Social Interaction Methods [*]

Chih-Sheng Hsieh[†] Xu Lin[‡] Eleonora Patacchini[§]

June 7, 2020

Abstract

This chapter is concerned with methods for analyzing social interaction effects. The attention is focused on how to estimate endogenous effects, where an individual's choice may depend on those of his/her contacts about the same activity. The analysis is guided by the data structure that is available to measure social interactions, an intuitive aspect that allows empirical researchers to understand whether and how they could study social interaction effects in their own data. First, the case where the information on social interaction patterns is limited to membership to a given group is considered, then the discussion moves to the case where the data contain information on specific relationships among pairs of individuals within each group, and the availability of data on the co-evolution of social structures and outcomes. This chapter also discusses some basic methods to deal with online social network data, and the novel literature estimating social interaction effects relying only on outcome data. For each data structure, the challenges and the main methods proposed in the literature to tackle them are reviewed.

JEL: C1, C23

---

[*] The authors are grateful to Alfonso Flores-Lagunes, Tom Snijders, and one anonymous referee for their comments.

[†] Department of Economics, National Taiwan University, Taipei 10617, Taiwan. E-mail: csh-sieh@ntu.edu.tw.

[‡] Department of Economics, Virginia Tech, 3016 Pamplin Hall, 880 West Campus Drive, Blacksburg, VA 24061. Email: xulin88@vt.edu.

[§] Department of Economics, Cornell University, 430 Uris Hall, Ithaca, NY 14853, EIEF, CEPR and IZA. E-mail: ep454@cornell.edu.

1 Introduction

This chapter is concerned with methods for analyzing social interaction effects. As pointed out by Manski (1993), there could be two types of social effects: endogenous effects, where an individual's choice may depend on those of his/her contacts about the same activity, and contextual effects, where an individual's activity may depend on the exogenous characteristics of his/her contacts. These two types of social effects are not only different in nature but also different in the methods that are used to estimate them. The main substantial difference between the two effects is that endogenous effects produce an externality that can lead to the clustering of outcomes above and beyond the clustering that might be expected based on individuals' observables. Such an externality is known as the "social multiplier". It emphasizes any effect of changes in observable characteristics due, for example, to policy interventions. For instance, if endogenous effects are present, a special tutoring program which benefits a small group of students may affect other students who are not directly targeted by the program. As a result, assessing the existence and magnitude of the social multiplier is paramount for policy purposes. In practice, such a task in not easy. First, separately identifying the two types of social effects is difficult since any exogenous peer attribute that can instrument for endogenous behaviors might also affect outcomes directly. Second, the interdependences between agents' actions make the OLS estimation method inappropriate (Anselin, 1988; Kelejian and Prucha, 1998; Lee, 2004) and thus estimation methods for social interaction effects of this sort need to be more sophisticated. These challenges are probably why most of the literature on social interactions estimates and labels social interaction effects as contextual effects.

This handbook chapter offers a first overview of the methods for the estimation of endogenous social interactions. There exist other surveys focusing on other areas of the literature on social interactions, see for example Advani and Malde (2018), Blume et al. (2011), Fortin and Boucher (2016), Graham (2015), and Topa and Zenou (2015). The organization of the chapter is based on the data structure employed in the social interaction study, an intuitive aspect that allows empirical researchers to understand whether and how they could study social interaction effects in their own data. For each data structure, the challenges and the main methods proposed in the literature to tackle them are reviewed. In Section 2, the attention is given to the case where the information on social interaction patterns is limited to membership to a given group, under the assumption that individuals interact with each and every other member within a group but not outside the group. In Section 3, the overview is expanded to the case where the data contain information on specific relationships among pairs of individuals within each group. Section 4 turns the attention to longitudinal data and Section 5 discusses some basic methods to deal with online social networks. Section 6 reviews the novel literature estimating social interaction effects without information on patterns of interactions at all, that is only relying on outcome data. Finally, Section 7 summarizes with some remarks on questions that are open for future research.

2 Group Data

Group interactions refer to the case where individuals interact with each and every other member within a group but not outside the group. Depending on the research framework and the data availability, group may be defined as region, district, county, village, school, census tract, birth cohort, etc. In general, there is no further information about the interaction patterns within groups.

Two main identification challenges faced by empirical researchers in using group structured data are the "reflection problem" (Manski, 1993) and the endogeneity issue due to self-selection into groups or unobserved group factors.

2.1 Reflection Problem

Without any further information about the interaction structure, the most conventional model employed to study social interaction effect is the linear-in-means model (Manski, 1993), where group mean outcome, $\mathbb{E}(Y_r|r)$ and group mean characteristics, $\mathbb{E}(X_r|r)$ are included to capture the endogenous and contextual effects, respectively:

$$y_{ir} = \lambda_0 \mathbb{E}\left(Y_r|r\right) + \beta_{10} x_{ir} + \beta_{20} \mathbb{E}\left(X_r|r\right) + \varepsilon_{ir}. \tag{1}$$

It can be shown that the two regressors $\mathbb{E}(Y_r|r)$ and $\mathbb{E}(X_r|r)$ are linearly dependent:

$$\mathbb{E}(Y_r|r) = \frac{\beta_{10}+\beta_{20}}{1-\lambda_0}\mathbb{E}(X_r|r). \tag{2}$$

And the reduced form of the model is

$$y_{ir} = \beta_{10}x_{ir} + \left(\frac{\lambda_0\beta_{10}+\beta_{20}}{1-\lambda_0}\right)\mathbb{E}(X_r|r) + \varepsilon_{ir}. \tag{3}$$

This is a representation of the "reflection problem": only β_{10} and $\frac{\lambda_0\beta_{10}+\beta_{20}}{1-\lambda_0}$ can be identified, one cannot separate the endogenous effects coefficient λ_0 from the contextual effects coefficient β_{20}.

To circumvent the "reflection problem", studies that employ the linear-in-means type of model usually focus on either endogenous effects or contextual effects, assuming the other type of social effects are not present. However, some recent papers, including Lee (2007) and Davezies et al. (2009), re-examine group interaction data and provide constructive insights on addressing the "reflection problem". As demonstrated in these studies, the key for solving the "reflection problem" is the exclusion of individual's own outcome/characteristics from the group mean as well as the variation of group size, which breaks down the linear relationship between the endogenous and contextual effects as shown in Equation (2).

Specifically, the model which consists of both endogenous and contextual effects is given as:

$$y_{ir} = \lambda_0\overline{Y}_{-ir} + \beta_{10}x_{ir} + \beta_{20}\overline{X}_{-ir} + \varepsilon_{ir}, \quad r = 1,\cdots,R, \tag{4}$$

where $\overline{Y}_{-ir} = \left(\frac{1}{m_r-1}\sum_{\substack{j=1\\ j\neq i}}^{m_r} y_{jr}\right)$ and $\overline{X}_{-ir} = \left(\frac{1}{m_r-1}\sum_{\substack{j=1\\ j\neq i}}^{m_r} x_{jr}\right)$ and m_r is the size of group r.

Solving for $\overline{Y}_{-ir}$ one can get a different version of Equation (2) as follows:

$$\begin{aligned}\overline{Y}_{-ir} = {} & \frac{\lambda_0}{(m_r-1)-\lambda_0(m_r-2)}y_{ir} + \frac{\beta_{10}(m_r-1)+\beta_{20}(m_r-2)}{(m_r-1)-\lambda_0(m_r-2)}\overline{X}_{-ir} \\ & + \frac{\beta_{20}}{(m_r-1)-\lambda_0(m_r-2)}x_{ir} + \frac{(m_r-1)\overline{\varepsilon}_{-ir}}{(m_r-1)-\lambda_0(m_r-2)},\end{aligned} \tag{5}$$

which is certainly not linear in $\overline{X}_{-ir}$. The reduced form of the model becomes:

$$\begin{aligned}y_{ir} = {} & \frac{\lambda_0\beta_{20}+[(m_r-1)-\lambda_0(m_r-2)]\beta_{10}}{(1-\lambda_0)(m_r-1+\lambda_0)}x_{ir} + \frac{(m_r-1)(\lambda_0\beta_{10}+\beta_{20})}{(1-\lambda_0)(m_r-1+\lambda_0)}\overline{X}_{-ir} \\ & + \frac{\lambda_0(m_r-1)}{(1-\lambda_0)(m_r-1+\lambda_0)}\overline{\varepsilon}_{-ir} + \frac{(m_r-1)-\lambda_0(m_r-2)}{(1-\lambda_0)(m_r-1+\lambda_0)}\varepsilon_{ir}.\end{aligned} \tag{6}$$

Therefore, as long as the group size, m_r, is different across groups in Equation (6), the magnitudes of the social interactions generated in each group will be different. Valuable information about the social interaction coefficients can be obtained from the differences in the social interaction patterns among the groups, which resolves the "reflection problem".

2.2 Correlated Effects

Turning to the second identification challenge, i.e., the endogeneity issue due to self-selection (sorting) into groups and the unobserved environmental factors (correlated effects) that confound with the endogenous and contextual peer effects, a strand of the literature adopts the strategy of including group fixed effects .

Consider a model with both endogenous and contextual effects as well as group fixed effects as follows:

$$y_{ir} = \lambda_0 \overline{Y}_{-ir} + \beta_{10} x_{ir} + \beta_{20} \overline{X}_{-ir} + \alpha_r + \epsilon_{ir}, \quad r = 1, \cdots, R, \tag{7}$$

where α_r denotes the fixed effect for group r. For Equation (7), Lee (2007) derives the between equation

$$\overline{Y}_r = \frac{(\beta_{10} + \beta_{20})}{(1 - \lambda_0)} \overline{X}_r + \frac{1}{(1 - \lambda_0)} \alpha_r + \frac{1}{(1 - \lambda_0)} \bar{\epsilon}_r, \quad r = 1, \cdots, R, \tag{8}$$

and the within equation

$$y_{ir} - \overline{Y}_r = \frac{(m_r - 1)\left(\beta_{10} - \frac{\beta_{20}}{m_r - 1}\right)}{(m_r - 1 + \lambda_0)} (x_{ir} - \overline{X}_r) + \frac{(m_r - 1)}{(m_r - 1 + \lambda_0)} (\epsilon_{ir} - \bar{\epsilon}_r), \tag{9}$$

where $\overline{Y}_r$, $\overline{X}_r$, and $\bar{\epsilon}_r$ are computed from means of all m_r individuals for the r^{th} group. Because Equation (8) contains the unknown group fixed effect α_r, there is no way to identify λ_0, β_{10}, and β_{20} separately from α_r. The possible identification only relies on Equation (9). If all group sizes are the same, i.e., $m_r = m$ for all r, identification fails because there are three structural form parameters but there is only one reduced-form equation. Thus, group size variation – at least three different group sizes – is a necessary condition for identification. This identification idea runs similarly to a fixed effects panel data analysis. In order to apply a within group transformation, variations over time are required, which are the variations between one's own outcomes and the average of the one's own outcomes over time. Here variations between average group outcomes without one's own outcome and average group outcomes with all individuals included (i.e., the fixed effect) are needed for identification.

Following this identification strategy, one can estimate Equation (9) by the conditional maximum likelihood approach or instrument variable approach which are proposed in Lee (2007) and demonstrated in Boucher et al. (2014). Nonetheless, it is important to note that the group size variation may only provide a tenuous source of identification. When the group size is large, the ratio $\frac{m_r-1}{m_r-1+\lambda_0}$ is close to one and λ_0 cannot be easily estimated from Equation (9). The intuition is that when the group is large the individual contribution to the mean is negligible, hence the variations in the leave-one-out group means will be small. In addition, when the group size variations are small or there are separate impacts of group size m_r on the outcome, the identification strategy based on group size variation will not work well.

Another standard strategy to deal with the endogeneity issue is the instrumental variable (IV) method. For instance, Evans et al. (1992) use metropolitan aggregate data as instruments to control for the endogeneity of neighborhood selection problem. The biggest concern about the IV strategy is the validity of the IVs because it is hard to guarantee that they are uncorrelated with the structural errors. Of course, data from experiment, quasi-experiment, and random assignment are effective in dealing with the endogeneity issue due to self-selection into groups and confounding correlated effects. For example, Sacerdote (2001), Zimmerman (2003), and Stinebrickner and Stinebrickner (2006) estimate peer effects on college students' academic outcome using random assignment of roommates at Dartmouth College, Williams College, and Berea college, respectively.

3 Network Data

Network data contain information of specific relationship among pairs of individuals within each group, and such relationship could be formed endogenously. The information on the social network of each individual in a group may be captured by a spatial weights matrix in a spatial autoregressive (SAR) model (Anselin, 1988; Kelejian and Prucha, 1998; Lee, 2004). The individual specific social network structure introduces additional non-linearity for identification of various social interaction effects beyond the variation of group sizes. Examples of using the SAR model with network data to study social interactions include Bandiera and Rasul (2006), Calvó-Armengol et al. (2009), Patacchini and Zenou (2009), Lin (2010), Banerjee et al. (2013), Cohen-Cole et al. (2014), Liu et al. (2014), Fortin and Yazbeck (2015), Hsieh and Lee (2016), Hsieh and Lin (2017) and Hsieh and Van Kippersluis (2018).

3.1 Identification Sources

Denote the element of the $m_r \times m_r$ network weights matrix W_r, $w_{ij,r} = 1$ if individuals j is one of $i's$ k friends; otherwise, $w_{ij,r} = 0$. Note that $w_{ii,r} = 0$ by default. The network interaction model with the endogenous and contextual peer effects, and the group fixed effects is specified as:

$$y_{ir} = \lambda_0 \sum_{j=1}^{m_r} \overline{w}_{ij,r} y_{jr} + \beta_{10} x_{ir} + \beta_{20} \sum_{j=1}^{m_r} \overline{w}_{ij,r} x_{jr} + \alpha_r + \epsilon_{ir}, \quad r = 1, \cdots, R, \tag{10}$$

where $\overline{w}_{ij,r} = \frac{w_{ij,r}}{\sum_j w_{ij,r}}$. To see where the identification comes from, consider the model in the matrix form but without group fixed effects:

$$Y_r = \lambda_0 \overline{W}_r Y_r + X_r \beta_{10} + \overline{W}_r X_r \beta_{20} + \epsilon_r, \quad r = 1, \ldots, R. \tag{11}$$

This model is identified if and only if $\mathbb{E}\left(\overline{W}_r Y_r | X_r\right)$ is not perfectly collinear with the other regressors $\left(X_r, \overline{W}_r X_r\right)$ so that some instruments can be found for the endogenous vector $\overline{W}_r Y_r$. Bramoullé et al. (2009) and Calvó-Armengol et al. (2009) show this condition is equivalent to the condition that the matrices, $I_{m_r}, \overline{W}_r, \overline{W}_r^2$ are linearly independent, where I_{m_r} denotes the identity matrix of dimension m_r. This will be true as long as the network links are intransitive, i.e., some individuals are not friends with her friends' friends. Then the characteristics of peers of peers, $\overline{W}_r^2 X_r$, may serve as instruments for peers' outcomes, $\overline{W}_r Y_r$. For instance, consider three individuals i, j, and k, where i and j are friends, and j and k are friends, whereas i and k are not. Then k's characteristics do not appear in person i's equation as i and k are not friends, and therefore can serve as a valid instrument for j's outcome, which appears in i's equation as the endogenous effect. Most networks have intransitive triads, which generates this type of natural exclusion restrictions induced by the network structure for identifying various social interaction effects in the model.

From Equation (11), the reduced form is

$$\begin{aligned} Y_r &= \left(I_{m_r} - \lambda_0 \overline{W}_r\right)^{-1} \left(\beta_{10} X_r + \beta_{20} \overline{W}_r X_r + \epsilon_r\right) \\ &= \left(I_{m_r} - \lambda_0 \overline{W}_r\right)^{-1} \left(H_r \delta_0 + \epsilon_r\right), \end{aligned} \tag{12}$$

where $H_r = (X_r, \overline{W}_r X_r)$ and $\delta_0 = (\beta_{10}, \beta_{20})'$.

$$\mathbb{E}(\overline{W}_r Y_r | \overline{W}_r, X_r) = \overline{W}_r \left(I_{m_r} - \lambda_0 \overline{W}_r\right)^{-1} (H_r \delta_0). \tag{13}$$

If $|\lambda_0| < 1$, the matrix $\left(I_{m_r} - \lambda_0 \overline{W}_r\right)$ is nonsingular and one can write

$$\left(I_{m_r} - \lambda_0 \overline{W}_r\right)^{-1} = I_{m_r} + \lambda_0 \overline{W}_r + \lambda_0^2 \overline{W}_r^2 + \cdots \tag{14}$$

Therefore,

$$\begin{aligned} \mathbb{E}(\overline{W}_r Y_r) &= \overline{W}_r H_r \delta_0 + \lambda_0 \overline{W}_r^2 H_r \delta_0 + \lambda_0^2 \overline{W}_r^3 H_r \delta_0 + \cdots \\ &= (\overline{W}_r H_r, \overline{W}_r^2 H_r, \overline{W}_r^3 H_r, \cdots)(\delta_0', \lambda_0 \delta_0', \lambda_0^2 \delta_0', \cdots)' \end{aligned} \tag{15}$$

and thus $(\overline{W}_r H_r, \overline{W}_r^2 H_r, \overline{W}_r^3 H_r, \cdots)$ can be used as IVs for $\overline{W}_r Y_r$.

Define $U_r = (H_r, \overline{W}_r H_r, \overline{W}_r^2 H_r, \cdots, \overline{W}_r^p H_r)$ as the IV matrix, then

$$\mathbb{E}(U_r' \epsilon_r) = 0. \tag{16}$$

Let $\epsilon_r(\lambda_0, \delta_0) = Y_r - \lambda_0 \overline{W}_r Y_r - H_r \delta_0$, the empirical counterpart of Equation (16) is

$$U_r' \epsilon_r(\lambda_0, \delta_0) = 0. \tag{17}$$

One can further incorporate quadratic moment conditions in the form, $\epsilon_r(\lambda_0, \delta_0)' A \epsilon_r(\lambda_0, \delta_0)$ with A denotes a constant square matrix in GMM to improve estimation efficiency (Liu and Lee, 2010) so that GMM can be asymptotically as efficient as MLE.

Now consider the matrix form of the model with group fixed effect as in Equation (10):

$$Y_r = \lambda_0 \overline{W}_r Y_r + X_r \beta_{10} + \overline{W}_r X_r \beta_{20} + l_{m_r} \alpha_r + \epsilon_r, r = 1, \ldots, R, \tag{18}$$

where l_{m_r} is the m_r-dimensional vector of ones.

As shown in Lee et al. (2010) and Lin (2010), the group fixed effect can be eliminated by the de-group-mean transformation, i.e., by multiplying Equation (18) with the matrix

$$J_{m_r} = I_{m_r} - \frac{1}{m_r} l_{m_r} l'_{m_r}. \tag{19}$$

Then, the model becomes

$$\hat{Y}_r = \lambda_0 J_{m_r} \overline{W}_r \hat{Y}_r + \hat{X}_r \beta_{10} + J_{m_r} \overline{W}_r \hat{X}_r \beta_{20} + \hat{\epsilon}_r \tag{20}$$

where $\hat{Y}_r = J_{m_r} Y_r, \hat{X}_r = J_{m_r} X_r,$ and $\hat{\epsilon}_r = J_{m_r} \epsilon_r$.

Similarly, this model is identified if and only if $\mathbb{E}\left(J_{m_r} \overline{W}_r \hat{Y}_r | X_r\right)$ is not perfectly collinear with the regressors $\left(\hat{X}_r, J_{m_r} \overline{W}_r \hat{X}_r\right)$ which is equivalent to the matrices, $I_{m_r}, \overline{W}_r, \overline{W}_r^2, \overline{W}_r^3$ are linearly independent. The condition is more demanding because some information has been used to get rid of the fixed effects. Bramoullé et al. (2009) show that if there are two agents i and j in the group separated by a link of distance 3, then $I_{m_r}, \overline{W}_r, \overline{W}_r^2, \overline{W}_r^3$ are linearly independent and the model is identified. This condition fails only for a few networks that have very restrictive structures, such as complete bipartite networks, where the population is divided into two groups such that all individuals in one group are friends with all people in the other group, while no friendship links exist within groups. To estimate the model, one can employ the eigendecomposition technique to eliminate the linear dependency among observations to obtain an effectively independent sample, and then perform the maximum likelihood estimation, as illustrated in Lee et al. (2010) and Lin (2010).

3.2 Endogenous Network Formation

3.2.1 Heckman-type Correction

Although using detailed relationship (link) information solves the "reflection problem", it inevitably introduces another identification problem, which is endogeneity of network matrix W_r. Considering the model in Equation (18), if there were unobserved factors (heterogeneity) affecting both formation of W_r and Y_r, which makes W_r endogenous, then the identification strategy discussed in Section 3.1 will fail. Endogenous linking decisions create selectivity bias in social effect estimates. Context specific features may suggest a way to control for it. For example, Horrace et al. (2016) study peer effects in productivity in the unique situation where a manager selects workers into teams (networks) to produce output. Team membership is perfectly observable, and the manager's selection decisions depend on the combination of individual characteristics at the team level, rather than individual-level characteristics. In team projects, the probability of selecting a worker for the project is not independent across workers. They exploit this interdependency for the identification and estimation of peer effects in network production by implementing a polychotomous Heckman-type correction where the network formation process at the first stage can simply be modeled as the choice of selecting a fixed number of players from a longer list. More recent studies adopt control function methods with more complex first stage network formation models.

Identification in Heckman selection models is notoriously difficult, since the factors driving friendship formation usually also drive behavioral choices. Battaglini et al. (2019c) implement an Heckman correction to study the importance of social networks in explaining U.S. Congress members' productivity using an original instrument: alumni connections. In this network, two politicians are connected if they attended the same educational institution in overlapping periods of time. These connections offer an instrument which rests on plausible assumptions: the network is exogenous to the political process, but still relevant even many years after the Congress members attend school (see Battaglini and Patacchini, 2018).

Goldsmith-Pinkham and Imbens (2013), Hsieh and Lee (2016), Auerbach (2019), and Moon and Johnsson (2019) include unobserved factors in the first stage network formation model. The main idea is to explicitly account for a possible correlation between unobserved factors driving network formation and outcomes, and to exploit the fact that individuals with similar characteristics are more likely to become friends ("homophily" Lazarsfeld et al. (1954)) in modeling network formation. Dyad-specific characteristics are included to reflect the homophily in the model, which are captured by either dummy variables or continuous variables that represent the similarity (or difference) between two individuals with respect to specific characteristics. For example, if two individuals are of the same gender, then the dyad gender dummy will be one, and zero if they are the opposite gender (see Fafchamps and Gubert, 2007b). Hsieh and Lee (2016) model homophily using both observed dyad-specific characteristics and the distance of latent variables between individuals. The distance of the latent variables captures the homophily of unobserved characteristics. The smaller the difference between individuals' unobserved characteristics, the more likely they will become friends. This extension complicates the estimation procedure. Specifically, they propose a network (link) formation model for W_r as follows:

$$P(w_{ij,r}|c_{i,r}, c_{j,r}, c_{ij,r}, z_{i,r}, z_{j,r}, \gamma, \eta) = \frac{\exp(w_{ij,r}\psi_{ij,r})}{1+\exp(\psi_{ij,r})},$$
$$\psi_{ij,r} = \gamma_0 + c_{i,r}\gamma_1 + c_{j,r}\gamma_2 + c_{ij,r}\gamma_3 + \eta_1|z_{i1,r} - z_{j1,r}| + \cdots + \eta_{\bar{d}}|z_{i\bar{d},r} - z_{j\bar{d},r}|. \tag{21}$$

In Equation (21), c_{ir} (c_{jr}) denotes an $\bar{s}$-dimensional vector of observed individual-specific variables; $c_{ij,r}$ is a $\bar{q}$-dimensional vector of observed dyad-specific variables, which can be dummy or continuous variables representing the similarity (difference) between two individuals. The idea of using exogenous variables $C_r = \{c_{ir}, c_{jr}, c_{ij,r}\}$ in explaining link decisions can be found in various empirical studies such as Fafchamps and Gubert (2007a,b). Notice that the individual-specific variables c_{ir} may overlap with x_{ir} used in Equation (18), as $x'_{ir}s$ also capture individual specific variables. However, dyad-specific variables $c_{ij,r}$ are naturally excluded from the set of individual variable x_{ir}, as these dyad-specific variables $c_{ij,r}$ measure the differences between individuals i and j, and none of these differences measures appears in Equation (18). The individual unobserved heterogeneity in Equation (21) is captured by a $\bar{d}$-dimensional vector of latent variables $z_{ir} = (z_{i1,r}, \cdots, z_{i\bar{d},r})$. The distance measures $|z_{id,r} - z_{jd,r}|$, $d = 1, \cdots, \bar{d}$, are used in Equation (21) to capture how differences between individuals (i, j) in these latent variables affect their friendship decisions. $Z_r = (z'_{1,r}, \cdots, z'_{m_r,r})'$ denotes the collection of unobserved latent variables in group r, which is assumed to be independent of C_r. Conditioning on observed variables C_r and unobserved variables Z_r, each network link decision is assumed to be independent and thus for the whole network,

$$P(W_r|C_r, Z_r, \gamma, \eta) = \prod_{i}^{m_r}\prod_{j\neq i}^{m_r} P(w_{ij,r}|C_r, Z_r, \gamma, \eta). \tag{22}$$

Assuming $\epsilon_r = Z_r\kappa_0 + u_r$, where κ_0 is the coefficient reflecting the effect of latent variable Z_r on the outcome equation, and $u_r \sim i.i.d.\ \mathcal{N}_{m_r}(0, \sigma_u^2 I_{m_r})$, where $\mathscr{N}_{m_r}$ represents a multivariate normal distribution of dimension m_r, the SAR model in Equation (18) can be rewritten as

$$Y_r = \lambda_0\overline{W}_r Y_r + X_r\beta_{10} + \overline{W}_r X_r\beta_{20} + Z_r\kappa_0 + l_{m_r}\alpha_r + u_r, \quad r = 1, \ldots, R. \tag{23}$$

Together with the network model in Equation (21), they form a two-equation system, which captures the network formation process and changes in outcomes due to unobservables (i.e., Z_r). The joint probability of Y_r and W_r is:

$$P(Y_r, W_r|X_r, C_r, \theta, \alpha_r) = \int_{Z_r} P(Y_r|W_r, X_r, Z_r, \theta, \alpha_r) \cdot P(W_r|C_r, Z_r, \theta) \cdot f(Z_r)dZ_r, \tag{24}$$

where $\theta = (\gamma', \eta', \lambda_0, \beta'_{10}, \beta'_{20}, \kappa'_0, \sigma_u^2)$.

The SAR model and the network formation model can be jointly estimated by Bayesian methods. First, the prior distributions of the unknown parameters and the unobservables in the models, such as unobserved characteristics $\{Z_r\}$ and groups fixed effects $\{\alpha_r\}$, need to be specified. Then, one can derive the posterior distributions of the parameters based on the specified prior distributions and the likelihood function of the model and simulate random draws by the Markov Chain Monte Carlo (MCMC) sampling to obtain the parameter estimates.

As the dimension of unobserved characteristics in the model is unknown, one can estimate the models with unobservables of different dimensions and choose the one by the model selection criterion, AICM (Akaike's information criterion - Monte Carlo) proposed by Raftery et al. (2007). AICM is the posterior simulation-based analogue of the conventional AIC. Since the maximum loglikelihood value may not be achievable during the Bayesian MCMC procedure so one cannot calculate AIC directly. AICM is derived based on the fact that the posterior distribution of the log-likelihood is approximately a gamma distribution:

$$\ell_{\max} - \ell_t \sim \text{Gamma}(d/2, 1),$$

where $\{\ell_t : t = 1, \cdots, T\}$ is a sequence of loglikelihoods from MCMC posterior draws with a proper thinning such that they are approximately independent. Based on the Gamma distribution, we know $\text{E}(\ell_{\max} - \ell_t) = d/2$ and $\text{Var}(\ell_t) = d/2$. Therefore, we can obtain the moment estimators $\hat{d} = 2s_\ell^2$ and $\hat{\ell}_{\max} = \bar{\ell} + s_\ell^2$, where $\bar{\ell}$ and s_ℓ^2 are the sample mean and variance of the ℓ_t's, respectively. The AICM is then given as

$$\text{AICM} = 2\hat{d} - 2\hat{\ell}_{\max} = 2(s_\ell^2 - \bar{\ell}).$$

3.2.2 Exponential Random Graph Model

Although the link formation model of Hsieh and Lee (2016) explains the unobserved individual heterogeneity during the network formation process, it does not consider the dependence among links. An obvious example of link dependence is the transitivity of friendships: when individuals i and j are both friends of individual k, i and j are likely to be friends too. Allowing link dependence abruptly complicates the link formation process because the conditional probability in Equation (21) would be changed to $P(w_{ij,r}|c_{i,r}, c_{j,r}, c_{ij,r}, z_{i,r}, z_{j,r}, W_{-ij,r})$, with $W_{-ij,r}$ denoting all links in W_r except $w_{ij,r}$. However, due to the dependence of links, one cannot compute the joint probability of W_r directly from the product of conditional likelihoods, i.e., $\prod_{i,j} P(w_{ij,r}|c_{i,r}, c_{j,r}, c_{ij,r}, z_{i,r}, z_{j,r}, W_{-ij,r})$ and thus the estimation becomes difficult. In the literature, the standard approach for modeling link dependence is the exponential random graph model (ERGM) which characterizes the probability of network realization by the Gibbs measure (i.e., an exponential family distribution that depends on the function of the configuration of network) (Frank and Strauss, 1986; Wasserman and Pattison, 1996; Robins et al., 2007). However, except two early works (Snijders, 2001; Butts, 2009), ERGMs typically explain link dependence in an atheoretical way without considering strategic individual decisions.

More recently, Mele (2017) applies the myopic best responses dynamics (Blume, 1993) to model link formation in which the unique stationary distribution can be characterized by the Gibbs measure and thus provides the microfoundations for the ERGM. Specifically, Mele (2017) considers that the utility of network W_r for player i is specified as

$$U_{ir}(W_r, C_r) = \underbrace{\sum_{j=1}^{m_r} w_{ij,r} u_{ij,r}^{\phi_u}}_{\text{direct links}} + \underbrace{\sum_{j=1}^{m_r} w_{ij,r} w_{ji,r} \nu_{ij,r}^{\phi_\nu}}_{\text{mutual links}} + \underbrace{\sum_{j=1}^{m_r} w_{ij,r} \sum_{k \neq i,j}^{m_r} w_{jk,r} h_{ik,r}^{\phi_h}}_{\text{indirect links}} + \underbrace{\sum_{j=1}^{m_r} \sum_{k \neq i,j}^{m_r} w_{ki,r} s_{kj,r}^{\phi_s}}_{\text{popularity}}, \tag{25}$$

where $u_{ij,r}^{\phi_u} \equiv u(c_{ir}, c_{jr}, c_{ij,r}; \phi_u)$; $\nu_{ij,r}^{\phi_\nu} \equiv \nu(c_{ir}, c_{jr}, c_{ij,r}; \phi_\nu)$; $h_{ij,r}^{\phi_h} \equiv h(c_{ir}, c_{jr}, c_{ij,r}; \phi_h)$; and $s_{ij,r}^{\phi_s} \equiv s(c_{ir}, c_{jr}, c_{ij,r}; \phi_s)$ are functions of individual attributes and unknown parameters $\phi = (\phi_u, \phi_\nu, \phi_h, \phi_s)$. Based on the utility function of Equation (25), when individual i creates a link to individual j, he receives a direct benefit measured by $u(c_{ir}, c_{jr}, c_{ij,r}; \phi_u)$; an additional benefit $\nu(c_{ir}, c_{jr}, c_{ij,r}; \phi_\nu)$ if the link is mutual; and externalities $h(c_{ir}, c_{jr}, c_{ij,r}; \phi_h)$

and $s(c_{ir}, c_{jr}, c_{ij,r}; \phi_s)$ from indirect links connected to j and i. By imposing the symmetry restrictions on the benefit of mutual links and externalities from indirect links, i.e., $\nu(c_{ir}, c_{jr}, c_{ij,r}; \phi_\nu) = \nu(c_{jr}, c_{ir}, c_{ij,r}; \phi_\nu)$ and $h(c_{kr}, c_{jr}, c_{kj,r}; \phi_h) = s(c_{kr}, c_{jr}, c_{kj,r}; \phi_s)$ for all i, j, and k in group r, one can model this network formation game as a potential game (Monderer and Shapley, 1996) and summarize the incentives of any individual in any state of the network by a potential function $Q(W_r, C_r, \phi)$,

$$Q(W_r, C_r, \phi) = \sum_{i=1}^{m_r} \sum_{j=1}^{m_r} w_{ij,r} u_{ij,r}^{\phi_u} + \sum_{i=1}^{m_r} \sum_{j>i}^{m_r} w_{ij,r} w_{ji,r} \nu_{ij,r}^{\phi_\nu} + \sum_{i=1}^{m_r} \sum_{j\neq i}^{m_r} \sum_{k\neq i,j}^{m_r} w_{ij,r} w_{jk,r} h_{ik,r}^{\phi_h}. \tag{26}$$

Mele (2017) shows that by assuming the network formation process follows a Markov chain process, with only two individuals i and j meeting at each time t and the meeting probability is determined by a function $\rho_{ij,r}(c_{ir}, c_{jr}, W_{-ij,r}^{t-1})$, which only depends on individual attributes, previous network configuration except link $w_{ij,r}^{t-1}$. When i and j meets at time t, individual i chooses a best response $w_{ij,r}^t = 1$ if and only if

$$U_{ir}(w_{ij,r}^t = 1, W_{-ij,r}^{t-1}; C_r, \phi) + \epsilon_{1t} \geq U_{ir}(w_{ij,r}^t = 0, W_{-ij,r}^{t-1}; C_r, \phi) + \epsilon_{0t}, \tag{27}$$

where ϵ_{1t} and ϵ_{0t} are idiosyncratic shocks received by individual i which are not observed by an econometrican. By further assuming that the shocks are i.i.d. Type-I extreme value distributed, the network formation game converges to a unique stationary distribution $\pi(W_r, C_r, \phi)$,

$$\pi(W_r, C_r, \phi) = \frac{\exp(Q(W_r, C_r, \phi))}{\sum_{\omega_r \in \Omega_r} \exp(Q(\omega_r, C_r, \phi))}, \tag{28}$$

where Ω_r denotes the set of all networks with m_r individuals. Equation (28) takes the form of ERGM, and the probability of observing the network W_r depends on the potential function in Equation (26). To estimate unknown parameters ϕ in Equation (28), a major challenge is to deal with the intractable normalizing constant in the denominator of the likelihood function. Mele (2017) provides a Bayesian MCMC sampling technique – approximate exchange algorithm and Mele and Zhu (2017) suggest an approximate variational estimation approach to estimate Equation (28). Other studies such as Chandrasekhar and Jackson (2014), and Boucher and Mourifié (2017) also investigate network formation by ERGM and propose different estimation procedures. Badev (2013), Boucher (2016), Hsieh et al. (2019a) and Hsieh et al. (2019b) further extend the approach of Mele (2017) to jointly model network formation and discrete and continuous behavior outcomes.

3.3 Heterogeneous Social Interaction Effects

Most existing social interaction studies focus on the average magnitude of peer influences. Some recent studies have started to explore the heterogeneity nature of peer effects, i.e., to investigate how the influences of peers could be different for different types of peers, which provides more informative policy implications related to issues like ability tracking, gender difference, racial difference (Kang, 2007; Griffith and Rask, 2014; Cools et al., 2019). The popular classification of peer types include gender, race, duration of friendships, ability, personality, and others. For instance, Patacchini et al. (2017) study the different roles of short-lived and long-lived network ties in education decisions. Hsieh and Lin (2017) introduce a social interaction model for individual i's outcome moderated by the within and cross gender (or race) peer effects. Given $\bar{p}$ subgroups in group r, the spatial weighting matrix W_r can be divided into $\bar{p}\times\bar{p}$ blocks, $\{W_r^{pq}\}_{p,q=1}^{\bar{p}} = (W_r^{11}, W_r^{12}, \cdots, W_r^{\bar{p}\bar{p}})$, where W_r^{pq} is the matrix of links between subgroups p and q in group r. The SAR model with different coefficients representing heterogeneous peer influences within and across blocks can be specified as:

$$Y_r = \overline{W}_r(\lambda)Y_r + X_r\beta_{10} + \overline{W}_r X_r\beta_{20} + l_{m_r}\alpha_r + \epsilon_r, \quad \epsilon_r \sim i.i.d.\ \mathcal{N}_{m_r}(0, \sigma_\epsilon^2 I_{m_r}), \quad r = 1, \cdots, R, \tag{29}$$

where

$$\overline{W}_r(\lambda) = \begin{pmatrix} \lambda_{11}\overline{W}_r^{11} & \cdots & \lambda_{1\bar{p}}\overline{W}_r^{1\bar{p}} \\ \vdots & \ddots & \vdots \\ \lambda_{\bar{p}1}\overline{W}_r^{\bar{p}1} & \cdots & \lambda_{\bar{p}\bar{p}}\overline{W}_r^{\bar{p}\bar{p}} \end{pmatrix}$$

with $\overline{W}_r^{pq}$ represents the normalized W_r^{pq} by the row-sum.

In Equation (29), λ_{pp}, $p = 1, \cdots, \bar{p}$, on the diagonal of $\overline{W}_r(\lambda)$ represent peer effects within the same subgroups, whereas λ_{pq}, $p \neq q$ represent peer effects across subgroups. Alternatively, Equation (29) can be rewritten as:

$$Y_r = \lambda_{11}\overline{W}_{11,r}Y_r + \cdots + \lambda_{\bar{p}\bar{p}}\overline{W}_{\bar{p}\bar{p},r}Y_r + X_r\beta_1 + \overline{W}_r X_r\beta_2 + l_{m_r}\alpha_r + \epsilon_r, \quad r = 1, \cdots, R, \tag{30}$$

where $\overline{W}_{pq,r}$ is a $m_r \times m_r$ matrix with the corresponding $(p,q)^{\text{th}}$ block equal to $\overline{W}_r^{pq}$ and 0 elsewhere. The terms $\{\overline{W}_{pq,r}Y_r\}_{p,q=1}^{\bar{p}}$ capture the contemporaneous outcomes of peers and the coefficients $\{\lambda_{pq}\}_{p,q=1}^{\bar{p}}$ represent the heterogeneous peer effects. The model in Equation (30) can be easily extended to incorporate heterogeneous contextual effects. However, the computation would be much more demanding given the large number of coefficients to be estimated.

To address the endogenous network formation, one can jointly estimate Equation (30) and the network model in Equations (21) and (22) as a system by Bayesian methods.

An alternative IV approach for estimating heterogeneous endogenous social interaction effects when social interactions take the linear-in-means form has been proposed by Arduini et al. (2019). They consider data from two stage randomization protocols, that is experimental designs where groups are assigned to treatment or control, and a subset of individuals are offered treatment within clusters assigned to treatment according to certain rules. The randomization procedure avoids dealing with endogenous network formation since it guarantees that the share of treated peers is random. In their context social interaction effects are different within treated and untreated individuals and also across groups. They show that heterogeneous externalities can be instrumented using non-linear functions of the shares of treated individuals across groups in a 2SLS estimation framework.

4 Longitudinal Data

4.1 Spatial Dynamic Panel Data Model

With longitudinal data on network and individual outcome, one can further capture the dynamic feature of social interactions through the spatial dynamic panel data (SDPD) model. Let $Y_{rt} = (y_{1rt}, y_{2rt}, \cdots, y_{m_r rt})'$ be the $m_r \times 1$ outcome vector of m_r individuals in group r at time t. $X_{rt} = (x_{1rt}, x_{2rt}, \cdots, x_{m_r rt})'$ denotes the $m_r \times k$ matrix of the exogenous characteristics. w_{ijrt} represents the network link from individual i to individual j, which equals 1 if i connects to j and 0 otherwise. Network links evolve and change over time. W_{rt} is the $m_r \times m_r$ network matrix which summarizes the connections between all m_r individuals in group r at time t. Assume network links are directed and thus, W_{rt} is not symmetric. The SDPD model is specified as follows:

$$Y_{rt} = \lambda_0 \overline{W}_{rt} Y_{rt} + \rho_0 Y_{r,t-1} + \mu_0 \overline{W}_{r,t-1} Y_{r,t-1} + X_{rt}\beta_{10} + \overline{W}_{rt} X_{rt}\beta_{20} + \tau_r + l_{m_r}\alpha_{rt} + \epsilon_{rt}, \quad (31)$$

where $\overline{W}_{rt}$ denotes the row-normalized W_{rt} for $r = 1, 2, \cdots, R$, $t = 1, 2, \cdots, T$. In Equation (31), λ_0, ρ_0, and μ_0 represent the contemporary peer (spillover) effect, persistency of activity outcomes, and temporal spillover or diffusion, respectively; β_{10} and β_{20} are the own and contextual effects of exogenous characteristics, respectively; $\tau_r = (\tau_{1r}, \tau_{2r}, \cdots, \tau_{m_r,r})'$ is the $m_r \times 1$ vector of time-invariant individual effects; $l_{m_r}\alpha_{rt}$ captures the scalar time-varying group effect and $\epsilon_{rt} = (\epsilon_{1t}, \epsilon_{2t}, \cdots, \epsilon_{m_r t})'$ denotes the $m_r \times 1$ vector of stochastic error terms. Specification and estimation of fixed and random effect specification on τ_r can be found in Lee and Yu (2010), Lee and Yu (2012b) and Lee and Yu (2015).

Let $S_{rt}(\lambda_0) = I_{m_r} - \lambda_0 \overline{W}_{rt}$ and $A_{rt}(\lambda_0, \rho_0, \mu_0) = S_{rt}^{-1}(\lambda_0)(\rho_0 I_{m_r} + \mu_0 \overline{W}_{r,t-1})$. The reduced form of Equation (31) is:

$$Y_{rt} = A_{rt}(\lambda_0, \rho_0, \mu_0) Y_{r,t-1} + S_{rt}^{-1}(\lambda_0)(X_{rt}\beta_{10} + \overline{W}_{rt} X_{rt}\beta_{20} + \tau_r + l_{m_r}\alpha_{rt} + \epsilon_{rt}). \quad (32)$$

Following Lee and Yu (2012a), the existence of Equation (32) requires two sets of stability conditions: $S_{rt}(\lambda_0)$ is invertible and $||A_{rt}(\lambda_0, \rho_0, \mu_0)||_\infty < 1$.

If the formation process of the time-varying spatial weights (network) matrix W_{rt} involves unobserved individual variables that correlate with ϵ_{rt}, then $\overline{W}_{rt}$ is potentially endogenous in Equation (31). The subsequent sections study the dynamic network formation model of W_{rt}, and the SDPD model which takes into account the endogeneity of $\overline{W}_{rt}$.

4.1.1 General Dynamic Network Formation Model

To model dynamic networks, Han et al. (2019) propose a dynamic network formation using latent variables to capture unobserved individual characteristics (Hoff et al., 2002; Hoff, 2005; Sarkar and Moore, 2005; Sewell and Chen, 2015). Let $z_{irt} = (z_{i1rt}, z_{i2rt}, \cdots, z_{i\bar{p}_1 rt})'$ and $v_{irt} = (v_{i1rt}, v_{i2rt}, \cdots, v_{i\bar{p}_2 rt})'$ denote two different vectors of latent variables with dimensions $\bar{p}_1$ and $\bar{p}_2$, respectively, which capture an individual i's position (unobserved characteristics) in group r at time t. z_{irt} and v_{irt} are assumed to be independent. $Z_{rt} = (z_{1rt}, z_{2rt}, \cdots, z_{m_r rt})'$

denotes the $m_r \times \bar{p}_1$ matrix of z_{irt}'s and $V_{rt} = (v_{1rt}, v_{2rt}, \cdots, v_{m_r rt})'$ denotes the $m_r \times \bar{p}_2$ matrix of v_{irt}'s. Given the binary network links w_{ijrt}'s, the probability of observing W_{rt} can be parameterized in a logit form as follows:

$$P(W_{rt}|Z_{rt}, V_{irt}, W_{r,t-1}, Y_{r,t-1}, \Gamma) = \prod_{i=1}^{m_r} \prod_{\substack{j=1 \\ j \neq i}}^{m_r} \left(\frac{\exp(w_{ijrt}\psi_{ijrt})}{1+\exp(\psi_{ijrt})} \right), \tag{33}$$

where Γ denotes the set of parameters in the score function ψ_{ijrt}, which is specified as follows,

$$\begin{aligned} \psi_{ijrt} = & \gamma_0 + \underbrace{c'_{irt}\gamma_1 + c'_{jrt}\gamma_2 + c'_{ijrt}\gamma_3}_{\text{direct and homophily effects on obs. chara.}} + \underbrace{\gamma_4 w_{ijr,t-1} + \gamma_5 \varpi_{ijr,t-1}}_{\text{persistency and transitivity}} \\ & + \underbrace{\gamma_6 y_{ir,t-1} + \gamma_7 y_{jr,t-1} + \gamma_8 |y_{ir,t-1} - y_{jr,t-1}|}_{\text{direct and homophily effects from activity outcomes}} \\ & + \underbrace{\sum_{p_1=1}^{\bar{p}_1} \delta_{p_1} |z_{ip_1r,t} - z_{jp_1r,t}|}_{\text{unobserved homophily}} + \underbrace{\sum_{p_2=1}^{\bar{p}_2} \xi_{p_2} v_{ip_2rt} + \sum_{p_2=1}^{\bar{p}_2} \zeta_{p_2} v_{jp_2rt}}_{\text{unobserved degree heterogeneity}} . \end{aligned} \tag{34}$$

The specification of ψ_{ijrt} in Equation (34) includes several important network features. First, the direct and homophily effects on the observed characteristics are captured by two $\bar{l}_1 \times 1$ vectors of exogenous individual-specific regressors, c_{irt} and c_{jrt}. The $\bar{l}_2$-dimensional vector of the dyad-specific regressors c_{ijrt} between individuals i and j describes the similarity (or difference) on their characteristics. Note that c_{irt} (c_{jrt}) can be overlapped with x_{irt} (x_{jrt}) as both capture individual level characteristics, but dyad-specific variables c_{ijrt} are naturally excluded from the set of individual level regressors used in the SDPD model of Equation (31), since dyad-specific variables c_{ijrt} capture the distances between individual pairs.

Second, the persistency of the network links and transitivity are captured by $w_{ijr,t-1}$ and the number of common friends shared by i and j, $\varpi_{ijrt} = \sum_{l=1}^{m_r} w_{ilrt} \times w_{jlrt}$ (Krivitsky et al., 2009), respectively. Graham (2016) and Patacchini et al. (2017) use $w_{ijr,t-1}$ to capture the effect of the "state-dependence" in link formation (i.e., the effect of past links on the present link formation). And for undirected links, Graham (2016) use $\varpi_{ijr,t-1}$ to reflect an individual's taste for transitive links. Both $w_{ijr,t-1}$ and $\varpi_{ijr,t-1}$ are realizations from previous networks so there is no statistical coherence problem on the probability function of Equation (33).

Third, it incorporates the direct and homophily effects with respect to activity outcomes as well as unobserved latent variable z_{ip_1rt}. The lagged dependent variables $y_{ir,t-1}$ and $y_{jr,t-1}$ separately capture the effects of past activity outcomes on sending and receiving network links. The difference between the lagged outcomes (i.e, $|y_{ir,t-1} - y_{jr,t-1}|$) captures the homophily with respect to activity outcomes. And $|z_{ip_1rt} - z_{jp_1rt}|$ captures the difference in the unobservables and the coefficients δ_{p_1}'s are expected to be negative.

Lastly, the degree heterogeneity is captured by time-varying individual latent variable v_{ip_2rt}. Graham (2017) and Dzemski (2019) argue that in many social networks, few individuals are well-known "hubs" with many links, whereas others only own a few links. This individual heterogeneity is captured by v_{ip_2rt} in the score function. Moreover, ξ_{p_2} and ζ_{p_2} represent the asymmetric effects for senders and receivers, respectively (Krivitsky et al., 2009). Equation (34) indicates that the set of parameters in ψ_{ijrt} is $\Gamma = (\gamma', \delta', \xi', \zeta')'$ with $\gamma = (\gamma_0, \gamma_1', \gamma_2', \gamma_3', \gamma_4, \gamma_5, \gamma_6, \gamma_7, \gamma_8)'$, $\delta = (\delta_1, \cdots, \delta_{\bar{p}_1})'$, $\xi = (\xi_1, \cdots, \xi_{\bar{p}_2})'$ and $\zeta = (\zeta_1, \cdots, \zeta_{\bar{p}_2})'$.

Following Sewell and Chen (2015), one can consider a Markov process for z_{irt} and v_{irt} to capture the dependency over time, with initial distributions as follows:

$$P(Z_{r1}|\sigma_{z_0}^2) = \prod_{i=1}^{m_r} \phi(z_{ir1}|0, \sigma_{z_0}^2 I_{\bar{p}_1}),\ P(V_{r1}|\sigma_{v_0}^2) = \prod_{i=1}^{m_r} \phi(v_{ir1}|0, \sigma_{v_0}^2 I_{\bar{p}_2}), \tag{35}$$

with ϕ denotes the normal density and the transition equations as follows:

$$P(Z_{rt}|Z_{r,t-1}, \sigma_z^2) = \prod_{i=1}^{m_r} \phi(z_{irt}|z_{ir,t-1}, \sigma_z^2 I_{\bar{p}_1}),\ P(V_{rt}|V_{r,t-1}, \sigma_v^2) = \prod_{i=1}^{m_r} \phi(v_{irt}|v_{ir,t-1}, \sigma_v^2 I_{\bar{p}_2}), \tag{36}$$

where $t = 2, 3, \cdots, T$. Also note that z_{irt} and v_{irt} are independent across individuals and dimensions. Let $\bar{p} = \bar{p}_1 + \bar{p}_2$, $h_{irt} = (z_{irt}', v_{irt}')'$ and $H_{rt} = (h_{1rt}, h_{2rt}, \cdots, h_{m_r rt})' = (Z_{rt}, V_{rt})$. Given Equations (35) and (36), the initial distribution and transition equation of H_{rt} are as follows:

$$P(H_{r1}|\Sigma_{h_0}) = \prod_{i=1}^{m_r} \phi\left(h_{ir1}|0, \Sigma_{h_0}\right) = \prod_{i=1}^{m_r} \phi\left(z_{ir1}|0, \sigma_{z_0} I_{\bar{p}_1}\right) \times \prod_{i=1}^{m_r} \phi\left(v_{ir1}|0, \sigma_{v_0}^2 I_{\bar{p}_2}\right), \tag{37}$$

and

$$p(H_{rt}|H_{r,t-1}, \Sigma_h) = \prod_{i=1}^{m_r} \phi\left(h_{irt}|h_{ir,t-1}, \Sigma_h\right) = \prod_{i=1}^{m_r} \phi\left(z_{irt}|z_{ir,t-1}, \sigma_z^2 I_{\bar{p}_1}\right) \times \prod_{i=1}^{m_r} \phi\left(v_{irt}|v_{ir,t-1}, \sigma_v^2 I_{\bar{p}_2}\right), \tag{38}$$

for $t = 2, 3, \cdots, T$, where $\Sigma_{h_0} = \begin{pmatrix} \sigma_{z_0}^2 I_{\bar{p}_1} & O \\ O & \sigma_{v_0}^2 I_{\bar{p}_2} \end{pmatrix}$ is the corresponding variance-covariance matrix for the initial period and $\Sigma_h = \begin{pmatrix} \sigma_z^2 I_{\bar{p}_1} & O \\ O & \sigma_v^2 I_{\bar{p}_2} \end{pmatrix}$ is the variance-covariance matrix of h_{irt}'s.

4.1.2 Correct Network Endogeneity in SDPD Model

Denote $\kappa_0 = (\kappa_{10}', \kappa_{20}')'$. Let $\epsilon_{rt} = Z_{rt}\kappa_{10} + V_{rt}\kappa_{20} + u_{rt} = H_{rt}\kappa_0 + u_{rt}$, where $u_{rt} = (u_{1t}, u_{2t}, \cdots, u_{m_r t})'$ with $u_{it} \sim i.i.d.\ \mathcal{N}_{m_r}(0, \sigma_u^2 I_{m_r})$. Then, the SDPD model of Equation (31) can be rewritten as:

$$Y_{rt} = \lambda_0 \overline{W}_{rt} Y_{rt} + \rho_0 Y_{r,t-1} + \mu_0 \overline{W}_{r,t-1} Y_{r,t-1} + X_{rt}\beta_{10} + \overline{W}_{rt} X_{rt} \beta_{20} + H_{rt}\kappa_0 + \tau_r + l_{m_r}\alpha_{rt} + u_{rt}, \tag{39}$$

where $r = 1, 2, \cdots, R$ and $t = 1, 2, \cdots, T$. The term $H_{rt}\kappa_0$ acts as a control function (Navarro, 2008; Wooldridge, 2015) to deal with the endogeneity issue of $\overline{W}_{rt}$. Compared to the static model of Equation (23) in Section 3.2 where time-invariant individual unobserved heterogeneity is captured by Z_r, Equation (39) employs τ_r and H_{rt} to capture time-invariant and time-varying unobserved individual heterogeneity, respectively. Equation (39) and general network formation model of Equations (33) and (34) can be jointly estimated by the Bayesian MCMC approach similar to the procedure discussed in Section 3.2.1, the details of which can be found in Han et al. (2019).

4.2 Stochastic Actor-based Model

Another well-known model for network dynamics is the stochastic actor-based model proposed by Snijders (1996), Snijders (2001) and Snijders et al. (2010a) in the statistical network literature, which explores the longitudinal data to differentiate social influence effect from homophily effect. From cross-sectional data, separately identifying social influence and homophily effects is difficult because the two effects exist simultaneously and complement each other. In the stochastic actor-based model, the simultaneity issue is resolved by utilizing the time lagged observations of network or outcomes.

4.2.1 Continuous-Time Markov Process Assumption

Assuming that researchers observe network, W_t, and H-dimensional behavioral outcomes, $y_t = (y_{1t}, \cdots, y_{Ht})$, of m individuals at two or more discrete points in time $t \in (t_1, t_2, \cdots, t_L | t_1 < t_2 < \cdots < t_L)$. The group subscripts are suppressed here for ease of presentation. Network W_t is represented by a $m \times m$ adjacency matrix, with each entry $w_{ij,t}$ equal to one if actor i connects to actor j; otherwise, it is equal to zero. Network links in W_t are assumed directed; thus, W_t is asymmetric. The discussion for undirected network links can be found in Snijders and Pickup (2017). Each y_{ht} is a $m \times 1$ vector, and element $y_{i,ht}$ is a binary variable that is equal to one if actor i performs behavior h, such as smoking, drinking, or using drugs; zero, otherwise. While only binary behavioral variables are discussed in this subsection, polychotomous behavioral variables are considered in Snijders et al. (2007). The stochastic actor-based model assumes that there are "micro" steps taking place at stochastically-determined moments between any two discrete time points t_ℓ and $t_{\ell+1}$, at which individuals can alter their network ties or behaviors. Denote the state variable $S_t = (W_t, y_t)$. Note that changes in the future state S_{t+r}, $r > 0$, depend only on the current S_t because of the Markov chain property.

In order to provide causal interpretations of the homophily effect and social influence, the model assumes that individual changes in network ties and behaviors are conditionally independent of each other at the given state of the process, and separates the co-evolutions of network ties and behaviors into network formation and social influence processes, respectively. Furthermore, to form a parsimonious and relatively simple Markov process, the model assumes that only one network tie or behavioral variable can be changed, and only by one unit, at a time.

The core of the model is a micro-mechanism where actors execute myopic changes on network ties or behaviors to optimize his/her utility function, which consists of an evaluation function $f_i(\beta, S_t)$, which depends on the current state configuration and the parameter vector β, and a stochastic error from the extreme type-I distribution. A Poisson process with the rate function $\rho_{i,\ell}$ determines when actor i should execute changes in the time period $t_\ell \leq t \leq t_{\ell+1}$. For simplicity, it is assumed that the rate functions of changing either network ties or behaviors are homogeneous for all m members in the same time period (i.e., $\rho^{[w]}_{i,t_\ell} = \rho^{[w]}_{t_\ell}$ and $\rho^{[y_h]}_{i,t_\ell} = \rho^{[y_h]}_{t_\ell}$ for all i's). For the network tie, the probability of actor i choosing to change the network tie with actor j, which is from $w_{ij,t}$ to $1 - w_{ij,t}$, takes the binary logit form,

$$\Pr(1 - w_{ij,t}, W_{-ij,t}|S_t) = \frac{\exp(f_i^{[w]}(\beta^{[w]}, 1 - w_{ij,t}, W_{-ij,t}, y_t))}{\exp(f_i^{[w]}(\beta^{[w]}, w_{ij,t} = 1, W_{-ij,t}, y_t)) + \exp(f_i^{[w]}(\beta^{[w]}, w_{ij,t} = 0, W_{-ij,t}, y_t))}, \tag{40}$$

where $W_{-ij,t}$ stands for W_t which excludes $w_{ij,t}$. And the probability of changing the behavioral variable $y_{i,ht}$, which is from $y_{i,ht}$ to $1 - y_{i,ht}$, is given by

$$\begin{aligned} &\Pr(1 - y_{i,ht}, y_{-i,ht}|S_t) \\ &\quad = \frac{\exp(f_i^{[y_h]}(\beta^{[y_h]}, 1 - y_{i,ht}, y_{-i,ht}, W_t))}{\exp(f_i^{[y_h]}(\beta^{[y_h]}, y_{i,ht} = 0, y_{-i,ht}, W_t)) + \exp(f_i^{[y_h]}(\beta^{[y_h]}, y_{i,ht} = 1, y_{-i,ht}, W_t))}, \end{aligned} \tag{41}$$

where $y_{-i,ht}$ stands for y_{ht} which excludes $y_{i,ht}$.

If the stationary transition probability of the state variable S from this continuous-time Markov chain exists, then it is fully described by the matrix of transition intensities:

$$q(s, s') = \lim_{d_t \downarrow 0} \frac{\Pr(S_{t+d_t} = s'|S_t = s)}{d_t}, \tag{42}$$

where s and s' denote the current and next state, respectively. Specifically,

$$q(s, s') = \begin{cases} \rho^{[w]} \Pr(1 - w_{ij,t}, W_{-ij,t}|S_t = s) & \text{if } s' = (1 - w_{ij,t}, W_{-ij,t}, y_t), \\ \rho^{[y_h]} \Pr(1 - y_{i,ht}, y_{-i,ht}|S_t = s) & \text{if } s' = (1 - y_{i,ht}, y_{-i,ht}, W_t), \\ -\sum_i \Big\{ \sum_{j \neq i} \rho^{[w]} \Pr(1 - w_{ij,t}, W_{-ij,t}|S_t = s) + & \\ \qquad \sum_h \rho^{[y_h]} \Pr(1 - y_{i,ht}, y_{-i,ht}|S_t = s) \Big\} & \text{if} s' = s, \\ 0 & \text{otherwise.} \end{cases} \tag{43}$$

Define the evaluation functions for network ties and behaviors in Equation (40) and Equation (41), respectively as follows

$$f_i^{[w]}(\beta^{[w]}, S_t) = \beta^{[w]} g_i^{[w]}(S_t), \tag{44}$$

$$f_i^{[y_h]}(\beta^{[y_h]}, S_t) = \beta^{[y_h]} g_i^{[y_h]}(S_t), \quad h = 1, \cdots, H, \tag{45}$$

By properly selecting $g_i^{[w]}(S_t)$ and $g_i^{[y_h]}(S_t)$ in the evaluation functions, the stochastic actor-based model can capture the homophily effect, social influence, and other effects from the network structure and exogenous variables. In particular, the social influence effect can be captured by the "average similarity effect", which expresses the conformity motive of an individual in behaving as his/her peers do, particularly when enough peers behave in the same manner (López-Pintado, 2008; Young, 2009). For actor i, the average similarity effect on behavior y_h is defined as

$$g_i^{[y_h]}(S_t) = w_{i+,t}^{-1} \sum_{j \neq i} w_{ij,t}(sim_{ij}^{[y_{ht}]} - \widehat{sim}^{[y_{ht}]}), \tag{46}$$

where $w_{i+,t} = \sum_{j \neq i} w_{ij,t}$, $sim_{ij}^{[y_{ht}]} - \widehat{sim}^{[y_{ht}]}$ is the centralized similarity score with $sim_{ij}^{[y_{ht}]} = 1 - |y_{i,ht} - y_{j,ht}|$ and $\widehat{sim}^{[y_{ht}]}$ is the mean of similarity scores across all pairs. The parameter $\beta_k^{[y_h]}$ for the average similarity effect reflects the size of the social influence effect. The homophily effect of behavior y_h on the evaluation function of network ties can also be captured by the "similarity effect":

$$g_i^{[w]}(S_t) = \sum_{j \neq i} w_{ij,t}(sim_{ij}^{[y_{ht}]} - \widehat{sim}^{[y_{ht}]}). \tag{47}$$

4.2.2 Model Estimation and Identification

Although the likelihood function of the stochastic actor-based model with the probability structure governed by the continuous-time Markov process can be used to estimate the unknown parameters $\theta = (\rho, \beta)$ in the rate and evaluation functions, the resulting likelihood function does not have a closed form, which renders implementing the maximum likelihood (ML) and Bayesian methods difficult (Koskinen and Snijders, 2007; Snijders et al., 2010b).

Snijders (2001) and Snijders et al. (2007) propose to estimate the model through the method of moments (MOM). Specifically, let $\mu(S)$ denote a vector of statistics based on the state variable S. MOM estimators are obtained by solving the moment equations in which the expected and observed sample statistics resemble each other (i.e., $\mathrm{E}_\theta(\mu(S)) = \mu(s)$, where $\mu(s)$ denotes the observed ones). Given that the expected sample statistics cannot be calculated explicitly, they are replaced by the averages of the statistics based on auxiliary samples simulated from the model. According to the micro-steps of the network and behavior evolution outlined in section 4.2.1, the changes in network ties and behaviors between any two network observations can be simulated from the model conditioning on the initial configuration of S.

The vector $\mu(S)$ has the same dimension as the parameter vector θ, which should be chosen so that a corresponding element in $\mu(S)$ sensitively responds to changes in each separated parameter in θ. Substantial variations in network ties and behaviors over time, which are reflected in $\mu(S)$, are necessary for parameter identification. The rate parameters (i.e., $\rho_\ell^{[w]}$ and $\rho_\ell^{[y_h]}$) influence only the Poisson process between t_ℓ and $t_{\ell+1}$. Thus, the moment equation is as follows:

$$\mathrm{E}_\theta \left\{\mu_\ell(S_{t_\ell}, S_{t_{\ell+1}})|S_{t_\ell} = s_{t_\ell}\right\} = \mu_\ell(s_{t_\ell}, s_{t_{\ell+1}}) \tag{48}$$

and the choices of $\mu_\ell(S_{t_\ell}, S_{t_{\ell+1}})$ in Equation (48) are $\sum_{ij} |w_{ij,t_{\ell+1}} - w_{ij,t_\ell}|$ and $\sum_i |y_{ih,t_{\ell+1}} - y_{ih,t_\ell}|$ for estimating $\rho_\ell^{[w]}$ and $\rho_\ell^{[y_h]}$, respectively. The parameters β in the evaluation functions are constant and appear in all of the statistics $\mu_\ell(S_{t_\ell}, S_{t_{\ell+1}})$, $\ell = 1, \cdots, L-1$. Thus, the empirical moment equation is

$$\sum_{\ell=1}^{L-1} \mathrm{E}_\theta \left\{\mu_\ell(S_{t_\ell}, S_{t_{\ell+1}}|S_{t_\ell} = s_{t_\ell})\right\} = \sum_{\ell=1}^{L-1} \mu_\ell(s_{t_\ell}, s_{t_{\ell+1}}), \tag{49}$$

and the sample statistics used to estimate $\beta^{[w]}$ and $\beta^{[y_h]}$ are based on $\sum_{i=1}^{m} g_i(S)$ with $g_i(S)$ from Equations (46) and (47). Increasing $\beta_\ell^{[w]}$ (or $\beta_\ell^{[y_h]}$) gives a larger influence of similarity effect on individual's evaluation function, which causes a higher chance of forming a link (or adopting a product) and leads to a higher similarity effect on all the actors in the subsequent time moments. Although these two sample statistics respond well to the change in $\beta^{[w]}$ and $\beta^{[y_h]}$, they are perfectly multicollinear and produce two identical moment equations. Following the concept of causality, Snijders et al. (2007) propose to exploit the time order of variables to prevent under-identification. In particular, homophily is reflected as a "later" change in the network tie, following the "earlier" configuration of behaviors, while social influence is reflected as a "later" change in behaviors following the "earlier" configuration of network ties. The homophily effect on network formation is estimated using the following sample statistic:

$$\mu_\ell(S_{t_\ell}, S_{t_{\ell+1}}) = \sum_i g_i^{[w]}(w_{t_{\ell+1}}, y_{t_\ell}). \tag{50}$$

The social influence effect on behavior y_h is estimated using

$$\mu_\ell(S_{t_\ell}, S_{t_{\ell+1}}) = \sum_i g_i^{[y_h]}(y_{h,t_{\ell+1}}, W_{t_\ell}). \tag{51}$$

Recently, Amati et al. (2019) extend the usual cross-lagged statistics to additional contemporaneous statistics depending jointly on the network and the behaviour and propose the generalized method of moments for estimation.

5 Online Social Network Data

During the past years, there has been extraordinarily rapid growth of online social networks, such as Facebook, Twitter, Yelp, and so on. The unprecedented ongoing innovation in Social networking technologies has made sharing daily activity through various online social medias almost costless and instantly. With the pervasiveness of online networking, it is very natural to ask how individuals' behaviors and decision making are influenced by their online friends. Understanding how online social influences work will be of great importance in improving marketing effectiveness, promoting desirable behaviors among large population, and so on. Given the limited availability of survey social network data, the ample availability of online social media data opens up a promising and valuable window for analyzing social interaction effects.

However, the estimation of online social interaction effects faces the same identification challenges as the survey data. Furthermore, the large sample size of social media data creates formidable computational burden. To address the computational issue and to apply the structural social interaction model which takes cares of both the reflection problem and the endogenous network formation as discussed in the previous subsection, two approaches can be employed. First, one can identify the clique (a clique is the maximum number of actors who have friendship ties present among themselves) or the component structure in the network. The focus will be on social influence effect within the clique or component. A potential problem is that the clique/component may still be too large due to the "tree" structure of the online network. The second approach is big data and machine learning related, which employs network community detection algorithm to identify the relevant network community. One possible network building algorithm is the iterative edge-removal approach in Newman and Girvan (2004). This method continuously finds and removes the edge with the largest betweenness measure to separate a network. Because the betweenness measures are recalculated after each removal, the computation is intense. Alternatively, Blondel et al. (2008) propose the Louvain method which involves two iterative steps. In the initial step, each node is placed in its own community. Then gains (or losses) of modularity (a scale value between -1 and 1 that measures the density of edges inside communities to edges outside communities) are calculated by moving each node into each of its neighbor's community. Associated with the largest modularity gain, the node will be merged into the nearby community. The second step regards the communities built in the previous step as new nodes and constructs inter- and intra-node links by the weighted regular links and self-loops. The steps are iterated until there is no modularity-increasing reassignment of communities. The Louvain method has been applied to community detection in web networks with 118 million nodes and 1 billion links (Blondel et al., 2008), Twitter social network with 2.4 million nodes and 38 million links (Pujol et al., 2009) and the mobile phone network with 4 million nodes and 100 million links (Greene et al., 2010).

After network communities are found, if certain community sizes (m_r) are still too large to directly apply the network interactions model in Equation (18) and the network formation model in Equation (21), one can further adopt the case-control approximate likelihood approach in Raftery et al. (2012). The concept of the case-control approach is adopted from Epidemiology where researchers intend to compare the group with the outcome of interest (case) to the control group. Since the cases are usually rare so that the case-control approach collects all available cases and samples from the corresponding control group. Even if the pairwise independence in the conditional distribution of network links of Equation (21) allows us to compute the joint likelihood in a convenient way, the required computation is still heavy when the network size is large. The case-control approach proposed by Raftery et al. (2012) could further alleviate the computational burden by reducing the computational cost from $O(m_r^2)$ to $O(m_r)$.

Given the probability function in Equation (22), the log likelihood function can be written as

$$\ell(W_r|C_r, Z_r, \gamma, \eta) = \sum_{i=1}^{m_r} \ell_i(w_{i.,r}|C_r, Z_r, \gamma, \eta),$$

where

$$\begin{aligned}
\ell_i(w_{i.,r}|C_r, Z_r, \gamma, \eta) &= \sum_{j=1}^{m_r} (w_{ij,r}\psi_{ij,r} - \ln(1 + \exp(\psi_{ij,r}))) \\
&= \sum_{j=1, w_{ij,r}=1}^{m_r} (\psi_{ij,r} - \ln(1 + \exp(\psi_{ij,r}))) + \sum_{j=1, w_{ij,r}=0}^{m_r} (-\ln(1 + \exp(\psi_{ij,r}))) \\
&= \ell_{i,1} + \ell_{i,0}.
\end{aligned}$$

Since the network links are sparse, the quantity $\ell_{i,0}$ can be viewed as a population total statistics. This population total can be estimated by a random sample of the population,

$$\widetilde{\ell}_{i,0} = \frac{m_{i0,r}}{n_{i0,r}} \sum_{j=1}^{n_{i0,r}} (-\ln(1 + \exp(\psi_{ij,r}))),$$

where $m_{i0,r}$ is the total number of zero's in the i^{th} row of the upper triangle of matrix W_r, and $n_{i0,r}$ is the number of samples selected from zero entries in the i^{th} row of the upper triangle of matrix W_r. $\widetilde{\ell}_{i,0}$ is an unbiased estimator of $\ell_{i,0}$ given the random samples. When the network size is large, one can choose a small $n_{i0,r}$ to compute $\widetilde{\ell}_{i,0}$ and reduce the amount of computation.

An alternative approach to distinguish peer influence and homophily effects in online social networks has been proposed by Easley et al. (2019). They infer preferences using a machine learning algorithm applied to previous adoption decisions, and then match agents using those inferred preferences. They develop a dynamic matched sample estimation algorithm that is suitable to estimate social interaction effects in product adoption in environments where networks are large, dynamic and there are numerous items diffusing simultaneously. Related methodologies in the computer science literature include Aral et al. (2009) and Sharma and Cosley (2016).

6 Estimating Social Interaction Effects Using Outcomes

In many environments social networks are not directly observed: only the activities of their members are observable and usable for estimation. The issue of estimating an *unobservable* social network using observable economic outcomes is that the number of network connections is much higher than the number of economic outcomes that can be used for the estimation: with m agents observed for T sessions, the first is on the order of m^2, the second just of order $m \cdot T$.

There is a small (but rapidly growing) literature looking at possible solutions. Two approaches have been proposed so far. The first is based on high-dimensional estimation techniques. The key assumption that leads to identification is called sparsity assumption, which is essentially a regularity condition on the topology of the network structure. De Paula et al. (2018) use the Adaptive Elastic Net GMM method (Caner and Zhang, 2014) which imposes a sparsity condition such that each agent may only connect with a few other agents on the network either directly or indirectly. Manresa (2013) focuses on the contextual effects and impose the assumption that contextual effects are sparse among agents on the network. Battaglini et al. (2019a) propose a Graphical LASSO estimator with normal errors that does not request to impose restrictions on the number of directly or indirectly connected agents. Instead, they assume the total number of links in the network is sparse. This allows more flexible network structures to be considered such as "star" shaped network. While identification relies on parametric assumptions, the assumption on the normality of the errors and the relaxation of the assumption of a row-sum normalization for the adjacency matrix provide more flexibility to incorporate correlated effects.

Importantly, these papers provide conditions for which the parameters of the social interaction model are consistently estimated when using the recovered network. This is a unique feature of the proposed methodologies since even the most sophisticated models of network formation in the existing literature (such as ERGMs) have no ability to estimate the impact on outcomes of the interactions mediated through the networks. Breza et al. (2017) propose a different network reconstruction procedure. Instead of being based on outcomes, this network elicitation procedure uses aggregated relational data, that is, responses to survey questions asking agents to report, for instance, the number of their social connections.

In practice, those methods require observing outcomes for a number of time periods that is very large (compared to the number of units, $m^2/T \to 0$). The implicit assumption is that the latent network remains stable over many time periods. This is a strong assumption for applications in environments where social networks change rapidly over time.

An alternative approach has been proposed by Battaglini et al. (2019b). They rely on the structure generated by a theoretical model of endogenous link formation to reduce the dimensionality of the estimation; and then to use the observable outcomes at just one or few points in time to structurally estimate the parameters of the model. And they use the analytical characterization of the equilibrium conditions to estimate the model by Bayesian methods. Because the analytical characterization of the equilibrium conditions makes it impossible to state an analytic likelihood function, the model is estimated by an *Approximate Bayesian Computation* method (henceforth, ABC), a computational approach that has proven useful in population genetics and other applications that require large scale models. Whether the characterization of the equilibrium is an acceptable compromise it is ultimately an empirical question that has to do with the ability of the model to fit the data better than alternative approaches.

7 Summary

Recent advances in data collection that allow researchers to better measure social connections among social actors, and advances in theoretical modeling have stimulated important progress on empirical methods to estimate social interaction effects. Still, important methodological challenges, such as problems concerning the endogeneity and the observability of social connections, have only started to be addressed.

This chapter reviews recent work studying how to estimate endogenous social interaction effects under different data structures. If there is one overarching message to emerge from this chapter, it is that the use of the existing econometric techniques cannot be mechanical. As discussed in this chapter, there are a variety of challenges and various possible solutions. Ultimately, the choice of the most appropriate model, identification and estimation strategy depends on the context under analysis and on the postulated mechanism underlying the presence of social interaction effects.

References

Advani, Arun and Bansi Malde (2018) "Credibly identifying social effects: Accounting for network formation and measurement error," *Journal of Economic Surveys*, Vol. 32, pp. 1016–1044.

Amati, Viviana, Felix Schönenberger, and Tom AB Snijders (2019) "Contemporaneous Statistics for Estimation in Stochastic Actor-Oriented Co-evolution Models," *Psychometrika*, Vol. 84, pp. 1068–1096.

Anselin, Luc (1988) *Spatial Econometrics: Methods and Models*, Vol. 4: Springer Science & Business Media.

Aral, Sinan, Lev Muchnik, and Arun Sundararajan (2009) "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *Proceedings of the National Academy of Sciences*, Vol. 106, pp. 21544–21549.

Arduini, Tiziano, Eleonora Patacchini, and Edoardo Rainone (2019) "Treatment Effects with Heterogeneous Externalities," *Journal of Business & Economic Statistics*, pp. 1–26.

Auerbach, Eric (2019) "Identification and estimation of a partially linear regression model using network data," *arXiv preprint arXiv:1903.09679.*

Badev, Anton (2013) "Discrete games in endogenous networks: Theory and policy," *Working paper.*

Bandiera, Oriana and Imran Rasul (2006) "Social networks and technology adoption in northern Mozambique," *Economic Journal*, Vol. 116, pp. 869–902.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson (2013) "The diffusion of microfinance," *Science*, Vol. 341, p. 1236498.

Battaglini, M., F. W. Crawford, E. Patacchini, and S. Peng (2019a) "Inferring networks from outcomes," *Working paper.*

Battaglini, Marco and Eleonora Patacchini (2018) "Influencing connected legislators," *Journal of Political Economy*, Vol. 126, pp. 2277–2322.

Battaglini, Marco, Eleonora Patacchini, and Edoardo Rainone (2019b) "Endogenous Social Connections in Legislatures," *NBER Working Paper No. w25988.*

Battaglini, Marco, Valerio Leone Sciabolazza, and Eleonora Patacchini (2019c) "Effectiveness of connected legislators," *American Journal of Political Science, forthcoming.*

Blondel, Vincent D, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre (2008) "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2008, p. P10008.

Blume, Lawrence E (1993) "The statistical mechanics of strategic interaction," *Games and Economic Behavior*, Vol. 5, pp. 387–424.

Blume, Lawrence E, William A Brock, Steven N Durlauf, and Yannis M Ioannides (2011) "Identification of social interactions," in *Handbook of Social Economics*, Vol. 1: Elsevier, pp. 853–964.

Boucher, Vincent (2016) "Conformism and self-selection in social networks," *Journal of Public Economics*, Vol. 136, pp. 30–44.

Boucher, Vincent, Yann Bramoullé, Habiba Djebbari, and Bernard Fortin (2014) "Do peers affect student achievement? Evidence from Canada using group size variation," *Journal of Applied Econometrics*, Vol. 29, pp. 91–109.

Boucher, Vincent and Ismael Mourifié (2017) "My friend far, far away: a random field approach to exponential random graph models," *Econometrics Journal*, Vol. 20, pp. S14–S46.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2009) "Identification of peer effects through social networks," *Journal of Econometrics*, Vol. 150, pp. 41–55.

Breza, Emily, Arun G Chandrasekhar, Tyler H McCormick, and Mengjie Pan (2017) "Using aggregated relational data to feasibly identify network structure without network data," *NBER Working Paper No. 23491.*

Butts, C. (2009) "Using Potential Games to Parameterize ERG Models," *Working Paper.*

Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou (2009) "Peer effects and social networks in education," *The Review of Economic Studies*, Vol. 76, pp. 1239–1267.

Caner, Mehmet and Hao Helen Zhang (2014) "Adaptive elastic net for generalized methods of moments," *Journal of Business & Economic Statistics*, Vol. 32, pp. 30–47.

Chandrasekhar, Arun G and Matthew O Jackson (2014) "Tractable and consistent random graph models," *Working paper.*

Cohen-Cole, Ethan, Andrei Kirilenko, and Eleonora Patacchini (2014) "Trading networks and liquidity provision," *Journal of Financial Economics*, Vol. 113, pp. 235–251.

Cools, Angela, Raquel Fernández, and Eleonora Patacchini (2019) "Girls, Boys, and High Achievers," *NBER Working Paper No. 25763.*

Davezies, Laurent, Xavier d'Haultfoeuille, and Denis Fougère (2009) "Identification of peer effects using group size variation," *The Econometrics Journal*, Vol. 12, pp. 397–413.

De Paula, Aureo, Imran Rasul, and Pedro Souza (2018) "Recovering social networks from panel data: identification, simulations and an application," *CEPR Discussion Paper No. DP12792.*

Dzemski, Andreas (2019) "An empirical model of dyadic link formation in a network with unobserved heterogeneity," *Review of Economics and Statistics, forthcoming.*

Easley, David, Eleonora Patacchini, and Christopher Rojas (2019) "Multidimensional Diffusion Processes in Dynamic Online Networks," *Working paper.*

Evans, William N, Wallace E Oates, and Robert M Schwab (1992) "Measuring peer group effects: A study of teenage behavior," *Journal of Political Economy*, Vol. 100, pp. 966–991.

Fafchamps, Marcel and Flore Gubert (2007a) “The formation of risk sharing networks,” *Journal of Development Economics*, Vol. 83, pp. 326–350.

——— (2007b) “Risk sharing and network formation,” *American Economic Review*, Vol. 97, pp. 75–79.

Fortin, Bernard and Vincent Boucher (2016) “Some challenges in the empirics of the effects of networks,” in Yann Bramoullé, Andrea Galeotti, and Brian Rogers eds. *The Oxford Handbook of the Economics of Networks*: Oxford University Press, pp. 277–302.

Fortin, Bernard and Myra Yazbeck (2015) “Peer effects, fast food consumption and adolescent weight gain,” *Journal of Health Economics*, Vol. 42, pp. 125–138.

Frank, Ove and David Strauss (1986) “Markov graphs,” *Journal of the American Statistical Association*, Vol. 81, pp. 832–842.

Goldsmith-Pinkham, Paul and Guido W Imbens (2013) “Social networks and the identification of peer effects,” *Journal of Business & Economic Statistics*, Vol. 31, pp. 253–264.

Graham, Bryan S (2015) “Methods of identification in social networks,” *Annual Review of Economics*, Vol. 7, pp. 465–485.

——— (2016) “Homophily and transitivity in dynamic network formation,” *Working paper.*

——— (2017) “An econometric model of network formation with degree heterogeneity,” *Econometrica*, Vol. 85, pp. 1033–1063.

Greene, Derek, Donal Doyle, and Padraig Cunningham (2010) “Tracking the evolution of communities in dynamic social networks,” in *2010 international conference on advances in social networks analysis and mining*, pp. 176–183, IEEE.

Griffith, Amanda L and Kevin N Rask (2014) “Peer effects in higher education: A look at heterogeneous impacts,” *Economics of Education Review*, Vol. 39, pp. 65–77.

Han, Xiaoyi, Chih-Sheng Hsieh, and Stanley Ko (2019) “Spatial Modeling Approach for Dynamic Network Formation and Interactions,” *Journal of Business and Economic Statistics, forthcoming.*

Hoff, P.D., A. E. Raftery, and M. S. Handcock (2002) "Latent Space Approaches to Social Network Analysis," *Journal of the American Statistical Association*, Vol. 97.

Hoff, Peter D (2005) "Bilinear mixed-effects models for dyadic data," *Journal of the American Statistical Association*, Vol. 100, pp. 286–295.

Horrace, William C, Xiaodong Liu, and Eleonora Patacchini (2016) "Endogenous network production functions with selectivity," *Journal of Econometrics*, Vol. 190, pp. 222–232.

Hsieh, Chih-Sheng, Michael D. Konig, and Xiaodong Liu (2019a) "A Structural Model for the Coevolution of Networks and Behavior," *Working paper.*

Hsieh, Chih-Sheng and Lung Fei Lee (2016) "A social interactions model with endogenous friendship formation and selectivity," *Journal of Applied Econometrics*, Vol. 31, pp. 301–319.

Hsieh, Chih-Sheng, Lung-Fei Lee, and Vincent Boucher (2019b) "Specification and Estimation of Network Formation and Network Interaction Models with the Exponential Probability Distribution," *Available at SSRN 2967867.*

Hsieh, Chih-Sheng and Xu Lin (2017) "Gender and racial peer effects with endogenous network formation," *Regional Science and Urban Economics*, Vol. 67, pp. 135–147.

Hsieh, Chih-Sheng and Hans Van Kippersluis (2018) "Smoking initiation: Peers and personality," *Quantitative Economics*, Vol. 9, pp. 825–863.

Kang, Changhui (2007) "Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea," *Journal of Urban Economics*, Vol. 61, pp. 458–495.

Kelejian, Harry H and Ingmar R Prucha (1998) "A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances," *Journal of Real Estate Finance and Economics*, Vol. 17, pp. 99–121.

Koskinen, Johan H and Tom AB Snijders (2007) "Bayesian inference for dynamic social network data," *Journal of Statistical Planning and Inference*, Vol. 137, pp. 3930–3938.

Krivitsky, P.N., M. S. Handcock, M. Handcock, A. E. Raftery, and P.D. Hoff (2009) "Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models," *Social Networks*, Vol. 31, pp. 204–213.

Lazarsfeld, Paul F, Robert K Merton et al. (1954) "Friendship as a social process: A substantive and methodological analysis," *Freedom and control in modern society*, Vol. 18, pp. 18–66.

Lee, Lung-Fei (2004) "Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models," *Econometrica*, Vol. 72, pp. 1899–1925.

Lee, Lung-fei (2007) "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," *Journal of Econometrics*, Vol. 140, pp. 333–374.

Lee, Lung-fei, Xiaodong Liu, and Xu Lin (2010) "Specification and estimation of social interaction models with network structures," *The Econometrics Journal*, Vol. 13, pp. 145–176.

Lee, Lung-fei and Jihai Yu (2010) "Some recent developments in spatial panel data models," *Regional Science and Urban Economics*, Vol. 40, pp. 255–271.

——— (2012a) "QML estimation of spatial dynamic panel data models with time varying spatial weights matrices," *Spatial Economic Analysis*, Vol. 7, pp. 31–74.

——— (2012b) "Spatial panels: Random components versus fixed effects," *International Economic Review*, Vol. 53, pp. 1369–1412.

——— (2015) "Spatial panel data models," in B. Baltagi ed. *The Oxford Handbook of Panel Data*: Northamptonshire, UK: Oxford University Press.

Lin, Xu (2010) "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables," *Journal of Labor Economics*, Vol. 28, pp. 825–860.

Liu, Xiaodong and Lung-fei Lee (2010) "GMM estimation of social interaction models with centrality," *Journal of Econometrics*, Vol. 159, pp. 99–115.

Liu, Xiaodong, Eleonora Patacchini, and Yves Zenou (2014) "Endogenous peer effects: local aggregate or local average?" *Journal of Economic Behavior & Organization*, Vol. 103, pp. 39–59.

López-Pintado, Dunia (2008) "Diffusion in complex social networks," *Games and Economic Behavior*, Vol. 62, pp. 573–590.

Manresa, Elena (2013) "Estimating the structure of social interactions using panel data," *Unpublished Manuscript. CEMFI, Madrid.*

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

Mele, Angelo (2017) "A structural model of dense network formation," *Econometrica*, Vol. 85, pp. 825–850.

Mele, Angelo and Lingjiong Zhu (2017) "Approximate variational estimation for a model of network formation," *arXiv preprint arXiv:1702.00308*.

Monderer, Dov and Lloyd S Shapley (1996) "Potential games," *Games and Economic Behavior*, Vol. 14, pp. 124–143.

Moon, Hyungsik Roger and Ida Johnsson (2019) "Estimation Of peer effects in endogenous social networks: control function approach," *conditionally accepted at The Review of Economics and Statistics.*

Navarro, S (2008) "Control functions," in Steven N. Durlauf and Lawrence E. Blume eds. *The New Palgrave Dictionary of Economics Second Edition*: London: Palgrave Macmillan Press.

Newman, Mark EJ and Michelle Girvan (2004) "Finding and evaluating community structure in networks," *Physical Review E*, Vol. 69, p. 026113.

Patacchini, E., E. Rainone, and Y. Zenou (2017) "Heterogeneous peer effects in education," *Journal of Economic Behavior & Organization*, Vol. 134, pp. 190–227.

Patacchini, Eleonora and Yves Zenou (2009) "Juvenile delinquency and conformism," *Journal of Law, Economics, & Organization*, Vol. 28, pp. 1–31.

Pujol, Josep M, Vijay Erramilli, and Pablo Rodriguez (2009) "Divide and conquer: Partitioning online social networks," *arXiv preprint arXiv:0905.4918*.

Raftery, Adrian E, Michael A Newton, Jaya M Satagopan, and Pavel N Krivitsky (2007) "Estimating the integrated likelihood via posterior simulation using the harmonic mean identity," in J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West eds. *Bayesian Statistics*, Vol. 8: Oxford University Press, p. 1–45.

Raftery, Adrian E, Xiaoyue Niu, Peter D Hoff, and Ka Yee Yeung (2012) "Fast inference for the latent space network model using a case-control approximate likelihood," *Journal of Computational and Graphical Statistics*, Vol. 21, pp. 901–919.

Robins, Garry, Tom Snijders, Peng Wang, Mark Handcock, and Philippa Pattison (2007) "Recent developments in exponential random graph (p*) models for social networks," *Social Networks*, Vol. 29, pp. 192–215.

Sacerdote, Bruce (2001) "Peer effects with random assignment: Results for Dartmouth roommates," *Quarterly Journal of Economics*, Vol. 116, pp. 681–704.

Sarkar, Purnamrita and Andrew W Moore (2005) "Dynamic social network analysis using latent space models," *ACM SIGKDD Explorations Newsletter*, Vol. 7, pp. 31–40.

Sewell, Daniel K and Yuguo Chen (2015) "Latent space models for dynamic networks," *Journal of the American Statistical Association*, Vol. 110, pp. 1646–1657.

Sharma, Amit and Dan Cosley (2016) "Distinguishing between personal preferences and social influence in online activity feeds," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pp. 1091–1103, ACM.

Snijders, Tom AB (1996) “Stochastic actor-oriented models for network change,” *Journal of Mathematical Sociology*, Vol. 21, pp. 149–172.

——— (2001) “The statistical evaluation of social network dynamics,” *Sociological Methodology*, Vol. 31, pp. 361–395.

Snijders, Tom AB, Gerhard G Van de Bunt, and Christian EG Steglich (2010a) “Introduction to stochastic actor-based models for network dynamics,” *Social Networks*, Vol. 32, pp. 44–60.

Snijders, Tom AB, Johan Koskinen, and Michael Schweinberger (2010b) “Maximum likelihood estimation for social network dynamics,” *The Annals of Applied Statistics*, Vol. 4, p. 567.

Snijders, Tom AB and Mark Pickup (2017) “Stochastic Actor Oriented Models for Network Dynamics,” in Jennifer Nicoll Victor, Alexander H. Montgomery, and Mark Lubell eds. *The Oxford Handbook of Political Networks*: Oxford University Press, p. 1–45.

Snijders, Tom AB, Christian Steglich, and Michael Schweinberger (2007) “Modeling the coevolution of networks and behavior,” in Kees van Montfort, Johan Oud, and Albert Satorra eds. *Longitudinal models in the behavioral and related sciences*: Routledge, pp. 41–71.

Stinebrickner, Ralph and Todd R Stinebrickner (2006) “What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds,” *Journal of Public Economics*, Vol. 90, pp. 1435–1454.

Topa, Giorgio and Yves Zenou (2015) “Neighborhood and network effects,” in *Handbook of Regional and Urban Economics*, Vol. 5: Elsevier, pp. 561–624.

Wasserman, Stanley and Philippa Pattison (1996) “Logit models and logistic regressions for social networks: I. An introduction to Markov graphs andp,” *Psychometrika*, Vol. 61, pp. 401–425.

Wooldridge, J. M. (2015) "Control function methods in applied econometrics," *Journal of Human Resources*, Vol. 50, pp. 420–445.

Young, H Peyton (2009) "Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning," *American Economic Review*, Vol. 99, pp. 1899–1924.

Zimmerman, David J (2003) "Peer effects in academic outcomes: Evidence from a natural experiment," *Review of Economics and Statistics*, Vol. 85, pp. 9–23.

Index