and j in the network. As a normalization, $w_{ii} = 0$ for all i. We assume the network topology captured by $\mathbf{W}$ is common knowledge among all individuals in the network.

Suppose all individuals in the network face m ordered alternatives. Individual i chooses alternative k, i.e. $y_i = k$, if and only if

$$\alpha_{k-1} < y_i^* \leq \alpha_k$$

where

$$y_i^* = \lambda \textstyle\sum_{j=1}^n w_{ij} y_j^{\mathrm{E(i)}} + \mathbf{x}_i' \boldsymbol{\beta} + \epsilon_i \tag{1}$$

and α_k's are threshold parameters such that $-\infty = \alpha_0 < \alpha_1 < \cdots < \alpha_{m-1} < \alpha_m = \infty$. In equation (1), $y_j^{\mathrm{E(i)}}$ denotes individual i's *subjective* expected value of y_j. $\sum_{j=1}^n w_{ij} y_j^{\mathrm{E(i)}}$ is the weighted sum of individual i's subjective expectations on her peers' choices, and the coefficient λ represents the *peer effect*. $\mathbf{x}_i$ is a column vector of exogenous variables that captures the characteristics of individual i and her peers. For instance, let $\overline{\mathbf{X}}$ be an $n \times q$ matrix of observations on exogenous individual characteristics. Then, a possible specification of $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_n]'$ is given by $\mathbf{X} = [\overline{\mathbf{X}}, \mathbf{W}\overline{\mathbf{X}}]$, with the coefficients of $\mathbf{W}\overline{\mathbf{X}}$ representing *exogenous contextual effects* (Manski, 1993). ϵ_i is a random shock that is independent of $\mathbf{W}$ and $\mathbf{X}$ and is independent and identically distributed (i.i.d.) with a distribution function $F(\cdot)$. Only individual i observes ϵ_i while everyone in the network observes

$\mathbf{X}$. Then, given $\mathbf{W}$ and $\mathbf{X}$, the probability that individual i chooses alternative k is[2]

$$\Pr(y_i = k|\mathbf{W}, \mathbf{X}) = F(\alpha_k - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E(i)}} - \mathbf{x}_i'\boldsymbol{\beta}) - F(\alpha_{k-1} - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E(i)}} - \mathbf{x}_i'\boldsymbol{\beta}),$$

for $k = 1, \cdots, m$, where $\mathbf{w}_i$ denotes the ith row of $\mathbf{W}$ and $\mathbf{y}^{\mathrm{E(i)}} = [y_1^{\mathrm{E(i)}}, \cdots, y_n^{\mathrm{E(i)}}]'$. Let y_i^{E} denote the mathematical expectation of y_i. Then,

$$y_i^{\mathrm{E}} \equiv \mathrm{E}(y_i|\mathbf{W}, \mathbf{X}) = \sum_{k=1}^{m} k \Pr(y_i = k|\mathbf{W}, \mathbf{X}) = m - \sum_{k=1}^{m-1} F(\alpha_k - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E(i)}} - \mathbf{x}_i'\boldsymbol{\beta}).$$

In the rational expectation equilibrium (see, e.g., Brock and Durlauf, 2001; Lee et al., 2014), the subjective expectation coincides with the mathematical expectation such that $\mathbf{y}^{\mathrm{E(i)}} = \mathbf{y}^{\mathrm{E}}$, where $\mathbf{y}^{\mathrm{E}} = [y_1^{\mathrm{E}}, \cdots, y_n^{\mathrm{E}}]'$, for all i. Thus, in the equilibrium, the rational expectation is characterized by the following system of equations,

$$\mathbf{y}^{\mathrm{E}} = \vec{h}(\mathbf{y}^{\mathrm{E}}; \boldsymbol{\delta}), \tag{2}$$

where $\vec{h}(\mathbf{y}^{\mathrm{E}}; \boldsymbol{\delta}) = [h_1(\mathbf{y}^{\mathrm{E}}; \boldsymbol{\delta}), \cdots, h_n(\mathbf{y}^{\mathrm{E}}; \boldsymbol{\delta})]'$ with

$$h_i(\mathbf{y}^{\mathrm{E}}; \boldsymbol{\delta}) = m - \sum_{k=1}^{m-1} F(\alpha_k - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E}} - \mathbf{x}_i'\boldsymbol{\beta}),$$

and $\boldsymbol{\delta} = (\alpha_1, \cdots, \alpha_{m-1}, \lambda, \boldsymbol{\beta}')'$. A sufficient condition for the existence of a unique solution to the system of equations (2) is given as follows. For an $n \times m$ matrix $\mathbf{A} = [a_{ij}]$, let the row sum and column sum matrix norms of $\mathbf{A}$ be denoted by $||\mathbf{A}||_\infty = \max_{i=1,\cdots,n} \sum_{j=1}^{m} |a_{ij}|$ and $||\mathbf{A}||_1 = \max_{j=1,\cdots,m} \sum_{i=1}^{n} |a_{ij}|$ respectively.

Assumption 1. (i) $\epsilon_1, \cdots, \epsilon_n$ *are random shocks that are independent of* $\mathbf{W}$ *and* $\mathbf{X}$ *and are i.i.d. with a continuous distribution function* $F(\cdot)$ *and density function* $f(\cdot)$. (ii) $|\lambda| < [(m-1)\sup_u f(u) \min\{||\mathbf{W}||_\infty, ||\mathbf{W}||_1\}]^{-1}$.

If $\mathbf{W}$ is row-normalized with $\sum_{j=1}^{m} w_{ij} = 1$ for all i, then $||\mathbf{W}||_\infty = 1$. If ϵ_i follows the standard normal distribution, then $\sup_u f(u) = 1/\sqrt{2\pi}$. If ϵ_i follows the logistic distribution, then

[2] Note that $F(\alpha_0 - \lambda \sum_{j=1}^{n} w_{ij} y_j^{\mathrm{E(i)}} - \mathbf{x}_i'\boldsymbol{\beta}) = 0$ and $F(\alpha_m - \lambda \sum_{j=1}^{n} w_{ij} y_j^{\mathrm{E(i)}} - \mathbf{x}_i'\boldsymbol{\beta}) = 1$.

$\sup_u f(u) = 1/4$. Hence, with a row-normalized adjacency matrix, Assumption 1 holds for an ordered *probit* social interaction model if $|\lambda| < \sqrt{2\pi}/(m-1)$, and holds for an ordered *logit* social interaction model if $|\lambda| < 4/(m-1)$. It is worth noting that, when $m = 2$, Assumption 1 coincides with the sufficient condition for the existence of a unique rational expectation equilibrium of the social interaction model with binary outcomes in Lee et al. (2014).

Proposition 1. *Under Assumption 1, the social interaction model with ordered outcomes has a unique rational expectation equilibrium.*

When Assumption 1 holds, the contraction mapping property of $\vec{h}(\cdot)$ not only guarantees the coherency and completeness of the model (Tamer, 2003), but also suggests the system of equations (2) can be solved by recursive iterations.

3 Identification

For identification, we follow Lee et al. (2014), Xu (2016), Lin and Xu (2016) and others by assuming $F(\cdot)$ is a strictly increasing distribution function with unity variance that is known to the econometrician.[3] Furthermore, as in a standard ordered choice model (McKelvey and Zavoina, 1975), we impose an identification constraint that $\mathbf{X}$ does not have a constant column (i.e., the intercept is set to be zero).[4]

Given the network topology $\mathbf{W}$, two sets of parameters, $\boldsymbol{\delta} = (\alpha_1, \cdots, \alpha_{m-1}, \lambda, \boldsymbol{\beta}')'$ and $\widetilde{\boldsymbol{\delta}} = (\widetilde{\alpha}_1, \cdots, \widetilde{\alpha}_{m-1}, \widetilde{\lambda}, \widetilde{\boldsymbol{\beta}}')'$, are observationally equivalent if

$$\Pr(y_i \leq k|\mathbf{W}, \mathbf{X}) = F(\alpha_k - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E}} - \mathbf{x}_i'\boldsymbol{\beta}) = F(\widetilde{\alpha}_k - \widetilde{\lambda}\mathbf{w}_i \widetilde{\mathbf{y}}^{\mathrm{E}} - \mathbf{x}_i'\widetilde{\boldsymbol{\beta}}),$$

for $k = 1, \cdots, m-1$ and for any $\mathbf{X}$ in its support, where, under Assumption 1, $\mathbf{y}^{\mathrm{E}}$ and $\widetilde{\mathbf{y}}^{\mathrm{E}}$ are

[3] In constrast, Brock and Durlauf (2007) consider the identification of binary choice group interaction models when the distribution of ϵ_i is unknown. In this paper, we discuss identification assuming $F(\cdot)$ is known as the proposed estimation procedure is parametric in nature.

[4] Instead of dropping the intercept, one could impose an identification constraint that one of the threshold parameters α_k is a known constant (e.g., $\alpha_1 = 0$).

uniquely determined by

$$\begin{aligned} y_i^{\mathrm{E}} &= m - \textstyle\sum_{k=1}^{m-1} F(\alpha_k - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E}} - \mathbf{x}_i' \boldsymbol{\beta}) \\ \widetilde{y}_i^{\mathrm{E}} &= m - \textstyle\sum_{k=1}^{m-1} F(\widetilde{\alpha}_k - \widetilde{\lambda} \mathbf{w}_i \widetilde{\mathbf{y}}^{\mathrm{E}} - \mathbf{x}_i' \widetilde{\boldsymbol{\beta}}) \end{aligned}$$

for $i = 1, \cdots, n$, respectively. Identification holds if observational equivalence of $\boldsymbol{\delta}$ and $\widetilde{\boldsymbol{\delta}}$ implies $\boldsymbol{\delta} = \widetilde{\boldsymbol{\delta}}$. Let $\mathbf{Z} = [\boldsymbol{\iota}_n, \mathbf{W}\mathbf{y}^{\mathrm{E}}, \mathbf{X}]$, where $\boldsymbol{\iota}_n$ is an $n \times 1$ vector of ones.[5]

Assumption 2. (i) $F(\cdot)$ *is a strictly increasing distribution function with unity variance that is known to the econometrician.* (ii) $\mathrm{plim}_{n\to\infty} n^{-1}\mathbf{Z}'\mathbf{Z}$ *exists and is a finite positive definite matrix of full rank.*

Proposition 2. *Under Assumptions 1 and 2, the social interaction model with ordered outcomes is identified.*

4 Estimation

One way to estimate this model is to use the NFXP algorithm. The NFXP algorithm was proposed by Rust (1987) for the estimation of dynamic discrete choice models, and has recently been applied in Lee et al. (2014) and Yang and Lee (2017) among others for the estimation of discrete choice social interaction models. For the estimation of the social interaction model with ordered choices, the NFXP algorithm use an inner loop that solves the system of equations (2) for $\mathbf{y}^{\mathrm{E}}$ by recursive iterations at each candidate parameter value of $\boldsymbol{\delta} = (\alpha_1, \cdots, \alpha_{m-1}, \lambda, \boldsymbol{\beta}')'$ in the search for the maximum of the log-likelihood function

$$\ln L(\boldsymbol{\delta}; \mathbf{y}^{\mathrm{E}}) = \sum_{i=1}^{n} \sum_{k=1}^{m} d_{ik} \ln \left[F(\alpha_k - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E}} - \mathbf{x}_i' \boldsymbol{\beta}) - F(\alpha_{k-1} - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E}} - \mathbf{x}_i' \boldsymbol{\beta}) \right] \tag{3}$$

where $d_{ik} = 1$ if $y_i = k$ and $d_{ik} = 0$ otherwise.

[5] Although $\mathbf{y}^{\mathrm{E}}$ is unobservable, under Assumptions 1 and 2, it is a well defined, known function of the observables, $\mathbf{W}$ and $\mathbf{X}$, and model parameters $\boldsymbol{\delta}$.

A computationally more efficient estimation method for this model is the NPL algorithm proposed by Aguirregabiria and Mira (2007). This method has recently been adopted by Lin and Xu (2016) and Liu (2017) for the estimation of large network games. For the estimation of the social interaction model with ordered choices, the NPL algorithm starts from an arbitrary initial value $\mathbf{y}^{\mathrm{E}}_{(0)}$ in its support and takes the following iterative steps:

Step 1 Given $\mathbf{y}^{\mathrm{E}}_{(t-1)}$, obtain $\widehat{\boldsymbol{\delta}}_{(t)} = \arg\max_{\boldsymbol{\delta}} \ln L(\boldsymbol{\delta}; \mathbf{y}^{\mathrm{E}}_{(t-1)})$, with the log-likelihood function defined in (3).

Step 2 Given $\mathbf{y}^{\mathrm{E}}_{(t-1)}$ and $\widehat{\boldsymbol{\delta}}_{(t)}$, obtain $\mathbf{y}^{\mathrm{E}}_{(t)} = \vec{h}(\mathbf{y}^{\mathrm{E}}_{(t-1)}; \widehat{\boldsymbol{\delta}}_{(t)})$, with the best response function defined in (2).

Step 3 Update $\mathbf{y}^{\mathrm{E}}_{(t-1)}$ to $\mathbf{y}^{\mathrm{E}}_{(t)}$ in Step 1 and repeat Steps 1 and 2 until the process converges.

Kasahara and Shimotsu (2012) show that a key determinant of the convergence of the NPL algorithm is the contraction property of (2), which is ensured by Assumption 1. When the NPL algorithm converges, the NPL estimator $\widehat{\boldsymbol{\delta}}$ satisfies $\widehat{\boldsymbol{\delta}} = \arg\max \ln L(\boldsymbol{\delta}; \widehat{\mathbf{y}}^{\mathrm{E}})$, where $\widehat{\mathbf{y}}^{\mathrm{E}}$ is the unique solution of the system of equations $\mathbf{y}^{\mathrm{E}} = \vec{h}(\mathbf{y}^{\mathrm{E}}; \widehat{\boldsymbol{\delta}})$. Under some standard regularity conditions, it follows by a similar argument as in Aguirregabiria and Mira (2007) and Lin and Xu (2016) that the NPL estimator is root-n consistent and asymptotically normal.

As in other discrete choice models, parameter estimates can be interpreted in terms of marginal effects. With social interaction, marginal effects should incorporate changes in expected choices of all individuals in the network caused by the change of an explanatory variable. There are two types of marginal effects — *direct* and *indirect* marginal effects. The direct marginal effect captures the impact of a unit change in an explanatory variable of an individual on her own choice probability. The indirect marginal effect captures the impact of a unit change in an explanatory variable of an individual on the choice probabilities of the other individuals in the network.

Formally, the average direct marginal effect (ADME) of a continuous explanatory variable x_{ih}

on the probability of individual i choosing alternative k is[6]

$$ADME_{k,h} = \frac{1}{n}\sum_{i=1}^{n}(f_{i,k-1} - f_{i,k})(\lambda \mathbf{w}_i \frac{\partial \mathbf{y}^{\mathrm{E}}}{\partial x_{ih}} + \beta_h)$$

and the average indirect marginal effect (AIME) of a continuous explanatory variable x_{ih} on the probability of individual j ($j \neq i$) choosing alternative k is

$$AIME_{k,h} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}(f_{j,k-1} - f_{j,k})\lambda \mathbf{w}_j \frac{\partial \mathbf{y}^{\mathrm{E}}}{\partial x_{ih}}$$

where $f_{i,k} = f(\alpha_k - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E}} - \mathbf{x}_i'\boldsymbol{\beta})$ and $f(\cdot)$ is the density function of ϵ_i. The marginal effect includes a term $\partial \mathbf{y}^{\mathrm{E}}/\partial x_{ih}$, which represents the impact of a unit change in an explanatory variable on the rational expectation of equilibrium choices. Using the implicit function theorem,

$$\frac{\partial \mathbf{y}^{\mathrm{E}}}{\partial x_{ih}} = (\mathbf{I}_n - \lambda \mathbf{D}\mathbf{W})^{-1}\mathbf{D}\frac{\partial \mathbf{X}}{\partial x_{ih}}\boldsymbol{\beta}.$$

where $\mathbf{D} = \mathrm{diag}_{i=1}^{n}(\sum_{k=1}^{m-1} f_{i,k})$.

5 Monte Carlo Experiment

To investigate the finite sample performance of the proposed estimation procedures, we conduct a limited simulation study with the latent variable given by

$$y_i^* = \lambda \sum_{j=1}^{n} w_{ij} y_j^{\mathrm{E(i)}} + \beta_1 \overline{x}_i + \beta_2 \sum_{j=1}^{n} w_{ij}\overline{x}_j + \epsilon_i,$$

where $\overline{x}_i$ and ϵ_i are, respectively, i.i.d. standard normal and logistic random variables. In the experiment, we consider a circular network where the n individuals are equidistantly located around a circle and are only connected with the nearest neighbors. The non-zero elements of the corresponding adjacency matrix $\mathbf{W} = [w_{ij}]$ are $w_{1,2} = w_{1,n} = w_{n,1} = w_{n,n-1} = 1/2$ and $w_{i,i-1} = w_{i,i+1} = 1/2$

[6] For notational simplicity, in the derivation of marginal effects, we assume the model does not include a contextual effect regressor such that $\partial x_{jg}/\partial x_{ih} = 0$ if $i \neq j$ or $g \neq h$.

for $i = 2, \cdots, n-1$. We set $\lambda = 0.5$, $\beta_1 = \beta_2 = 1$, and $m = 3$ with the threshold parameters $\alpha_1 = 0$ and $\alpha_2 = 1$. We conduct 1000 repetitions in the simulation with $n \in \{500, 1000, 2000\}$. We report the mean and standard deviation (SD) of the empirical distributions of the NFXP and NPL estimates in Table 1. We also report the average computation time (in seconds) per repetition of the Monte Carlo simulation.[7]

Table 1: Monte Carlo Simulation Results

	$\lambda = 0.5$	$\beta_1 = 1.0$	$\beta_2 = 1.0$	$\alpha_1 = 0.0$	$\alpha_2 = 1.0$	computation time
NFXP						(seconds)
$n = 500$	0.438(0.483)	1.023(0.150)	1.031(0.263)	−0.151(1.070)	0.862(1.077)	0.347
$n = 1000$	0.453(0.336)	1.018(0.109)	1.025(0.190)	−0.114(0.752)	0.894(0.751)	0.416
$n = 2000$	0.454(0.241)	1.014(0.078)	1.022(0.133)	−0.105(0.536)	0.900(0.536)	0.538
NPL						
$n = 500$	0.471(0.533)	1.016(0.156)	1.012(0.287)	−0.078(1.181)	0.935(1.188)	0.233
$n = 1000$	0.490(0.380)	1.010(0.116)	1.005(0.208)	−0.031(0.849)	0.977(0.849)	0.257
$n = 2000$	0.496(0.274)	1.004(0.083)	1.002(0.147)	−0.012(0.609)	0.993(0.609)	0.311

Mean(SD)

First, we notice from the simulation results that the estimates of the peer effect coefficient λ and threshold parameters α_1 and α_2 are downwards biased. The bias is more severe for the NFXP estimator. The bias reduces as the sample size increases. When $n = 2000$, the NPL estimator is essentially unbiased. Second, as the NPL estimator maximizes the "pseudo" log-likelihood function evaluated at $\mathbf{y}^{\mathrm{E}}_{(t-1)}$ that is not necessarily the equilibrium $\mathbf{y}^{\mathrm{E}}$ associated with $\boldsymbol{\delta}$ in its every iteration, the NPL estimator may not be asymptotically as efficient as the NFXP estimator. This is reflected by the slightly larger standard deviations of the NPL estimator. Finally, as the NPL algorithm does not require to solve the fixed point mapping (2) at each candidate parameter value of $\boldsymbol{\delta}$ in the search for the maximum of the log-likelihood function, it is faster than the NFXP algorithm. From the last column of Table 1, we can see that the computational advantage of the NPL estimator is more pronounced when the sample size is large.

[7] The computation is conducted on a PC with an Intel(R) Core(TM) i7-6700 CPU @ 3.4 GHz and 32 GB RAM.

6 Summary

In this paper, we introduce a new econometric model for the analysis of social networks with ordered choices. The specification of the econometric model can be motivated by an incomplete information network game. We provide the sufficient condition for the existence of a unique equilibrium of the underlying network game. We discuss the identification of the model and propose to estimate the model by the NFXP and NPL algorithms. We derive the marginal effect formula that facilitates the interpretation of the estimated parameters. We also conduct simulation experiments to study the finite sample performance of the proposed NFXP and NPL estimators. We find the NFXP estimator has a smaller standard deviation, while the NPL estimator is computationally more efficient.

References

Aguirregabiria, V. and Mira, P. (2007). Sequential estimation of dynamic discrete games, *Econometrica* **75**: 1–53.

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

Brock, W. A. and Durlauf, S. N. (2001). Discrete choice with social interaction, *Review of Economic Studies* **68**: 235–260.

Brock, W. A. and Durlauf, S. N. (2002). A multinomial choice model with neighborhood effects, *American Economic Review* **92**: 298–303.

Brock, W. A. and Durlauf, S. N. (2006). Multinomial choice with social interactions, *in* L. E. Blume and S. N. Durlauf (eds), *The Economy as an Evolving Complex System*, Vol. III, Oxford University Press, pp. 175–206.

Brock, W. A. and Durlauf, S. N. (2007). Identification of binary choice models with social interactions, *Journal of Econometrics* **140**: 52–75.

Kasahara, H. and Shimotsu, K. (2012). Sequential estimation of structural models with a fixed point constraint, *Econometrica* **80**: 2303–2319.

Lee, L. F., Li, J. and Lin, X. (2014). Binary choice models with social network under heterogeneous rational expectations, *The Review of Economics and Statistics* **96**: 402–417.

Lee, L. F., Liu, X. and Lin, X. (2010). Specification and estimation of social interaction models with network structures, *The Econometrics Journal* **13**: 145–176.

Lin, Z. and Xu, H. (2016). Estimation of social-influence-dependent peer pressures in a large network game. Working paper, Emory University.

Liu, X. (2017). Simultaneous equations with binary outcomes and social interactions. Working paper, University of Colorado Boulder.

Liu, X. and Lee, L. F. (2010). GMM estimation of social interaction models with centrality, *Journal of Econometrics* **159**: 99–115.

Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem, *The Review of Economic Studies* **60**: 531–542.

McKelvey, R. and Zavoina, W. (1975). A statistical model for the analysis of ordinal level variables, *Journal of Mathematical Sociology* **4**: 103–120.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher, *Econometrica* **55**: 999–1033.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria, *The Review of Economic Studies* **70**: 147–165.

Xu, H. (2016). Social interactions on large networks: A game theoretic approach. Forthcoming in *International Economic Review*.

Yang, C. and Lee, L. F. (2017). Social interactions under incomplete information with heterogeneous expectations, *Journal of Econometrics* **198**: 65–83.

Online Supplement to

"A Social Interaction Model with Ordered Choices"

Xiaodong Liu and Jiannan Zhou

Department of Economics, University of Colorado, Boulder, CO 80309, U.S.A.

September 20, 2017

A Proofs

Proof of Proposition 1. Under Assumption 1 (i), $\vec{h}(\cdot)$ is continuous, and thus the existence of a solution to the system of equations

$$\mathbf{y}^{\mathrm{E}} = \vec{h}(\mathbf{y}^{\mathrm{E}}), \tag{A.1}$$

is guaranteed by the Brouwer fixed-point theorem. By the contraction mapping theorem, the system of equations (A.1) has a unique solution if $\vec{h}(\cdot)$ is a contraction mapping with respect to some matrix norm $||\cdot||$. Thus, to show the desired result, we only need to show that, under Assumption 1 (ii), $\vec{h}(\cdot)$ is a contraction mapping.

First, we show that $\vec{h}(\cdot)$ is a contraction mapping with respect to a (submultiplicative) matrix norm $||\cdot||$, if $||\vec{h}'(\cdot)|| < 1$ where $\vec{h}'(\mathbf{x}) = \partial\vec{h}(\mathbf{x})/\partial\mathbf{x}'$. Let $\vec{k}(t) = \vec{h}(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1))$. If $||\vec{h}'(\cdot)|| < 1$, then

$$\begin{aligned} ||\vec{h}(\mathbf{x}_2) - \vec{h}(\mathbf{x}_1)|| &= ||\vec{k}(1) - \vec{k}(0)|| = \left\| \int_0^1 \vec{k}'(t)dt \right\| = \left\| \int_0^1 (\mathbf{x}_2 - \mathbf{x}_1)\vec{h}'(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1))dt \right\| \\ &\leq \int_0^1 ||\mathbf{x}_2 - \mathbf{x}_1|| \cdot ||\vec{h}'(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1))||dt < ||\mathbf{x}_2 - \mathbf{x}_1||, \end{aligned}$$

i.e. $\vec{h}(\cdot)$ is a contraction mapping.

Next, as the (i,j)-th element of $\partial\vec{h}(\mathbf{y}^{\mathrm{E}})/\partial\mathbf{y}^{\mathrm{E}\prime}$ is given by

$$\partial h_i(\mathbf{y}^{\mathrm{E}})/\partial y_j^{\mathrm{E}} = \lambda w_{ij} \textstyle\sum_{k=1}^{m-1} f(\alpha_{m-k} - \lambda \sum_{j=1}^{n} w_{ij} y_j^{\mathrm{E}} - \mathbf{x}_i'\boldsymbol{\beta}),$$

we have

$$|\partial h_i(\mathbf{y}^{\mathrm{E}})/\partial y_j^{\mathrm{E}}| \leq (m-1)|\lambda| \cdot |w_{ij}| \cdot \sup_u f(u).$$

Hence, under Assumption 1 (ii), either

$$||\partial\vec{h}(\mathbf{y}^{\mathrm{E}})/\partial\mathbf{y}^{\mathrm{E}\prime}||_\infty \leq (m-1)|\lambda| \cdot ||\mathbf{W}||_\infty \sup_u f(u) < 1,$$

or

$$||\partial\vec{h}(\mathbf{y}^{\mathrm{E}})/\partial\mathbf{y}^{\mathrm{E}\prime}||_1 \leq (m-1)|\lambda| \cdot ||\mathbf{W}||_1 \sup_u f(u) < 1.$$

Therefore, under Assumption 1 (ii), the system of equations (A.1) is a contraction mapping with respect to the $||\cdot||_\infty$ or $||\cdot||_1$ norm, whichever is a smaller matrix norm of $\mathbf{W}$. $\square$

Proof of Proposition 2. Given the network topology $\mathbf{W}$, observational equivalence requires

$$\Pr(y_i \leq k|\mathbf{W}, \mathbf{X}) = F(\alpha_k - \lambda\mathbf{w}_i\mathbf{y}^{\mathrm{E}} - \mathbf{x}_i'\boldsymbol{\beta}) = F(\widetilde{\alpha}_k - \widetilde{\lambda}\mathbf{w}_i\widetilde{\mathbf{y}}^{\mathrm{E}} - \mathbf{x}_i'\widetilde{\boldsymbol{\beta}}),$$

for $k = 1, \cdots, m-1$ and for any $\mathbf{X}$ in its support, where, under Assumption 1, $\mathbf{y}^{\mathrm{E}}$ and $\widetilde{\mathbf{y}}^{\mathrm{E}}$ are uniquely determined by

$$y_i^{\mathrm{E}} = m - \textstyle\sum_{k=1}^{m-1} F(\alpha_k - \lambda\mathbf{w}_i\mathbf{y}^{\mathrm{E}} - \mathbf{x}_i'\boldsymbol{\beta}) \quad \text{(A.2)}$$

$$\widetilde{y}_i^{\mathrm{E}} = m - \textstyle\sum_{k=1}^{m-1} F(\widetilde{\alpha}_k - \widetilde{\lambda}\mathbf{w}_i\widetilde{\mathbf{y}}^{\mathrm{E}} - \mathbf{x}_i'\widetilde{\boldsymbol{\beta}}) \quad \text{(A.3)}$$

for $i = 1, \cdots, n$, respectively. The model is identified if observational equivalence of $\boldsymbol{\delta}$ and $\widetilde{\boldsymbol{\delta}}$ implies $\boldsymbol{\delta} = \widetilde{\boldsymbol{\delta}}$. That is,

$$F(\alpha_k - \lambda\mathbf{w}_i\mathbf{y}^{\mathrm{E}} - \mathbf{x}_i'\boldsymbol{\beta}) = F(\widetilde{\alpha}_k - \widetilde{\lambda}\mathbf{w}_i\widetilde{\mathbf{y}}^{\mathrm{E}} - \mathbf{x}_i'\widetilde{\boldsymbol{\beta}}), \quad \text{(A.4)}$$

for $k = 1, \cdots, m-1$, implies $\boldsymbol{\delta} = \widetilde{\boldsymbol{\delta}}$. If (A.4) holds for $k = 1, \cdots, m-1$, then the right hand sides of (A.2) and (A.3) are identical, i.e., $y_i^{\mathrm{E}} = \widetilde{y}_i^{\mathrm{E}}$, for $i = 1, \cdots, n$. Therefore, (A.4) can be rewritten as

$$F(\alpha_k - \lambda \mathbf{w}_i \mathbf{y}^{\mathrm{E}} - \mathbf{x}_i' \boldsymbol{\beta}) = F(\widetilde{\alpha}_k - \widetilde{\lambda} \mathbf{w}_i \mathbf{y}^{\mathrm{E}} - \mathbf{x}_i' \widetilde{\boldsymbol{\beta}}), \tag{A.5}$$

for $k = 1, \cdots, m-1$. Under Assumption 2, (A.5) implies

$$(\alpha_k - \widetilde{\alpha}_k)\boldsymbol{\iota}_n + (\lambda - \widetilde{\lambda})\mathbf{W}\mathbf{y}^{\mathrm{E}} + \mathbf{X}(\boldsymbol{\beta} - \widetilde{\boldsymbol{\beta}}) = 0,$$

which, in turn, implies $\alpha_k = \widetilde{\alpha}_k$, $\lambda = \widetilde{\lambda}$ and $\boldsymbol{\beta} = \widetilde{\boldsymbol{\beta}}$, for $k = 1, \cdots, m-1$. Hence, observational equivalence of $\boldsymbol{\delta}$ and $\widetilde{\boldsymbol{\delta}}$ implies $\boldsymbol{\delta} = \widetilde{\boldsymbol{\delta}}$, i.e., the model is identified. □