



Munich Personal RePEc Archive

**Specification and Estimation of Network
Formation and Network Interaction
Models with the Exponential Probability
Distribution**

Hsieh, Chih-Sheng and Lee, Lung fei

Chinese University of Hong Kong, Ohio State University

8 April 2017

Online at <https://mpra.ub.uni-muenchen.de/60726/>

MPRA Paper No. 60726, posted 15 May 2017 04:34 UTC

Specification and Estimation of Network Formation and Network Interaction Models with the Exponential Probability Distribution*

Chih-Sheng Hsieh[†], Lung-Fei Lee[‡]

April 8, 2017

*A previous version of this paper was circulated under the title, “A structural modeling approach for network formation and social interactions with applications to students’ friendship choices and selectivity on activities.” The authors are grateful to Matthew Jackson as well as conference participants at 2013 AMES in Singapore and seminar participants at Academia Sinica, Chicago Booth, Chinese University of Hong Kong, Lingnan University, National Taiwan University, Ohio State University, Shanghai University of Finance and Economics, and Xiamen University for helpful comments.

[†]Department of Economics, The Chinese University of Hong Kong, Room 911 Esther Lee Building, CUHK, Shatin, HK

[‡]Department of Economics, The Ohio State University, Columbus, Ohio, 43210

Abstract

In this paper, we model network formation and network interactions under a unified framework. The key feature of our model is to allow individuals to respond to incentives stemming from interaction benefits on certain activities when they choose friends (network links), while capturing homophily in terms of unobserved characteristic variables in network formation and activities. There are two advantages of this modeling approach: first, one can evaluate whether incentives from certain interactions are important factors for friendship formation or not. Second, in addition to homophily effects in terms of unobserved characteristics, inclusion of incentive effects in the network formulation also corrects possible friendship selection bias on activity outcomes under network interactions. A theoretical foundation of this unified model is based on a complete information cooperative game. A tractable Bayesian MCMC approach is proposed for the estimation of the model. We apply the model to empirically study American high school students' friendship networks with the Add Health data. We consider two activity variables, GPA and smoking frequency, and find a significant incentive effect from GPA, but not from smoking, on friendship formation. These results suggest that the benefit of interactions in academic learning is an important factor for friendship formation, while the interaction benefit in smoking is not, even though homophily in smoking behavior is important for a smoker to link to other smokers. On the other hand, from the perspective of network interactions, both GPA and smoking frequency are subject to significant positive interaction (peer) effects.

JEL classification: C21, C25, I21, J13

Keywords: Social Networks, social interaction, selectivity, spatial autoregression, Bayesian estimation

1 Introduction

Economic research on social networks and interactions has grown over the past two decades. For many economic issues, the role of a social network as a channel to disseminate information or facilitate activities is revealed.¹ Accompanying a wide application of network concepts in economics, both academic researchers and practitioners are interested in understanding how networks are formed. This question is not only interesting in its own right, but is also important to understand how network structures may affect economic activities. In the context of social interactions, regardless of whether research subjects are workers, students, or delinquents, one likes to know how individuals choose their friends in order to take into account the advantage of peer effects on economic outcomes. As a friendship network might be formed in order to achieve favorable economic consequences, when studying the result of network (or peer) effects on economic activities, there is a need to correct for possible endogeneity biases due to friendship selection. Besides, the choice of friendships might amplify observed peer interaction effects due to unobserved factors behind both decisions of friendship and economic activities (Weinberg, 2007). With regard to the latter, Goldsmith-Pinkham and Imbens (2013), Hsieh and Lee (2016), and Johnsson and Moon (2016) study possibly important unobserved driving factors and use them to link network formation and network interactions on economic activities.

As mentioned, favorable outcomes due to peer interactions might be motivating factors for network formation from an economic perspective. In this paper, we propose an extended modeling approach for static networks and interactions among individuals, with a special interest on whether economic outcomes under interactions play roles in network formation. A static network refers to a cross-sectional case in which only one observation of a network

¹For example, job finding and labor force participation (Calvo-Armengol and Jackson, 2004; Calvo-Armengol and Jackson, 2007; Bayer et al., 2008); social learning and knowledge diffusion (Conley and Udry, 2001, 2010); risk sharing and insurance (Fafchamps and Gubert, 2007a,b); obesity transmission (Christakis and Fowler, 2007; Fowler and Christakis, 2008); peer effects on students' academic achievement (Calvo-Armengol et al., 2009); sport and club participation (Bramoullé et al., 2009; Liu et al., 2014); and juvenile delinquencies or criminal activities (Ballester et al., 2010; Bayer et al., 2009; Patacchini and Zenou, 2008, 2012)

is available.² We assume the possibility that economic outcomes give individuals utility values so that individuals may have incentives stemming from interaction benefits on certain activities when making friendship decisions. With an empirical survey data for friendship, it remains interesting to see from data which economic activities would provide significant incentives for forming friendships.

The advantage of modeling both network formation and network interactions under a unified framework is twofold: first, one can evaluate the importance of individuals' incentives stemming from economic activity interactions on friendship formation; second, the resulting model can correct possible friendship selection biases when studying network interactions on economic activities. We apply this model to study American high school students' friendship networks with the Add Health data. Two activity variables, student's GPA and how frequently a student smokes in a usual week, are of special interest and are considered in this paper. We find a significant incentive effect from GPA but not from smoking, which suggests that the interaction benefit in academic learning is a factor for building friendships, while the interaction benefit in smoking is not. Our results also reveal significant homophily effects from both observed and unobserved characteristics for network formation. Moreover, we find a significant homophily effect on smoking activity, which confirms that why smokers like to make friends with other smokers. Unobserved characteristics in network formation have significant influences in activity outcomes, i.e., outcomes with peer interactions are subject to selection biases in unobserved characteristics related to friendship formation. The outcome incentives and unobserved characteristics provide two sources of selection biases in activity outcomes. Though we found both GPA and smoking activity are subject to peer effects, but the estimated endogenous (peer) effects in activity outcomes from our model become smaller than those of a network interaction model where the network is assumed to be exogenously given, which indicates that our modeling approach can effectively detect and correct selection biases in outcome interactions due to endogenous friendship formation.

²We focus on a static setting because widely used social network data are cross-sectional ones without dynamics, e.g., Add Health data (Udry, 2003) and Indian rural village data (Banerjee et al., 2013). Few students' friendship network data which have panel waves can be found in the literature of stochastic actor-based dynamic network modeling proposed by Snijders (2001) and Snijders et al. (2010).

To model static networks, one approach is to assume pairwise independence between network links. For example, [Fafchamps and Gubert \(2007a,b\)](#) and [Comola \(2007\)](#) apply the pairwise independence assumption, which allows them to focus on individual and dyad-specific variables to explain network formation. A further extension is available in the latent position model ([Hoff et al., 2002](#); [Handcock et al., 2007](#)) and the model with degree heterogeneity ([Graham, 2014](#)), where individuals are assumed having unobserved positions (or fixed effects) in the network which reflect heterogeneity of their social or economic statuses. These unobserved positions allow researchers to control homophily effect in terms of unobserved individual characteristics. Under pairwise independence, conditional on unobserved individual effects, the likelihood of the whole network is the product of likelihoods from all pairwise links.³ However, as noted by [Bramoullé and Fortin \(2009\)](#), the assumption of pairwise independence is strong because it requires that the latent utility for each pairwise link be separable. This means that an individual’s utility derived from a network is equal to the sum of utilities from each of her links and each link utility is not affected by any other links in the network. However, while one is easier to provide economic justification in terms of game theoretical consideration in network formation, it ignores finding from the statistical literature that relevant network characteristics can be important in the estimation of network formation probabilities.

A statistical investigation on static networks, without imposing the pairwise independence assumption, is to treat an observed network as a polychotomous choice with $2^{m(m-1)}$ alternatives made jointly by individuals, where m is the size of the network. The Exponential Random Graph (ERG) model proposed by [Frank and Strauss \(1986\)](#), or more generally, the p^* model by [Wasserman and Pattison \(1996\)](#), are the models of this type. In either an ERG or a p^* model, several selected network statistics, such as the number of reciprocal links, the number of k -stars, $k \geq 2$, are specified in an exponential probability distribution to capture how likely those network structures are to appear in a network. The parameters of those network statistics in ERG and p^* models do not provide casual interpretations. In terms of estimation, the likelihood function of an ERG model involves an intractable

³The unconditional likelihood will involve integration of the conditional likelihood with respect to the unobservables’ distribution.

normalizing term in the denominator which requires evaluation of network statistics on all possible network realizations. To handle the intractable normalizing term during estimation, researchers need to use the classical M estimation incorporated with simulation (Geyer and Thompson, 1992; Snijders, 2002) or the Bayesian approach with auxiliary Markov chain Monte Carlo (MCMC) (Liang, 2010; Murray et al., 2006). These estimation methods are generally more computational intensive compared to pairwise independent network models. However, these methods are still tractable and manageable.

In this paper, we go beyond the pairwise independence specification and consider the exponential probability distribution to model network data. However, different from standard ERGMs, we motivate the model specification from economic reasoning. Meanwhile, we control unobserved individual heterogeneity through latent variables as we did in Hsieh and Lee (2016). Our proposed network formation model can handle both unobserved individual heterogeneity and endogenous economic activities as incentives in link decisions. Few papers in the literature are relevant to us (Christakis et al., 2010; Steglich et al., 2010; Badev, 2013; Sheng, 2014; Chandrasekhar and Jackson, 2014; Mele, 2016). In Christakis et al. (2010), Sheng (2014), Chandrasekhar and Jackson (2014), and Mele (2016), they model link formation but without incorporating network interaction effects on activity outcomes; and they do not have economic outcomes as incentive factors in link decisions. In Steglich et al. (2010) and Badev (2013), network links and activity outcomes are modelled jointly. However, the difference between their approaches and ours is that they define peer effect as the effect of peers' activities on individual's utility, while we define it as the effect of peers' activities on individual's activity following the conventional social interaction literature with both endogenous and contextual interactions. Besides, our approach additionally capture individual unobserved heterogeneity during network formation while they do not.

This paper is organized as follows. Section 2 presents a unified modeling approach for both network formation and network interactions on economic activities. A Bayesian estimation method for the proposed model is discussed in Section 3. Section 4 provides an application of the model to high school students' friendship networks and activities with the Add Health data. Section 5 concludes the paper. We leave a simulation study for showing model identification in the appendix. Some more technical details for estimation are in a

supplementary file.

2 Models of network formation and network interactions on economic activities

Our research subjects are individuals in a closed group, such as students in a school-grade or workers in a company. Let W_g be a $m_g \times m_g$ matrix (spatial weights matrix; adjacency matrix; sociomatrix) representing a friendship network of m_g individuals (size) in group g , where $g = 1, \dots, G$, with G being the total number of groups.⁴ The $(i, j)^{\text{th}}$ entry of W_g , denoted as $w_{ij,g}$, is a dichotomous indicator which equals one if individual i sends a link to individual j and zero, if not. The notation $w_{i.,g}$ stands for the i^{th} row of W_g and $W_{-i.,g}$ stands for W_g excluding $w_{i.,g}$. The links are all directed without imposing reciprocity.⁵ Therefore, it is possible that individual i sends a link to j but j does not send a link to i , i.e., W_g is not symmetric. Diagonal elements, $w_{ii,g}$, $i = 1, \dots, m_g$, are zeros, à priori. Let $x_{i,g}$ be a k -dimensional row vector containing individual i 's exogenous characteristics and the $m_g \times k$ dimensional matrix X_g with $x_{i,g}$ as its i^{th} row be a collection of such vectors in group g . For economic activities, we consider two types of variables – continuous and Tobit-type (continuous but left censored at zero).⁶ Let $y_{i,cg}$ and $y_{i,tg}$ denote, respectively, individual i 's continuous and Tobit-type activity variables in group g ; $Y_{cg} = (y_{1,cg}, y_{2,cg}, \dots, y_{m_g,cg})'$ and $Y_{tg} = (y_{1,tg}, y_{2,tg}, \dots, y_{m_g,tg})'$ be, respectively, the m_g -dimensional column vectors for all members' continuous and Tobit-type activity variables in group g .

We assume that individuals make their decisions on friendship links and economic activities in two stages with complete information. In the first stage, individuals choose their friends and in the second stage, they choose economic activities with network interactions.

⁴We do not rule out the case that there is only a single group, i.e., $G = 1$.

⁵We do not impose reciprocity of friendship links in our model because in our empirical data set – Add Health, friendships are nominated by individuals privately without mutual consent.

⁶We do not consider the case of binary variables in this paper because it might involve the issue of multiple equilibria if network interactions are based on observed binary variables (Krauth, 2006; Soetevent and Kooreman, 2007). The modeling of binary variables is of interest and challenging. We will leave it for future research.

This two-stage process is characterized as a two-stage static game with complete information. Individuals adopt strategies on choosing friends and economic activities and obtain utilities as payoffs of the game. We further assume this two-stage game is cooperative. Cooperative behaviors among individuals in a closed group are argued by economists and biologists with theories from iterated prisoner’s dilemma and cooperative strategies (i.e., tit-for-tat) (See e.g., [Peck, 1993](#); [Hruschka and Henrich, 2006](#); [Majolo et al., 2006](#); [Ule, 2008](#); [Fu et al., 2008](#); and [Fosco and Mengel, 2011](#)) and social preference (See, e.g., [Fehr and Schmidt, 1999](#)). Also, [Jackson \(2010\)](#) indicates that studying allocation rules behind the cooperative network formation game is rational in many economic applications, such as students’ friendship network in schools, where favors can be exchanged between students.⁷

We further assume there is complete information between the two stages. The equilibrium of this two-stage game satisfies the principle of sequential rationality, i.e., a player’s strategy should specify optimal actions at every point in the game tree ([Mas-Colell et al., 1995](#)). Hence, one can solve the equilibrium of this game by backward induction. First, determine equilibrium activities in the second stage and calculate corresponding optimum utilities of economic activities for each possible network pattern. Then by incorporating the optimum utilities from economic activities under network interactions into the utilities of links in the first stage, solve for the equilibrium network.⁸

⁷By assuming the game is cooperative, we can obtain one unique game equilibrium from maximizing the transferable utility defined by the aggregate of individual utilities and that of a coordinator, which directly links to formulation of our empirical model in terms of the exponential probability distribution. On the other hand, when assuming the game is non-cooperative, one needs to handle multiplicity of game equilibria in the empirical study by either specifying a partially identified model ([Sheng, 2014](#)) or changing the model from the static one to an evolutionary one with the equilibria characterized by the potential function ([Christakis et al., 2010](#); [Badev, 2013](#); [Mele, 2016](#)). In both cases, specification of individual utility in the empirical study will be restricted, owing to the curse of dimensionality and the requirement of symmetry for the potential function.

⁸The reason why we assume a complete information between the two stages instead of specifying the expectation of activity outcomes in the network formation process is because we want the disturbance term from the activity outcome to have a direct impact on network formation, which forms a mechanism of selection. If using the expectations of activity outcomes, the disturbances will be irrelevant in the network formation process, and we lose the channel to correct the selection bias on the estimated peer effect in the

2.1 Network interactions on activities

Following the equilibrium solving rule of this two-stage game, we first provide the details of the network interaction process in the second stage. We adapt the utility specification from [Ballester et al. \(2006\)](#) and [Calvó-Armengol et al. \(2009\)](#) which considers that, given the network W_g , an individual faces a quadratic utility from choosing activity $y_{i,cg}$,

$$u_{i,cg}(y_{i,cg}, Y_{-i,cg}, W_g) = \mu_{i,g}y_{i,cg} - \frac{1}{2}y_{i,cg}^2 + \lambda_c y_{i,cg} \sum_{j=1}^{m_g} w_{ij,g} y_{j,cg}, \quad (1)$$

for $i = 1, \dots, m_g$, where $\mu_{i,g}$ captures individual exogenous heterogeneity. The first and second terms of Eq. (1) show that the utility is concave in individuals' activity. The third term reflects a complementary (or competitive) effect from peers' activities if $\lambda_c \geq 0$ ($\lambda_c \leq 0$). The sub-game cooperative equilibrium can be found from maximizing the transferable utility, defined as the aggregate of individual utilities,

$$U(Y_{cg}, W_g) = \sum_{i=1}^{m_g} u_{i,cg}(y_{i,cg}, Y_{-i,cg}, W_g) = \sum_{i=1}^{m_g} \mu_{i,g}y_{i,cg} - \frac{1}{2} \sum_{i=1}^{m_g} y_{i,cg}^2 + \lambda_c \sum_{i=1}^{m_g} \left(y_{i,cg} \sum_{j=1}^{m_g} w_{ij,g} y_{j,cg} \right). \quad (2)$$

From the first order condition,

$$\frac{\partial U(Y_{cg}, W_g)}{\partial y_{i,cg}} = \mu_{i,g} - y_{i,cg} + \lambda_c \sum_{j=1}^{m_g} (w_{ij,g} + w_{ji,g}) y_{j,cg} = 0, \quad (3)$$

we have the corresponding individual activity characterized by

$$y_{i,cg}^*(Y_{-i,cg}, W_g) = \lambda_c \sum_{j=1}^{m_g} (w_{ij,g} + w_{ji,g}) y_{j,cg} + \mu_{i,g}. \quad (4)$$

For the individual exogenous heterogeneity, we model it as $\mu_{i,g} = x_{i,g}\beta_{1c} + \sum_{j=1}^{m_g} w_{ij,g}x_{j,g}\beta_{2c} + \alpha_{cg} + \epsilon_{i,cg}$, where α_{cg} represents an unobserved group effect of group g ; and $\epsilon_{i,cg}$ is a shock activity outcome equation.

 Though we can still estimate the incentive effect from the expectation of activity outcome in the network formation, the formed network would be independent with the disturbances in the outcome equations. Our model is relatively more complicated in that a formed network may correlate with disturbances (shocks) of the outcome equation. In some cases, the model can also be interpreted as joint decisions in a single stage instead of two stages. However, the two-stage decision setting would be more flexible in general. Descriptions will be subsequently presented.

for i , assumed known to all players but unknown to econometricians. By the theorem of [Ballester et al. \(2006\)](#), as long as $|\lambda_c|$ is less than the inverse of the largest eigenvalue of $(W_g + W'_g)$, the unique interior cooperative equilibrium activity vector will take the form as⁹

$$Y_{cg}^*(W_g) = (I_{m_g} - \lambda_c(W_g + W'_g))^{-1} (X_g\beta_{1c} + W_gX_g\beta_{2c} + l_g\alpha_{cg} + \epsilon_{cg}), \quad (5)$$

for $g = 1, \dots, G$, where I_{m_g} is an $m_g \times m_g$ identity matrix; l_g is the m_g -dimensional vector of ones; and $\epsilon_{cg} = (\epsilon_{1,cg}, \epsilon_{2,cg}, \dots, \epsilon_{m_g,cg})'$.

The activity vector of Eq. (5) matches the reduced form of the spatial autoregressive (SAR) model ([Bramoullé et al., 2009](#); [Lee et al., 2010](#); [Lin, 2010](#)) for social interactions except that in our current case, the observed W_g will be endogenous and its elements correlate with the disturbance ϵ_{cg} .¹⁰ The coefficient λ_c in Eq. (5) represents the endogenous (peer) effect, which is the key parameter of interest to us. The vector of coefficients, $\beta_c = (\beta'_{1c}, \beta'_{2c})'$, will capture effects from individuals' own and friends' exogenous characteristics, i.e., own and contextual effects, on Y_{cg} . Specifying group effects in Eq. (5) will capture specific group characteristics in addition to contextual variables and will be helpful to handle the identification problem caused by correlated effects. [Moffitt et al. \(2001\)](#) argues that correlated unobservables (to econometricians) in a group may contribute to correlations of Y_{cg} across elements and cause an identification problem by confounding the endogenous effect. Here, group effects refer to effects from (unobserved) environmental factors shared by all members in the same group.¹¹ Note that with the strategy in Eq. (5), the group can obtain the

⁹It is interesting to note that in the non-cooperative environment considered in [Ballester et al. \(2006\)](#), the unique interior Nash equilibrium activity vector will be given by

$$Y_{cg}^*(W_g) = (I_{m_g} - \lambda_c W_g)^{-1} (X_g\beta_{1c} + W_gX_g\beta_{2c} + l_g\alpha_{cg} + \epsilon_{cg}).$$

Compared to Eq. (5), individuals only receive the endogenous peer effect from outward links (friendships nominated) but not from inward links (friendships received) in a non-cooperative environment. In a cooperative environment, externality is taken into account.

¹⁰The original SAR model is specified as

$$Y_{cg} = \lambda_c(W_g + W'_g)Y_{cg} + X_g\beta_{1c} + W_gX_g\beta_{2c} + l_g\alpha_{cg} + \epsilon_{cg}.$$

¹¹For a single group or a network, the group effect would be absorbed by the intercept term. Group effects

corresponding aggregated utility $U(Y_{cg}^*(W_g), W_g) = \frac{1}{2}Y_{cg}^{*'}(W_g)Y_{cg}^*(W_g)$ for a given network structure W_g . In certain cases, activity variables might be continuous, but nonnegative, i.e., a Tobit-type variable which is left-censored at the value zero, such as smoking frequency considered in the empirical study of this paper. A Tobit-type activity outcome would be the case of activity outcomes subject to non-negative constraints. Since network interactions on such Tobit-type variables are similarly captured by a SAR model with non-negative constraints, we delegate details of network interactions on Tobt-type variables to Appendix A.

One concern of using a conventional SAR model in studying network interactions is the possible endogeneity of the weight matrix, W_g . If W_g is endogenous and it correlates with the disturbance term ϵ_g of the model, estimating a conventional SAR model by treating W_g as exogenous will result in biases on the estimated endogenous effect, as well as other effects. A standard instrumental variable (IV) approach would suggest finding instruments for the endogenous weight matrix. However, without utilizing information provided by structures of W_g or its formation process, an effective instrument might be difficult to find. Instead of pursuing an IV approach, we propose a structural model in this paper which unifies the SAR with a network formation model. In the next subsection we will present the first stage of the game and introduce individual utility functions for network formation. The equilibrium condition of the network formation game motivates specified network statistics in ERGM with economic reasoning. The key to combining the network formation and interaction processes on economic activities, as mentioned before, is to allow individuals to consider potential benefits, which they can earn from interacted economic activities, when choosing friends.

are of interest only when there are many groups.

2.2 Network formation with incentives from interaction benefits

We consider that each individual i , $i = 1, \dots, m_g$, obtains her utility from network links in W_g as:

$$v_{i,g}(W_g) = \underbrace{\sum_{j=1}^{m_g} w_{ij,g} \psi_{ij,g}}_{\text{Exogenous Effects}} + \underbrace{\varpi_{i,g}(w_{i.,g}, W_{-i.,g}) \eta}_{\text{Network Structure Effects}} + \underbrace{\sum_{d=1}^{\bar{d}} \frac{\delta_d}{2} y_{i,dg}^{*2}}_{\text{Incentive Effects}}. \quad (6)$$

In Eq. (6), the exogenous effects capture influences from individual-specific and dyad-specific exogenous characteristics on the link utility. The function, $\psi_{ij,g}$, has the specification,

$$\psi_{ij,g} = c_{i,g} \gamma_1 + c_{j,g} \gamma_2 + c_{ij,g} \gamma_3 + \sum_{\ell=1}^{\bar{\ell}} \gamma_{4\ell} |z_{i\ell,g} - z_{j\ell,g}|. \quad (7)$$

The variables, $c_{i,g}$ and $c_{j,g}$, in Eq. (7) are \bar{s} -dimensional row vectors of individual-specific characteristics and the variable, $c_{ij,g}$, is a \bar{q} -dimensional row vector of dyad-specific characteristics, such as the same age, sex, or race shared by each pair of individuals (i, j) in group g . The individual and dyadic characteristics $C_g = \{(c_{i,g}, c_{j,g}, c_{ij,g}) : i = 1, \dots, m_g, j = 1, \dots, m_g, i \neq j\}$ provide the controls of homophily from observed (for econometricians) characteristics in friendship formation process (see e.g., [Fafchamps and Gubert \(2007a,b\)](#) on the study of risk-sharing network formation). To further capture unobserved (for econometricians) individual heterogeneity during the network formation process, we follow [Hsieh and Lee \(2016\)](#) to introduce multi-dimensional individual latent variables $z_{i,g} = (z_{i1,g}, \dots, z_{i\bar{\ell},g})'$ and use $|z_{i\ell,g} - z_{j\ell,g}|$, $\ell = 1, \dots, \bar{\ell}$ in Eq. (7) to capture the homophily of unobserved characteristics. We expect the coefficients $\gamma'_{4\ell}$ s to be negative to reflect the fact that the larger the differences between individual unobserved characteristics, the less likely that two individuals become friends. We assume that the individual latent variable $z_{i,g}$ are components in $\epsilon_{i,cg}$ (as well as $\epsilon_{i,tg}$) in the activity outcome. To be explicit, we assume $\epsilon_{cg} = Z_g \rho_{1c} + W_g Z_g \rho_{2c} + \xi_{cg}$, where $Z_g = (z_{1,g}, \dots, z_{m_g,g})'$. Correspondingly, the activity outcome of Eq. (5) should be modified into

$$Y_{cg}^*(W_g) = (I_{m_g} - \lambda_c(W_g + W_g'))^{-1} (X_g \beta_{1c} + W_g X_g \beta_{2c} + Z_g \rho_{1c} + W_g Z_g \rho_{2c} + l_g \alpha_{cg} + \xi_{cg}), \quad (8)$$

where we assume $\xi_{cg} \sim \mathcal{N}_{m_g}(0, \sigma_{\xi_{cg}}^2 I_{m_g})$ for estimation. In Eq (8), Z_g and $W_g Z_g$ are regarded as controls of individual and contextual unobserved correlated effects. [Fruehwirth \(2014\)](#) argues that any omitted contextual effect will be picked up by the endogenous peer effect. This issue is similar to [Manski \(1993\)](#)'s reflection problem in social interactions.

Furthermore, it is possible to allow $|y_{i,dg}^*(W_g) - y_{j,dg}^*(W_g)|$ to appear in $\psi_{ij,g}$ for capturing the homophily effect from activity outcomes. Such an extension provides a channel for activity variables to feedback network formation. For example, individuals may take into account the difference of smoking intensities in forming friendships – heavy smokers may like to make friends with other heavy smokers and a non-smoker may not like to make friends with heavy smokers.

The network structure effects in Eq. (6) capture influences from some patterns of link dependence on the link utility, where $\varpi_{i,g}(w_{i.,g}, W_{-i.,g})$ represents a \bar{h} -dimensional row vector of summary statistics constructed from W_g which are relevant to individual i 's utility and η is a corresponding vector of coefficients. By considering network structure effects, our network model differs substantially from the pairwise network link case ([Bramoullé and Fortin, 2009](#)) and connects to ERG models in the statistical literature. The empirical specification of network structure effects used in this paper will be discussed later in Section 4.1.

The incentive effects in Eq. (6) are innovative in this paper, which represent benefits from network interactions, i.e., utilities obtained from the activity outcomes with interaction. For example, students may want to make friends with someone who is doing well in school in order to learn from him or her to improve their own education performance. As there is assumed complete information across two stages, individuals can exactly foresee these benefits during the network formation process. The incentive effects are functions of optimized individual's activity responses, which are intuitive. The coefficients δ_d 's capture how important are benefits of economic outcomes as factors in network formation decision. Moreover, the aggregate of individual incentive effects equals to the aggregated utility generated from economic activities under network interactions, which is similar to a dynamic inter-temporal utility optimization setting with δ_d 's as discounting factors of the second stage utility.

There may be several (\bar{d}) economic activities which provide incentives for forming friendships. For simplicity, these incentive effects are assumed to be separable as utility compo-

nents. As noted by [Ballester et al. \(2006\)](#), utilities from economic outcomes with network interactions always increase with the number of links in the network if interactions provide complementary effects on economic activities. Since the utility from network links contains incentive effects, individuals might choose to add as many links as possible if there were no cost on link formation. This is also related to the problem of network degeneracy as discussed in [Snijders et al. \(2006\)](#). To mitigate such a strong incentive to form links, we rely on nontrivial negative effects from some exogenous or network structural effects to represent possible costs of forming friendship links.¹²

Similar to network interactions in the second stage of the game, we define a transferable utility of this cooperative network formation game from the aggregate of individual utilities, which is

$$\begin{aligned} T(W_g) &= V(W_g) + \tau_{W_g} = \sum_{i=1}^{m_g} v_{i,g}(W_g) + \tau_{W_g} \\ &= \sum_{i=1}^{m_g} \sum_{j=1}^{m_g} w_{ij,g} \psi_{ij,g} + \sum_{i=1}^{m_g} \varpi_{i,g}(w_{i.,g}, W_{-i.,g}) \eta + \sum_{d=1}^{\bar{d}} \frac{\delta_d}{2} Y_{dg}^*(W_g)' Y_{dg}^*(W_g) + \tau_{W_g}, \end{aligned} \quad (9)$$

where $Y_{dg}^*(W_g)$ represents a d^{th} vector of activity outcomes, such as $Y_{cg}^*(W_g)$ discussed in Eq. (8) or $Y_{tg}^*(W_g)$ discussed in Appendix A. τ_{W_g} stands for other value that the coordinator considers in addition to the aggregated utilities of individuals, which is uncorrelated with the network or the activity outcome. Thus, W_g is the formed network if and only if $T(W_g)$ reaches the maximum value over all possible network patterns, i.e., $T(W_g) = \max_{\widetilde{W}_g \in \Omega_g} (V(\widetilde{W}_g) + \tau_{\widetilde{W}_g})$, where Ω_g is the set of possible network patterns for the individuals in a group g . By assuming that $\tau_{\widetilde{W}_g}$ is from an i.i.d. type I extreme value random variable indexed by \widetilde{W}_g in Ω_g , we have the exponential probability specification for W_g :

$$P(W_g) = \frac{\exp(V(W_g))}{\sum_{\widetilde{W}_g \in \Omega_g} \exp(V(\widetilde{W}_g))}, \quad (10)$$

which relates our network formation model to the ERG framework. A typical ERG model ([Frank and Strauss, 1986](#); [Wasserman and Pattison, 1996](#); [Snijders, 2002](#)) assumes a probability

¹²[Bhamidi et al. \(2011\)](#), [Chatterjee et al. \(2013\)](#), and [Mele \(2016\)](#) also argue that the most basic ERGs are statistically equivalent to an Erdős-Rényi random graph in the limit of large size unless the model contains at least one non-trivial negative network structural effect.

function of W_g as:

$$P(W_g) = \frac{\exp(Q(W_g))}{\sum_{\widetilde{W}_g \in \Omega_g} \exp(Q(\widetilde{W}_g))},$$

where a specification of the function, $Q(W_g)$, can accommodate various network statistics of interest. But in order to apply this type of a model in empirical studies, researchers tend to specify link dependencies sparingly so that the resulting probability is simple and practical (Jackson, 2010). Network statistics such as the number of k -stars, $k \geq 2$, and the number of triangles are used in $Q(W_g)$ to measure how likely those network structures appear in observed networks. However, the coefficients of those network statistics do not represent causal relationships. In contrast to typical ERG models, our specification of $V(W_g)$ in Eq. (10) is related to an economically motivated utility function for network formation. In fact, this model can also be modeled as a one shot game with a joint decision for W_g and Y_{dg} 's by maximizing the aggregate utility

$$T(W_g, Y_{1g}, \dots, Y_{\bar{d},g}) = \sum_{i=1}^{m_g} \sum_{j=1}^{m_g} w_{ij,g} \psi_{ij,g} + \sum_{i=1}^{m_g} \varpi_{i,g}(w_{i.,g}, W_{-i,g}) \eta + \sum_{d=1}^{\bar{d}} \delta_d U(Y_{dg}, W_g) + \tau_{W_g} \quad (11)$$

for the case that all δ_d 's are strictly positive. This follows because

$$\begin{aligned} & \max_{W_g, Y_{1g}, \dots, Y_{\bar{d},g}} T(W_g, Y_{1g}, \dots, Y_{\bar{d},g}) \\ &= \max_{W_g} \left\{ \sum_{i=1}^{m_g} \sum_{j=1}^{m_g} w_{ij,g} \psi_{ij,g} + \sum_{i=1}^{m_g} \varpi_{i,g}(w_{i.,g}, W_{-i,g}) \eta + \sum_{d=1}^{\bar{d}} \delta_d \max_{Y_{dg}} U(Y_{dg}, W_g) + \tau_{W_g} \right\} \\ &= \max_{W_g} T(W_g). \end{aligned} \quad (12)$$

However, jointly modeling the endogenous network formation and activity variables under a unified framework as a two stage game has a flexibility. In the event that individuals do not respond to incentives from economic activities when choosing their friends, the coefficients δ_d 's can be zero. In that case, a one shot game can provide only the solution for network formation, but not for economic activities. Optimum economic activities can be determined only if a second stage game is formulated.

Even though the i.i.d. assumption on $\tau_{\widetilde{W}_g}$ gives the conditional logit form in Eq. (10), which is relatively simple for estimation, it inevitably exhibits the property of independence

of irrelevant alternatives (IIA). One may change the distribution of $\tau_{\widetilde{W}_g}$ to the generalized extreme value (GEV) distribution

$$F(\tau_{W_{g(1)}}, \tau_{W_{g(2)}}, \dots, \tau_{W_{g(J)}}) = \exp[-\Gamma(e^{-\tau_{W_{g(1)}}}, e^{-\tau_{W_{g(2)}}}, \dots, e^{-\tau_{W_{g(J)}}})]$$

so that it permits correlations among the total of J network alternatives, where J is the number of networks in Ω_g . If the j^{th} network alternative $W_{g(j)}$ is chosen with the highest value of $T(W)$, we obtain the GEV model

$$P(W_{g(j)}) = \frac{\exp(V(W_{g(j)}))\Gamma_j[e^{V(W_{g(1)})}, e^{V(W_{g(2)})}, \dots, e^{V(W_{g(k)})}]}{\Gamma[e^{V(W_{g(1)})}, e^{V(W_{g(2)})}, \dots, e^{V(W_{g(k)})}]},$$

where Γ_j denotes $\partial\Gamma/\partial \exp(V(W_{g(j)}))$. Within the GEV model family, the function Γ can be specified to relax the IIA property in the resulting probability function, for example, nested logit is a widely used one if the data exhibits a nested structure. However, it is not obvious in our empirical context that friendship network alternatives may exhibit a nested structure. Therefore, we maintain the i.i.d. assumption on τ_{W_g} for our empirical applications in this paper.

The model specification alleviates the problem of friendship selection bias on the economic activities under interaction in two ways. First, the correlation of W_g and ξ_{dg} is explicitly modeled. Secondly, the disturbance terms ξ_{dg} 's appear in both the network formation and network interaction process. Hence, they capture unobserved factors which contribute to both friendship and economic activity decisions. In the following section, we will discuss how to estimate this model.

3 Model estimation

3.1 Likelihood function of the model

To give a clear but succinct illustration on how the likelihood function of our model is constructed, we consider a model with an incentive effect from a continuous activity variable as an example. For readers who are interested in the Tobit activity case or models with multiple activity variables,¹³ we provide additional discussion on their likelihood functions

¹³In Section 2.2 we consider the optimization of activities is separable (in utility). Therefore, when we consider a single activity outcome in the model, but in fact there are other omitted activity outcomes, then

in Appendix B.

With an incentive effect from the continuous variable Y_{cg} , the joint probability of the activity variable Y_{cg} and the network W_g is¹⁴

$$\begin{aligned}
P(W_g, Y_{cg} | \theta_{cg}, \alpha_{cg}, Z_g) &= P(Y_{cg} | W_g, \theta_{cg}, \alpha_{cg}, Z_g) \cdot P(W_g | \theta_{cg}, \alpha_{cg}, Z_g) \\
&= |S_{cg}(W_g)| \cdot f(\xi_{cg} | W_g, \theta_{cg}, \alpha_{cg}, Z_g) \cdot P(W_g | \theta_{cg}, \alpha_{cg}, Z_g) \\
&= |S_{cg}(W_g)| \cdot f(\xi_{cg}, W_g | \theta_{cg}, \alpha_{cg}, Z_g) \\
&= |S_{cg}(W_g)| \cdot f(\xi_{cg} | \theta_{cg}, \alpha_{cg}, Z_g) \cdot P(W_g | \xi_{cg}, \theta_{cg}, \alpha_{cg}, Z_g) \\
&= |S_{cg}(W_g)| \cdot f(\xi_{cg} | \theta_{cg}, \alpha_{cg}, Z_g) \cdot \frac{\exp(V(W_g, \xi_{cg}, \theta_{cg}, \alpha_{cg}, Z_g))}{\sum_{\widetilde{W}_g \in \Omega_g} \exp(V(\widetilde{W}_g, \xi_{cg}, \theta_{cg}, \alpha_{cg}, Z_g))},
\end{aligned} \tag{13}$$

where $S_{cg}(W_g) = I_{m_g} - \lambda_c(W_g + W'_g)$, and

$$f(\xi_{cg} | \theta_{cg}, \alpha_{cg}) = (2\pi)^{-\frac{m_g}{2}} \left(\sigma_{\xi_{cg}}^2 \right)^{-\frac{m_g}{2}} \exp \left(-\frac{1}{2\sigma_{\xi_{cg}}^2} \xi_{cg}' \xi_{cg} \right),$$

with $\xi_{cg} = S_{cg}(W_g)Y_{cg} - X_g\beta_{1c} - W_gX_g\beta_{2c} - Z_g\rho_{1c} - W_gZ_g\rho_{2c} - l_g\alpha_{cg}$ and the parameter vector $\theta_{cg} = (\gamma', \eta', \delta_c, \lambda_c, \beta'_c, \rho'_c, \sigma_{\xi_{cg}}^2)$.

The main issue we will encounter during estimation is to calculate the likelihood function of the exponential distribution for the network. When the network size is moderate, its calculation is almost impossible since it requires evaluating all network patterns in Ω_g for the denominator of the exponential distribution function. For example, even in a network with just 5 individuals, it needs to evaluate $2^{4 \times 5} = 2^{20}$ terms of possible network

the resulted optimal utility components from other omitted activity outcomes, given W , become a part in τ_W – disturbance in network utility, which is independent of the included optimum activity outcome. However, if disturbances of activity outcomes are correlated, then we need to include both to avoid correlation of the optimum utility components from activity outcomes.

¹⁴In Eq. (13), we use the observed Y_{cg} to evaluate the function $V(W_g, \xi_{cg}, \theta_{cg}, \alpha_{cg}, Z_g)$ in the numerator. To evaluate $V(\widetilde{W}_g, \xi_{cg}, \theta_{cg}, \alpha_{cg}, Z_g)$ in the denominator for any \widetilde{W}_g , we follow the reduced form outcome equation in Eq (8) to get

$$Y_{cg}^*(\widetilde{W}_g) = \left(I_{m_g} - \lambda_c(\widetilde{W}_g + \widetilde{W}'_g) \right)^{-1} \left(X_g\beta_{1c} + \widetilde{W}_gX_g\beta_{2c} + Z_g\rho_{1c} + \widetilde{W}_gZ_g\rho_{2c} + l_g\alpha_{cg} + \xi_{cg} \right).$$

alternatives (patterns) for the denominator. Hence, any estimation method involving direct likelihood evaluation would be infeasible. This problem applies to all ERGMs for networks and can be traced back to the spatial analysis in [Besag \(1974\)](#).

To deal with this problem, several estimation methods have been proposed. The first is the maximum pseudo-likelihood approach (MPL). This approach was first mentioned in [Besag \(1974\)](#) and later applied to the network study in [Strauss and Ikeda \(1990\)](#). A pseudo-likelihood simply uses the product of conditional probabilities for estimation. The estimates from the MPL would not be the MLE. One may use the estimates from the MPL as initial values for other estimation approaches. Another approach is the Monte Carlo maximum likelihood (MCML) estimation approach which simulates auxiliary networks for approximating the denominator of the exponential distribution function with its simulated counterpart ([Geyer and Thompson, 1992](#)). One shortcoming of the MCML approach is that the choice of initial values during the optimization algorithm plays a critical role. They have to be close enough to the true parameter values, otherwise, the convergence of the algorithm might not be attained ([Bartz et al., 2009](#); [Caimo and Friel, 2011](#)). The Robbins-Monro approach used in [Snijders \(2002\)](#) to simulate auxiliary networks for constructing simulated moments usually accepts a wide range of initial values which may lead to a convergent algorithm. In this paper we turn to the Bayesian estimation with an effective MCMC technique (discussed in [Section 3.3](#)) developed to handle an intractable normalizing term in the posterior density function.

Regarding identification of parameters in our models, one may focus on the coefficients of incentive effects, δ_d 's, in the network formation model, and the endogenous effects, λ_d 's, in the network interaction models. The remaining parameters in the network formation model will be identified as long as corresponding regressors are not linearly dependent ([Mele, 2016](#)). So are the coefficients of regressors in the network interaction models ([Bramoullé et al., 2009](#)). In Bayesian theory, a parameter is identifiable if data can update the conditional posterior distribution of the parameter to distinguish it from the prior distribution ([Drèze, 1974](#); [Poirier, 1998](#)) and the posterior distribution becomes more concentrated at a single peak when data are more informative ([Koop et al., 2013](#)). To show that the incentive and endogenous effects in our model are identified, we conduct a simulation experiment which

shows that the conditional posterior distributions of these parameters collapse to the true values when the sample size increases. The details of this simulation experiment are left in Appendix D.

3.2 Posterior distributions and the MCMC

The posterior distribution of parameters considered here is based on the model with one continuous activity variables.¹⁵ To facilitate the Bayesian posterior analysis, it is natural to include the sampling of unobservables (including group fixed effects and latent variables) during the MCMC procedure as an augmentation (Albert and Chib, 1993). We use the notation $\{A_g\}$ to represent the collection of variable A_g across G groups, i.e., $\{A_g\} := (A_1, \dots, A_G)$. By Bayes' theorem, the joint posterior distribution of the parameters and unobservables in the model with exogenous variables, $\{X_g\}$ and $\{C_g\}$ suppressed for simplicity is

$$P(\{\theta_{cg}\}, \{\alpha_{cg}\}, \{Z_g\} | \{Y_{cg}\}, \{W_g\}) \propto \pi(\{\theta_{cg}\}, \{\alpha_{cg}\}, \{Z_g\}) \cdot \prod_{g=1}^G P(Y_{cg}, W_g | \theta_{cg}, \alpha_{cg}, Z_g), \quad (14)$$

where $\pi(\cdot)$ represents the density function of a prior distribution. We assume independence across prior distributions of common parameters, group effects, and latent variables, namely, $\pi(\{\theta_{cg}\}, \{\alpha_{cg}\}, \{Z_g\}) = \pi_1(\{\theta_{cg}\})\pi_2(\{\alpha_{cg}\})\pi_3(\{Z_g\})$. It is not easy to directly simulate draws from the joint posterior density in Eq. (14). But one can use the Gibbs sampling algorithm and work on the conditional posterior densities of parameters.

By properly blocking parameters in $\{\theta_{cg}\}$ into subgroups, we define prior distributions for parameters and other unknown variables in the model as follows:

- (i) Latent variables in both network formation and economic activity equations,

$$z_{i,g} \sim \mathcal{N}_{\bar{\ell}}(0, I_{\bar{\ell}}), \quad i = 1, \dots, m_g, \quad g = 1, \dots, G.$$

- (ii) Coefficients of network formation utility,

$$\phi = (\gamma', \eta', \delta_c) \sim \mathcal{N}_{2\bar{s}+\bar{q}+\bar{\ell}+\bar{h}+1}(\phi_0, \Phi_0 I_{2\bar{s}+\bar{q}+\bar{\ell}+\bar{h}+1}), \quad \phi \in O = \{\phi \in \mathbb{R}^{2\bar{s}+\bar{q}+\bar{\ell}+\bar{h}+1} | \delta_c \geq 0\}.$$

¹⁵We provide the posterior analysis in correspondence to the model with one continuous and one Tobit-type activity variables in Appendix C.

(iii) Coefficient of endogenous effect in the economic activity equation,

$$\lambda_c \sim U[-1/\tau_G, 1/\tau_G].$$

(iv) Coefficients of own and contextual effects in the economic activity equation,

$$\beta_c \sim \mathcal{N}_{2k}(\beta_0, B_0 I_{2k}).$$

(v) Coefficients of own and contextual correlated effects in the economic activity equation,

$$\rho_c \sim \mathcal{N}_{2\bar{\ell}}(\rho_0, R_0 I_{2\bar{\ell}}).$$

(vi) Variance of disturbances in the economic activity equation,

$$\sigma_{\xi_{cg}}^2 \sim \mathcal{IG}\left(\frac{\kappa_0}{2}, \frac{\nu_0}{2}\right), \quad g = 1, \dots, G.$$

(vii) Group fixed effects in the economic activity equation,

$$\alpha_{cg} \sim \mathcal{N}(\alpha_0, A_0), \quad g = 1, \dots, G.$$

In the above prior distributions, except the one for λ_c , they are conjugate priors commonly used in the Bayesian literature. We assign γ , η , and δ_c into the group, ϕ , since they are all (linear) coefficients in the function, $V(W_g)$, for network formation. We require the incentive effect, δ_c , to be non-negative. This constraint helps us to rule out the case of negative incentive effects, which is not reasonable for the utility specification. Thus, the prior distribution of ϕ is a truncated multivariate normal on the convex cone, O , with ϕ_0 and Φ_0 being, respectively, a prior mean vector and a variance before truncation. For λ_c , we employ a uniform prior as suggested in [Smith and LeSage \(2004\)](#) on the area $A = [-1/\tau_G, 1/\tau_G]$, with $\tau_G = \max\{\tau_1^*, \dots, \tau_G^*\}$ and $\tau_g^* = \{\max_{1 \leq i \leq m_g} \sum_{j=1}^{m_g} |w_{ij,g} + w_{ji,g}|\}$. As pointed out by [Kelejian and Prucha \(2010\)](#), $I_{m_g} - \lambda_c(W_g + W_g')$ is nonsingular for all values of λ_c in this interval. The prior of $\sigma_{\xi_{cg}}^2$ is specified as an inverse Gamma distribution with the shape and scale parameters governed by $\frac{\kappa_0}{2}$ and $\frac{\nu_0}{2}$. In the specification of (vii), we treat the group effects α_{cg} as fixed effects with the hyperparameters, α_0 and A_0 , fixed in their prior distributions. The distinction between fixed and random effects in a Bayesian approach lies on prior assignment at the second and the third levels of hierarchy ([Lancaster, 2004](#); [Rendon, 2013](#)). For a fixed effect model, a Bayesian approach updates distributions of fixed effect parameters, while for a random effect model it updates distributions of hyperparameters in the prior distribution of random effect parameters. If it were preferred to explicitly model

the correlation between covariates and group effects, one may follow [Mundlak \(1978\)](#) to have a random effect specification which allows the mean of random effect (i.e., the mean hyperparameter in the prior distribution of the random effect) to be a linear function of covariates.¹⁶ To show if there is any impact due to the specification of random group effects, we also examine the estimation results of our model based on the correlated random effect specification for a robustness check.

Within each Gibbs sampling step, random draws can be simulated from the conditional posterior distribution for a parameter group. We provide the list of conditional posterior distributions used by the Gibbs sampler in the supplementary appendix. Since any of the conditional posterior distributions is not available in a closed form, we use the Metropolis-Hastings (M-H) algorithm to draw from those conditional distributions. [Tierney \(1994\)](#) and [Chib and Greenberg \(1996\)](#) have shown that the combination of Markov chains (Metropolis-within-Gibbs) is still a Markov chain with the invariant distribution being the correct objective distribution. The procedure of the MCMC sampling starts with arbitrary initial values for $\{\alpha_{cg}^{(0)}\}$, $\{Z_g^{(0)}\}$, and $\{\theta_{cg}^{(0)}\}$, and then the sampling proceeds sequentially from the above set of conditional posterior distributions. A further detail about implementing the MCMC sampling based on steps (i) to (vii) is available in the supplementary appendix of this paper. In the following section, we will discuss a relative new version of the M-H algorithm which can be used when the likelihood function contains an intractable normalizing term.

3.3 Double M-H algorithm

From Section 3.1, the likelihood function of $y = (\{Y_{cg}\}, \{W_g\})$, given the vector of parameters and unknown variables θ , which refers to the vector $(\{\theta_{cg}\}, \{\alpha_{cg}\}, \{Z_g\})$, is $P(y|\theta) =$

¹⁶We specify α_{cg} as follows:

$$\alpha_{cg} = \bar{X}_g \beta_{3c} + \bar{Z}_g \rho_{3c} + \zeta_{cg}, \quad \zeta_{cg} \sim \mathcal{N}(0, \sigma_{\alpha,c}^2),$$

where \bar{X}_g and \bar{Z}_g are, respectively, the group averages of X_g and Z_g . The β_{3c} , ρ_{3c} , and $\sigma_{\alpha,c}^2$ are unknown parameters, and we also specify prior distributions for them such that $\beta_{3c} \sim \mathcal{N}_k(\beta_0, B_0 I_k)$, $\rho_{3c} \sim \mathcal{N}_{\bar{\ell}}(\rho_0, R_0 I_{\bar{\ell}})$, and $\sigma_{\alpha,c}^2 \sim \mathcal{IG}(\frac{\kappa_0}{2}, \frac{\nu_0}{2})$. As a result, we form a hierarchical prior for α_{cg} where the prior mean and the prior variance of α_{cg} follow other prior distributions.

$\prod_{g=1}^G P(Y_{cg}, W_g | \theta_{cg}, \alpha_{cg}, Z_g)$, where $P(Y_{cg}, W_g | \theta_{cg}, \alpha_{cg}, Z_g)$ is from Eq (13). It takes the form $P(y|\theta) = f(y; \theta)/D(\theta)$, where $D(\theta)$ is an intractable normalizing term. The standard M-H algorithm to simulate random draws of θ runs as follows: given an old draw, θ_{old} , one proposes a new one, θ_{new} , from a proposal distribution, $q(\cdot|\theta_{old})$, and then updates the old draw to the new draw with an acceptance probability, $\alpha(\theta_{new}|\theta_{old})$. Denoting $\pi(\theta)$ as the prior distribution of θ , the acceptance probability needs to be computed as

$$\begin{aligned} \alpha(\theta_{new}, \theta_{old}) &= \min \left\{ 1, \frac{P(\theta_{new}|y)q(\theta_{old}|\theta_{new})}{P(\theta_{old}|y)q(\theta_{new}|\theta_{old})} \right\} \\ &= \min \left\{ 1, \frac{\pi(\theta_{new})f(y; \theta_{new})q(\theta_{old}|\theta_{new})}{\pi(\theta_{old})f(y; \theta_{old})q(\theta_{new}|\theta_{old})} \cdot \frac{D(\theta_{old})}{D(\theta_{new})} \right\}. \end{aligned} \quad (15)$$

One can see that in Eq. (15), the normalizing terms, $D(\theta_{old})$ and $D(\theta_{new})$, are left in both the numerator and denominator and will not cancel out, so the evaluation of the acceptance-rejection criterion based on α in Eq. (15) would be intractable. Murray et al. (2006) consider to include auxiliary variables, $\tilde{y} = (\{\tilde{Y}_{cg}\}, \{\tilde{W}_g\})$, into the acceptance probability, i.e., the acceptance probability can be written as

$$\begin{aligned} \alpha(\theta_{new}, \theta_{old}, \tilde{y}) &= \min \left\{ 1, \frac{\pi(\theta_{new})P(y|\theta_{new})q(\theta_{old}|\theta_{new})}{\pi(\theta_{old})P(y|\theta_{old})q(\theta_{new}|\theta_{old})} \cdot \frac{P(\tilde{y}|\theta_{old})}{P(\tilde{y}|\theta_{new})} \right\} \\ &= \min \left\{ 1, \frac{\pi(\theta_{new})f(y; \theta_{new})q(\theta_{old}|\theta_{new})}{\pi(\theta_{old})f(y; \theta_{old})q(\theta_{new}|\theta_{old})} \cdot \frac{f(\tilde{y}; \theta_{old})}{f(\tilde{y}; \theta_{new})} \right\}, \end{aligned} \quad (16)$$

where \tilde{y} are simulated from the likelihood function, $P(\tilde{y}|\theta_{new}) = f(\tilde{y}; \theta_{new})/D(\theta_{new})$ with the exact sampling (Propp and Wilson, 1996). In the conditional acceptance probability of Eq (16), all normalizing terms cancel out and the other terms left are computable. This algorithm is called the “exchange algorithm” because a swapping operation between (θ_{old}, y) and $(\theta_{new}, \tilde{y})$ is involved (Geyer, 1991). The exchange algorithm is different from the conventional M-H algorithm by adding a randomization component into the proposal density, which changes $q(\theta_{new}|\theta_{old})$ into $q(\theta_{new}|\theta_{old})P(\tilde{y}|\theta_{new})$. The exchange algorithm defines a valid Markov chain for simulating from $P(\theta|y)$ (Murray et al., 2006; Liang, 2010; Liang et al., 2016). However, to implement the exchange algorithm is time consuming because it requires the exact sampling of \tilde{y} from $P(\tilde{y}|\theta_{new})$. In order to save time on the computation, Liang (2010) proposes a “double M-H algorithm” which utilizes the reversibility condition and shows that when \tilde{y} is simulated by the M-H algorithm starting from y with m iterations, the conditional

acceptance probability in Eq. (16) can be obtained regardless of the value of m . This gives the double M-H algorithm an advantage, as a small value of m can be used, removing the need for the exact sampling. Due to this computational efficiency compared to the exact sampling for our model, we adopt the double M-H algorithm in this study.

One thing worth mentioning is that in this paper we provide a technical modification on the double M-H algorithm which can simplify the simulation and better fit into our application. Using the double M-H algorithm to update θ from $P(\theta|y)$ requires simulating auxiliary variable, \tilde{y} . However, the auxiliary activity variables, $\{\tilde{Y}_{cg}\}$ in \tilde{y} are redundant as they can be fully replaced by a function of auxiliary networks, $\tilde{w} = \{\tilde{W}_g\}$. Therefore, we modify the conditional acceptance probability in Eq. (16) to

$$\begin{aligned} \alpha(\theta_{new}, \theta_{old}, \tilde{w}) &= \min \left\{ 1, \frac{\pi(\theta_{new})P(y|\theta_{new})q(\theta_{old}|\theta_{new})}{\pi(\theta_{old})P(y|\theta_{old})q(\theta_{new}|\theta_{old})} \cdot \frac{P(\tilde{w}|\theta_{old})}{P(\tilde{w}|\theta_{new})} \right\} \\ &= \min \left\{ 1, \frac{\pi(\theta_{new})f(y; \theta_{new})q(\theta_{old}|\theta_{new})}{\pi(\theta_{old})f(y; \theta_{old})q(\theta_{new}|\theta_{old})} \cdot \frac{f(\tilde{w}; \theta_{old})}{f(\tilde{w}; \theta_{new})} \right\}. \end{aligned} \quad (17)$$

To evaluate $\alpha(\theta_{new}, \theta_{old}, \tilde{w})$ in Eq. (17), we only simulate the auxiliary networks, \tilde{w} , from the probability density function, $P(\tilde{w}|\theta_{new}) = f(\tilde{w}; \theta_{new})/D(\theta_{new})$, which shares the same normalizing term, $D(\theta_{new})$, with $P(\tilde{y}|\theta_{new})$. We leave the details of simulating \tilde{w} in the supplementary appendix.

4 Friendship network, academic and smoking behaviors

We apply our model to study American high school students' friendship networks in the Add Health data, which is a national survey based on grades 7 through 12 in 132 schools (Udry, 2003).¹⁷ Four waves of surveys were conducted between 1994 and 2008. In the wave I in-school survey, a total of 90,182 students were interviewed. Respondents answered questions

¹⁷This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health

about their demographic backgrounds, academic performances, and health-related behaviors. Most uniquely, students were asked to nominate up to five male and five female friends. This provides information about their friendship networks. In the following waves of in-home surveys, more information about students' families and neighborhoods in which they live are available for a subset of the total sample. To accommodate most of the students' nominated friends into our framework, the sample used in this study is constructed from the wave I in-school survey. We consider two activities which may be relevant for friendship formation. One is a student's academic performance (measured by GPA, which is regarded as a proxy for studying activity), which is represented by a continuous variable. The other is how frequently a student smokes in a week, which is represented by a Tobit-type variable.¹⁸

In the context of social interactions, students' academic performance and smoking behavior are extensively studied as they have important long-term consequences on students' future lives and health. Studies of peer effects on students' academic performance, including [Hoxby \(2000\)](#), [Sacerdote \(2001\)](#), [Hanushek et al. \(2003\)](#), and [Zimmerman \(2003\)](#), use the linear-in-means model; and [Calvó-Armengol et al. \(2009\)](#), [Lin \(2010\)](#), [Boucher et al. \(2014\)](#), and [Liu et al. \(2014\)](#) use the network interactions model. For studies of peer effects on students' smoking behaviors, evidence of peer effects can be found from [Gaviria and Raphael \(2001\)](#), [Powell et al. \(2005\)](#), [Lundborg \(2006\)](#), [Clark and Lohéac \(2007\)](#), [Fletcher \(2010\)](#), and [Hsieh and Van Kippersluis \(2016\)](#). When studying interaction (peer) effects, researchers face difficulty in identifying correlated effects from group-level unobservables and endogenous selection into groups ([Moffitt et al., 2001](#)), and separating the endogenous interaction effect from contextual effects in a linear model (the reflection problem by [Manski \(1993\)](#)). With various approaches (e.g., randomization, fixed effects, etc.) to avoid these difficulties, researchers generally provide evidences for the existence of peer effects. [Hsieh and Lee \(2016\)](#) further consider the problem of endogenous friendship selection on peer effects on economic website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from grant P01-HD31921 for this analysis.

¹⁸Discussions about how academic performance and smoking affect friendship selections, see, e.g., [Kiuru et al. \(2010\)](#), [Lomi et al. \(2011\)](#), [Flashman \(2012\)](#), [Schaefer et al. \(2013\)](#). Other activities may affect friendship choices. We focus on academic performance and smoking because they are the key subjects of interest discussed in social interaction studies.

activities by modeling unobservables in both the network interaction and network formation process. They find that the endogenous effect on academic performance obtained from the SAR model without controlling the endogeneity of the spatial weight matrix can be upward biased. In the present study, we control individual unobservables in formation of friendship networks as in [Hsieh and Lee \(2016\)](#). Moreover, we find that the benefit of interactions from academic learning is an important factor for students to form friendships.

4.1 Structural network effects in the link-associated utility

For our empirical study, we consider the following specification of the structural network effects in the link-associated utility of Eq. (6),

$$\begin{aligned}
& \varpi_{i,g}(w_{i.,g}, W_{-i.,g})\eta \\
&= \underbrace{\eta_1 \sum_{j=1}^{m_g} w_{ij,g} w_{ji,g}}_{\text{Reciprocity Effect}} + \underbrace{\eta_2 \sum_{j=1}^{m_g} w_{ij,g} \left(\sum_{k \neq j}^{m_g} w_{ik,g} \right) + \eta_3 \sum_{j=1}^{m_g} w_{ij,g} \left(\sum_{k \neq j}^{m_g} w_{ik,g} \right)^2}_{\text{Sender's Expansiveness Effect}} + \underbrace{\eta_4 \sum_{j=1}^{m_g} w_{ij,g} \left(\sum_{k \neq i}^{m_g} w_{kj,g} \right)}_{\text{Receiver's Popularity Effect}} \\
&+ \underbrace{\eta_{51} \sum_{j=1}^{m_g} w_{ij,g} \left(\sum_k^{m_g} w_{ik,g} w_{kj,g} \right) + \eta_{52} \sum_{j=1}^{m_g} w_{ij,g} \left(\sum_k^{m_g} w_{ki,g} w_{kj,g} \right) + \eta_{53} \sum_{j=1}^{m_g} w_{ij,g} \left(\sum_k^{m_g} w_{ik,g} w_{jk,g} \right)}_{\text{Transitive Triads Effect}} \\
&+ \underbrace{\eta_6 \sum_{j=1}^{m_g} w_{ij,g} \left(\sum_k^{m_g} w_{jk,g} w_{ki,g} \right)}_{\text{Three Cycles Effect}}. \tag{18}
\end{aligned}$$

In Eq. (18), the reciprocity effect reflects the utility from reciprocal friendships. The sender's expansiveness effect in Eq. (18) reflects the utility from being an outgoing person who actively nominates friends. The statistics involved in the sender's expansiveness effect are the sender's outdegree and the outdegree square. We expect the coefficient, η_3 , would be negative to reflect the reality that individuals might not make too many friends due to limited resources, e.g., limited time, energy, and money. The receiver's indegree is used to measure the receiver's popularity effect in Eq. (18), which reflects the utility from making friends with someone who is popular. The transitive triads effect and the three cycles effect both reflect the utility from engaging in a transitive relationship, i.e., friends of my friends are my friends. However, they are distinguished by directions of links. From [Kovářík and van der Leij \(2014\)](#), transitive triads effects may be linked to an individual's sense of risk aversion.

The three-cycles effect can be interpreted as an opposite hierarchy effect (Snijders et al., 2010). If the coefficient η_6 is negative, it implies a local hierarchy among linked individuals. The effects that we specify in Eq. (18) are motivated by which network structural statistics represent the potential benefits and costs of network connections. They can be categorized into the first, second, and third order interaction terms of network links. More discussion on the categorization is provided in the supplementary appendix.

Given $\varpi_{i,g}(w_{i.,g}, W_{-i.,g})\eta$ in Eq. (18), the term $\sum_{i=1}^{m_g} \varpi_{i,g}(w_{i.,g}, W_{-i.,g})\eta$ in the transferable utility of Eq.(9) can be written as

$$\begin{aligned} \sum_{i=1}^{m_g} \varpi_{i,g}(w_{i.,g}, W_{-i.,g})\eta &= \eta_1 \text{tr}(W_g^2) + \eta_2 (l'_g W'_g W_g l_g - l'_g W_g l_g) \\ &\quad + \eta_3 (l'_g W'_g \text{Diag}(W_g l_g) W_g l_g - 2l'_g W'_g W_g l_g + l'_g W_g l_g) \\ &\quad + \eta_4 (l'_g W_g W'_g l_g - l'_g W_g l_g) + (\eta_{51} + \eta_{52} + \eta_{53}) \text{tr}(W_g^2 W'_g) + \eta_6 \text{tr}(W_g^3), \end{aligned} \tag{19}$$

where l_g is a column vector of ones, and $\text{Diag}(A)$ is a $n \times n$ diagonal matrix with its diagonal elements formed by the entries of a $n \times 1$ vector of A . One can see that parameters η_{51} , η_{52} and η_{53} are not separately identified from Eq. (19). Hence, without loss of generality, we will use η_5 for $\eta_{51} + \eta_{52} + \eta_{53}$ hereafter.

4.2 Data summary

To ease the computation burden, we only work with small networks in this study. The following steps are used to construct the sample. First, we group students by their school and grade level and consider friendships only inside the same group. In the Add Health data, about 80% of friendship nominations happen within the same grade level. Hence, about 20% of links are missed due to the design of network boundary. So effectively, we are investigating friendship formation within a school grade, and such a friendship network's effects on student activities. Second, we focus on senior high school students from 9th to 12th grades. Third, we restrict our network sample to those groups with sizes between 10 and 50 (10 and 60 for the smoking case). A total of 1,290 (1,510 for the smoking case) respondents from 47 networks (44 networks for the smoking case) are utilized for analysis.

There are 113 (9.6%) missing observations for GPA and 34 (2%) missing observations for smoking.¹⁹ These networks have the average size equals to 25.043 (33.546 for the smoking case), average density equals to 0.142 (0.108 for the smoking case), average outdegree equals to 2.564 (2.866 for the smoking case), and average clustering coefficient equal to 0.327 (0.332 for the smoking case). In the network formation model, we capture an individual-specific effect by a dummy variable of whether a student is older than the group average or not. Three other dummy variables – whether a pair of students has the same age, same sex, or same race – are used to capture dyad-specific effects.

For the activity under network interaction model, the continuous variable, GPA, is calculated by the average of a respondent’s reported grades from several subjects, including language, social science, mathematics, and science, of which each has a value between 1 and 4. The Tobit-type variable, smoking, is obtained from student’s response to the survey question, “*During the past twelve months, how often did you smoke cigarettes?*”, which has a value between 0 and 7. We follow [Lin \(2010\)](#), [Lee et al. \(2010\)](#), and [Hsieh and Lee \(2016\)](#) to choose the independent variables. A complete list of variables is provided in [Table 1](#). In [Figures 1](#) and [2](#), we plot two typical networks from our sample – one is from the GPA sample and the other is from the smoking sample. From these two figures, one can observe that students who have higher GPAs tend to receive more friendship nominations than those who have lower GPAs. This observation does not seem to be evident for smoking behaviors, but one can see that smokers are friends with each other. Our estimation results shown in the following section provide evidence for the incentive stemming from interaction benefits on academic learning, but not from the pleasure of smoking together, on friendship decisions. Smokers are making friends with other smokers through homophily on smoking.

To obtain estimates from the Bayesian estimation in this empirical study, the values of hyperparameters in the prior distributions are set as follows:²⁰ $\phi_0 = 0$; $\Phi_0 = 10$; $\beta_0 = 0$; $B_0 = 10$; $\rho_0 = 0$; $R_0 = 10$; $\sigma_0 = 0$; $\Sigma_0 = 10$; $\alpha_0 = 0$; $A_0 = 400$. These specified values of hyperparameters are chosen to form very flat prior densities over the ranges of parameter

¹⁹Individuals with missing economic activities are eventually kept in sample for estimation as Bayesian methods can easily handle missing observations on dependent variables via simulation.

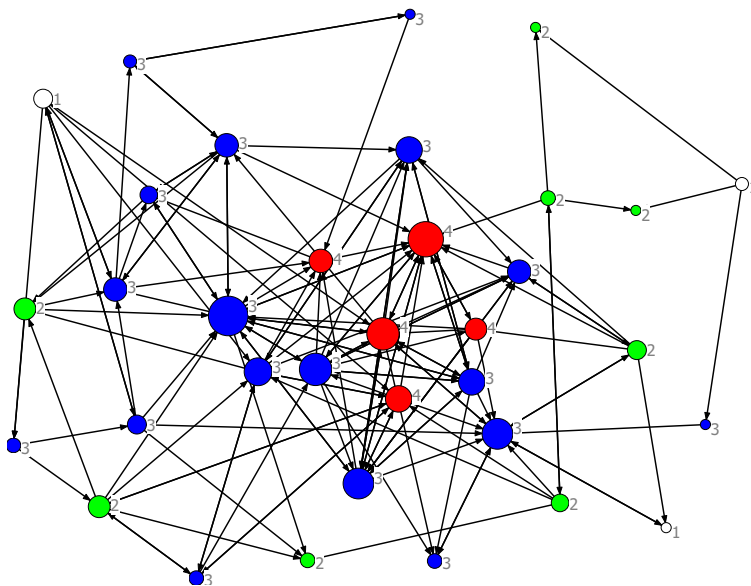
²⁰See [Appendix C](#) for the prior distribution assumptions for a full model with continuous and Tobit-type activities.

Table 1: Summary Statistics

variable	min	max	GPA				Smoking	
			drop missing		keep missing		drop missing	
			mean	s.d.	mean	s.d.	mean	s.d.
GPA	1	4	2.910	0.734	-	-	-	-
Smoking	0(57.86%)	7	-	-	-	-	1.257	2.511
Age	10	19	16.004	1.285	15.960	1.294	15.997	1.269
Male	0	1	0.493	0.500	0.490	0.500	0.482	0.499
<i>Female</i>	0	1	0.507	0.500	0.510	0.500	0.517	0.499
<i>White</i>	0	1	0.611	0.487	0.522	0.500	0.629	0.483
Black	0	1	0.246	0.430	0.326	0.469	0.230	0.421
Asian	0	1	0.016	0.125	0.013	0.114	0.016	0.123
Hispanic	0	1	0.068	0.251	0.076	0.265	0.067	0.250
Other race	0	1	0.059	0.236	0.063	0.243	0.058	0.233
Both parents	0	1	0.725	0.447	0.697	0.460	0.733	0.442
Less HS	0	1	0.114	0.318	0.110	0.313	0.109	0.312
<i>HS</i>	0	1	0.340	0.473	0.353	0.478	0.341	0.474
More HS	0	1	0.398	0.490	0.374	0.484	0.402	0.490
Edu missing	0	1	0.068	0.252	0.073	0.260	0.067	0.250
Professional	0	1	0.248	0.432	0.238	0.426	0.249	0.432
<i>Staying home</i>	0	1	0.220	0.414	0.227	0.419	0.228	0.419
Other Jobs	0	1	0.366	0.481	0.339	0.473	0.356	0.479
Job missing	0	1	0.076	0.265	0.093	0.291	0.077	0.266
Welfare	0	1	0.011	0.103	0.120	0.111	0.010	0.100
Num. of students at home	0	6	0.580	0.818	0.584	0.837	0.568	0.793
Network size			25.043	13.146	27.447	13.790	33.546	16.551
Network density			0.142	0.100	0.137	0.102	0.108	0.076
Outdegree			2.564	2.294	2.652	2.353	2.866	2.406
Indegree			2.564	2.418	2.652	2.487	2.866	2.596
Clustering Coef.			0.327	0.120	0.344	0.105	0.332	0.086
Sample size			1,177		1,290		1,476	
Num. of networks			47		47		44	

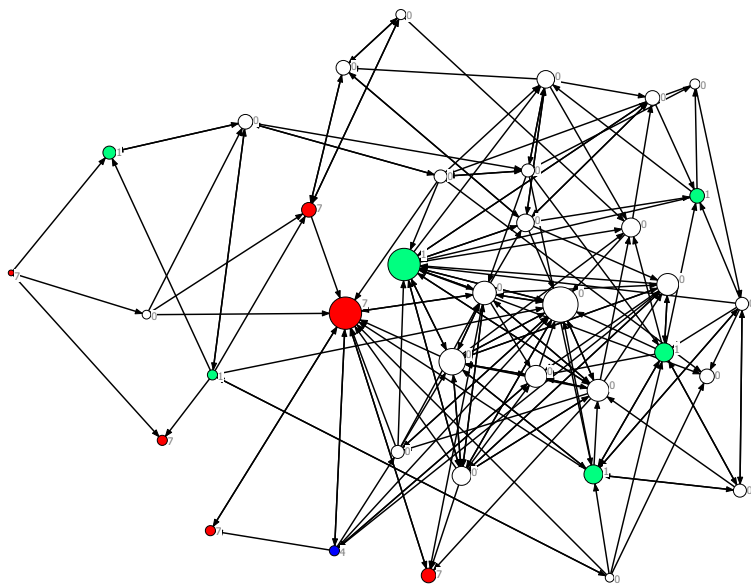
Note: ‘Both parents’ means living with both parents. ‘Less HS’ means mother’s education is less than high school. ‘Edu missing’ means mother’s education level is missing. ‘Professional’ means mother’s job is either scientist, teacher, executive, director and the like. ‘Other jobs’ means mother’s occupation is not among Professional or Staying home. ‘Welfare’ means mother participates in social welfare programs. ‘Num. of students at home’ means number of other students from grade 7 to 12 living in the same household with the student. The variables in italics are omitted categories during estimation. We drop missing observations in the ‘drop missing’ case. There are only 2% of missing observations on smoking and therefore we do not consider the ‘keep missing’ case in the smoking sample.

Figure 1: A friendship network from the GPA sample



Note: The number (and color) for each node indicates the value of GPA. Nodes with a larger size means they have higher indegrees.

Figure 2: A friendship network from the Smoking sample



Note: The number (and color) for each node indicates the frequency of smoking. Nodes with a larger size means they have higher indegrees.

spaces so that estimation results are less influenced by our choice of priors. The reported estimation results are based on the MCMC sampling draws which pass the convergence test provided by [Geweke \(1992\)](#).

4.3 Estimation results

4.3.1 Case of GPA

We first estimate the model with a single incentive effect from the academic activity measured by GPA, which is a continuous variable. The sample contains 9.6% of observations that have missing values on GPA and thus we apply the Bayesian data imputation approach to recover them in estimation. An intuition of this data imputation approach is in [Appendix E](#). We report the results in [Table 2](#). From columns 1 to 4, we separately consider the full model, the model without latent characteristic variables, the model without the incentive effect in network formation, and the activity outcome equation under network interaction assuming exogenous network links. We compare the full model with possibly misspecified models in order to see how each model misspecification may affect estimates of the full model, particularly for the estimate of endogenous interaction effect on outcomes (λ_c). In the table, the mean and the standard deviation (in parenthesis) of MCMC posterior draws are reported as point estimates for each parameter.

The result of network formation in the full model shows that whether being older than the group average or not does not have a significant effect on sending or receiving friendship nominations. The exogenous dyad-specific effects are all positive and significant, where the effect of the same race (0.4904) is strongest, followed by the effects of the same sex (0.4308) and the same age (0.1081). The distances of three latent characteristic variables have significant negative effects (-0.4546, -0.3459, and -0.1913) on network formation, confirming the existence of homophily with respect to unobservables ([Goldsmith-Pinkham and Imbens, 2013](#); [Hoff et al., 2002](#); [Hsieh and Lee, 2016](#)).²¹

²¹We follow [Hsieh and Lee \(2016\)](#) to specify the dimension of latent variables to be three. We have also tried the fourth dimension and found there is no further change on the estimates of endogenous peer effect and other coefficients. Therefore, the results suggest that three dimensions of latent variables suffice to correct the endogeneity bias. Also notice that we are unable to evaluate the exact likelihood values of the

Among structural network effects, the positive and strong reciprocity effect (1.1632) is consistent with findings in the literature (Badev, 2013; Snijders et al., 2010; Mele, 2016), which reflects that mutual friendship nominations among students are pervasive (49.8% of friendship links) in our sample. The sender’s expansiveness effect is concave, as the coefficient of the first order term is positive (0.2314) and the coefficient of the second order term is negative (-0.0270). This result confirms our conjecture that limited resources, e.g., time, energy, and money, might constrain students from making too many friends. The receiver’s popularity effect is negative (-0.0274), which suggests that students between 9th and 12th grades in our sample are less willing to make friends with someone who is popular. The positive and strong transitive triads effect (0.5735) shows that students value transitive relationships. When the positive triads effect is accompanied by the negative three cycles effect (-0.2285), as discussed in Snijders et al. (2010), a certain degree of local hierarchy among students is revealed. The incentive effect from GPA is found to be large and significant (0.7010). Therefore, for high school students in our sample, the potential benefit of learning from others in school work is a factor which influences their friendship decisions.

Our important finding on academic performance under network interactions is that, by controlling network endogeneity through both the latent characteristic variables and the incentive effect, the estimated endogenous effect on academic performance drops from 0.0138 in the network interaction model alone (column 4) to 0.0077 in the full model (column 1). This result demonstrates the effectiveness of our joint modeling approach for network and activity outcome on correcting the selection bias inherited in the activity outcome equation. Results of two restricted models in columns 2 and 3 further show that the bias correction comes from both the incentive effect and the unobserved latent characteristic variables. 35% of the bias is due to omitting the incentive effect and 10% of bias is due to omitting the latent variables. Our estimate implies that the social multiplier effects, measured by the vectors $(I_{m_g} - \lambda_c(W_g + W'_g))^{-1}l_{m_g}$ across individuals in a group and across groups, have the average equals to 1.0121 and the standard deviation equals to 0.005.

From the estimates of own effects of individual characteristics, we observe that students who are older and male tend to have lower GPAs. On the contrary, students who are

model and therefore we cannot perform the same AICM model selection criterion as in Hsieh and Lee (2016).

Hispanic, Asian, or whose mothers have education more than high school tend to have higher GPAs. We also see two latent characteristic variables (2^{nd} and 3^{rd} dimensions) show significant positive effects on GPA. Estimates of all contextual effects, observed or not, are found insignificant in the full model.

4.3.2 Case of smoking

Next, we consider the model with smoking activity. The results are reported in Table 3. As there are only 2% of missing observations on the variable of smoking, the corresponding individual observations are dropped.²² From the part of network formation in the full model, we find that students who are older than the group average tend to receive fewer friendship nominations (-0.0858). The estimates of exogenous dyad-specific effects show significant homophily in terms of sex and race, but not age, for friendship formation. In order to capture the possible feature that smokers tend to make friends of each other, we add the absolute difference of smoking frequencies between individuals as an additional control of homophily effect in network formation. Since smoking frequency is an activity outcome under interactions, the difference of smoking frequencies in network formation is an endogenous factor, which represents the feedback of activity outcome in network formation as a homophily effect instead of an incentive effect. The results confirm that there is a significant homophily effect in terms of smoking activity (-0.0399). The effects of homophily on other three unobserved characteristics (-0.1557, -0.1331, and -0.1152) are also significant here, even those estimates are lower in absolute value than those in the GPA case. Structural network effects are generally similar to those in the GPA case. An important finding is that the incentive effect from smoking is small and insignificant (0.0041). Hence, we conclude that students in our sample do not consider the interaction benefit from smoking as a factor for their friendship decisions but the homophily effect in smoking behavior matters.

For the equation of smoking activity under network interactions, the estimated endogenous (peer) effect drops from 0.0620 in the network interaction model alone (column 4) to 0.0534 in the full model (column 1). This change again shows that our modeling approach can correct the selection bias. We also see that the correction is more due to the latent

²²We do not expect data imputation would have an effect on estimates.

variable (column 3) instead of the incentive effect (column 2). This result echoes our finding from the network formation model that the incentive effect of smoking is small and insignificant; therefore, it does not cause a selection bias on the estimate of endogenous effect in the smoking activity equation. Our estimate implies that the social multipliers have the average and standard deviation equal to 1.0943 and 0.0388, respectively.

The estimated own effects of characteristics show that students who are Black, Asian, or who live with both parents tend to smoke less than their counterparts. On the contrary, students who are older, male, and whose mothers participated in welfare programs tend to smoke more than others. From contextual effects, a student may smoke more if he or she is surrounded by more friends who are Black, or whose mothers have higher education levels. A student may smoke less by owning more male friends.

4.3.3 Case with both GPA and smoking

Lastly, we estimate the model with incentive effects from both GPA and smoking.²³ From the results reported in Table 4, the parameter estimates in the network formation model are close to those in Tables 2 and 3. Particularly, the joint modeling of both incentive effects from GPA and smoking does not affect the estimate of each single effect, which supports the separability of these incentive effects.

For the outcome equation of GPA, the estimated endogenous interaction effect changes from 0.0159 in the equation under network interaction by ignoring network endogeneity (column 4) to 0.0121 in the full model (column 1), which again shows a significant correction on the friendship selection bias. When further comparing columns 2 and 3, we can see the correction comes from both the incentive effect and the latent characteristic variables. For the outcome equation of smoking, there is also a correction on the estimated endogenous interaction effect from 0.0579 in the equation assuming an exogenous network (column 4) to 0.0475 in the full model (column 1). This correction relies more on the inclusion of latent variables instead of the inclusion of the incentive effect (as seen from a larger drop from 0.0579 in column 4 to 0.0490 in column 3, compared to a smaller drop to 0.0513 in column

²³The sample used to estimate this model is based on the original GPA sample where we remove missing observations on both GPA and smoking.

2.) Finally, it is interesting to note that the jointing modeling of two activity outcomes allows us to estimate the covariance of disturbances in the outcome equations between GPA and smoking, which is found to be around -0.74 with some variations across groups.

For a robustness check, we additionally estimate the model with correlated random group effects as discussed in Footnote 16 of Section 3.2. The full estimation results are available in Table A1 of the Appendix for interested readers to check. We find the coefficient estimates in the network formation model and the endogenous effects in the activity outcome equations remain similar to those in Table 4. There are few significant changes on estimated own effects. Taking the full model for example, the own effect of Asian on GPA drops from 0.2132 in Table 4 to 0.0669 in Table A1. Other significant changes include the own effect of welfare on GPA, and the own effect of Hispanic on smoking. There is no significant change on estimated contextual effects between two group effect specifications. Therefore, our estimation results seem quite robust between fixed and random group effect settings.

Table 2: Estimation result based on GPA

	Full		Without Latent		Without Incentive		Activity Alone	
Network Formation								
Higher sender age (γ_1)	-0.0075	(0.0440)	-0.0256	(0.0480)	-0.0284	(0.0323)	-	
Higher receiver age (γ_2)	0.0295	(0.0239)	-0.0197	(0.0464)	-0.0357	(0.0384)	-	
Constant (γ_{31})	-3.5055	(0.0693)	-4.4825	(0.0953)	-3.3828	(0.0942)	-	
Same age (γ_{32})	0.1081	(0.0580)	0.0922	(0.0416)	0.1657	(0.0234)	-	
Same sex (γ_{33})	0.4308	(0.0380)	0.3846	(0.0400)	0.4278	(0.0272)	-	
Same race (γ_{34})	0.4904	(0.0300)	0.4475	(0.0509)	0.5200	(0.0357)	-	
Latent dist. (γ_{41})	-0.4546	(0.0277)	-		-0.5246	(0.0675)	-	
Latent dist. (γ_{42})	-0.3459	(0.0194)	-		-0.2936	(0.0334)	-	
Latent dist. (γ_{43})	-0.1913	(0.0590)	-		-0.1475	(0.0329)	-	
Reciprocity (η_1)	1.1632	(0.0468)	1.2943	(0.0394)	1.1733	(0.0402)	-	
Expansiveness (η_2)	0.2314	(0.0268)	0.2486	(0.0229)	0.1968	(0.0156)	-	
Expansiveness (η_3)	-0.0270	(0.0025)	-0.0289	(0.0024)	-0.0241	(0.0017)	-	
Popularity (η_4)	-0.0274	(0.0059)	-0.0296	(0.0054)	-0.0323	(0.0052)	-	
Trans. triads (η_5)	0.5735	(0.0167)	0.5873	(0.0192)	0.6011	(0.0133)	-	
Three cycles (η_6)	-0.2285	(0.0159)	-0.2594	(0.0169)	-0.2447	(0.0138)	-	
Economic incentive (δ_c)	0.7010	(0.1119)	1.0027	(0.2120)	-		-	
Activity Outcome								
Endogenous (λ_c)	0.0077	(0.0024)	0.0089	(0.0015)	0.0124	(0.0026)	0.0138	(0.0028)
Own (X)								
Age	-0.1644	(0.0161)	-0.2172	(0.0145)	-0.2242	(0.0147)	-0.2398	(0.0314)
Male	-0.1131	(0.0375)	-0.1096	(0.0327)	-0.0949	(0.0380)	-0.1057	(0.0402)
Black	-0.0041	(0.0847)	-0.0252	(0.0574)	0.0282	(0.0656)	-0.0012	(0.0767)
Asian	0.1708	(0.1800)	0.2088	(0.1571)	0.1478	(0.1811)	0.1363	(0.1862)
Hispanic	0.2395	(0.1057)	0.2320	(0.0701)	0.2593	(0.0917)	0.2021	(0.1084)
Other race	-0.0501	(0.0938)	-0.1254	(0.0592)	-0.0252	(0.0960)	-0.0580	(0.0928)
Both Parents	0.0761	(0.0467)	0.0590	(0.0326)	0.0697	(0.0455)	0.0643	(0.0450)
Less HS	-0.0878	(0.0713)	-0.0792	(0.0470)	-0.0721	(0.0650)	-0.0930	(0.0678)
More HS	0.1362	(0.0479)	0.1711	(0.0412)	0.1153	(0.0461)	0.1559	(0.0469)
Edu missing	-0.1082	(0.0845)	-0.0680	(0.0542)	-0.0705	(0.0788)	-0.0476	(0.0815)
Welfare	0.1180	(0.1954)	-0.0635	(0.0959)	0.1115	(0.1963)	0.0857	(0.1939)
Job missing	-0.0858	(0.0769)	-0.0635	(0.0562)	-0.0885	(0.0811)	-0.0843	(0.0731)
Professional	-0.0383	(0.0609)	-0.0444	(0.0477)	-0.0433	(0.0571)	-0.0641	(0.0581)
Other Jobs	-0.0006	(0.0530)	-0.0089	(0.0498)	-0.0130	(0.0498)	-0.0233	(0.0491)
Num. of students at home	-0.0050	(0.0249)	-0.0009	(0.0226)	-0.0050	(0.0277)	0.0051	(0.0242)
latent (ρ_{c11})	-0.0430	(0.1144)	-		0.0981	(0.0707)	-	
latent (ρ_{c12})	0.4662	(0.0876)	-		0.5007	(0.0640)	-	
latent (ρ_{c13})	0.2117	(0.0695)	-		0.1620	(0.0486)	-	
Contextual (WX)								
Age	-0.0002	(0.0014)	-0.0007	(0.0009)	-0.0028	(0.0024)	-0.0020	(0.0023)
Male	-0.0002	(0.0155)	-0.0195	(0.0093)	0.0110	(0.0260)	-0.0185	(0.0260)
Black	-0.0078	(0.0135)	-0.0439	(0.0105)	-0.0273	(0.0192)	-0.0303	(0.0210)
Asian	0.0038	(0.0869)	-0.0072	(0.0477)	0.0077	(0.1563)	0.1459	(0.1611)
Hispanic	-0.0019	(0.0385)	0.0075	(0.0192)	0.0215	(0.0634)	0.0062	(0.0582)

Continued on Next Page

Table – Continued

Other race	-0.0301	(0.0362)	0.0076	(0.0247)	-0.0708	(0.0479)	-0.0842	(0.0547)
Both Parents	0.0074	(0.0180)	0.0466	(0.0122)	0.0405	(0.0268)	0.0438	(0.0296)
Less HS	-0.0133	(0.0298)	-0.0259	(0.0187)	-0.0546	(0.0467)	-0.0688	(0.0452)
More HS	-0.0013	(0.0170)	-0.0053	(0.0109)	-0.0116	(0.0274)	-0.0034	(0.0283)
Edu missing	0.0037	(0.0343)	-0.0156	(0.0223)	0.0484	(0.0535)	0.0313	(0.0562)
Welfare	-0.0110	(0.0992)	-0.0125	(0.0418)	-0.0600	(0.1532)	-0.1113	(0.1649)
Job missing	-0.0015	(0.0324)	-0.0320	(0.0198)	0.0054	(0.0459)	-0.0062	(0.0526)
Professional	0.0010	(0.0211)	-0.0054	(0.0130)	-0.0001	(0.0369)	0.0020	(0.0365)
Other Jobs	0.0004	(0.0185)	-0.0155	(0.0121)	0.0022	(0.0315)	-0.0031	(0.0320)
Num. of students at home	-0.0034	(0.0091)	-0.0097	(0.0067)	-0.0174	(0.0137)	-0.0069	(0.0155)
latent (ρ_{c21})	0.0051	(0.0588)	-		0.0330	(0.0313)	-	
latent (ρ_{c22})	-0.0275	(0.0481)	-		-0.0735	(0.0202)	-	
latent (ρ_{c23})	0.0009	(0.0258)	-		0.0288	(0.0326)	-	
Group Fixed Effect	Yes		Yes		Yes		Yes	
$\sigma_{\epsilon_c}^{2(*)}$	0.4588	(0.1166)	0.4361	(0.1776)	0.4432	(0.1121)	0.4561	(0.1931)

Note: The full model contains the activity outcome equation with interactions and the network formation model, where the network formation model involves the latent characteristic variables and the incentive effect. In the second column, we remove the latent variables from the network formation model. In the third column, we remove the incentive effect from the network formation model. In the fourth column, we only estimate the activity outcome equation. The MCMC runs for 150,000 iterations and the first 50,000 runs are dropped for the burn-in. Values in parentheses are standard deviations of draws from MCMC. $\sigma_{\epsilon_c}^{2(*)}$ denote the average of $\sigma_{\epsilon_c}^2$ estimated from different groups and the value in the parenthesis is the average of standard deviations.

Table 3: Estimation result based on Smoking

	Full		Without Latent		Without Incentive		Activity Alone	
Network Formation								
Higher sender age (γ_1)	-0.0794	(0.0457)	-0.0621	(0.0467)	-0.0748	(0.0434)	-	
Higher receiver age (γ_2)	-0.0858	(0.0328)	-0.0616	(0.0457)	-0.0879	(0.0243)	-	
Constant (γ_{31})	-3.9726	(0.0341)	-4.2461	(0.0652)	-3.9671	(0.0295)	-	
Same age (γ_{32})	0.0356	(0.0379)	0.0463	(0.0417)	0.0425	(0.0328)	-	
Same sex (γ_{33})	0.3701	(0.0400)	0.3698	(0.0373)	0.3718	(0.0383)	-	
Same race (γ_{34})	0.3307	(0.0357)	0.3378	(0.0462)	0.3402	(0.0284)	-	
Diff. in smoking (γ_{35})	-0.0399	(0.0061)	-0.0358	(0.0077)	-0.0396	(0.0050)	-	
Latent dist. (γ_{41})	-0.1557	(0.0156)	-	-	-0.1567	(0.0142)	-	
Latent dist. (γ_{42})	-0.1331	(0.0126)	-	-	-0.1325	(0.0119)	-	
Latent dist. (γ_{43})	-0.1152	(0.0092)	-	-	-0.1156	(0.0087)	-	
Reciprocity (η_1)	1.3354	(0.0433)	1.3568	(0.0404)	1.3021	(0.0363)	-	
Expansiveness (η_2)	0.1932	(0.0167)	0.2186	(0.0211)	0.2053	(0.0187)	-	
Expansiveness (η_3)	-0.0228	(0.0017)	-0.0251	(0.0021)	-0.0240	(0.0018)	-	
Popularity (η_4)	-0.0290	(0.0054)	-0.0312	(0.0053)	-0.0324	(0.0045)	-	
Trans. triads (η_5)	0.6123	(0.0189)	0.6230	(0.0187)	0.6218	(0.0154)	-	
Three cycles (η_6)	-0.2613	(0.0161)	-0.2709	(0.0157)	-0.2625	(0.0138)	-	
Economic incentive (δ_t)	0.0041	(0.0027)	0.0046	(0.0027)	-	-	-	
Activity Outcome								
Endogenous (λ_t)	0.0534	(0.0071)	0.0598	(0.0067)	0.0535	(0.0072)	0.0620 (0.0085)	
Own (X)								
Age	0.2260	(0.0848)	0.3853	(0.0710)	0.2288	(0.0923)	0.4294 (0.0350)	
Male	0.3336	(0.1039)	0.2767	(0.0981)	0.3561	(0.1001)	0.2503 (0.0882)	
Black	-1.2733	(0.1673)	-1.3346	(0.2242)	-1.3025	(0.1459)	-1.8383 (0.1820)	
Asian	-1.4314	(0.2410)	-1.5606	(0.1515)	-1.4860	(0.1966)	-1.2254 (0.1837)	
Hispanic	0.0439	(0.2158)	-0.1062	(0.1873)	0.0370	(0.2167)	0.0786 (0.1517)	
Other race	0.1670	(0.2073)	0.0675	(0.1449)	0.1993	(0.1835)	0.4337 (0.1193)	
Both Parents	-0.2843	(0.1331)	-0.2532	(0.1306)	-0.3119	(0.1276)	-0.2185 (0.1091)	
Less HS	-0.1726	(0.1791)	-0.0531	(0.1827)	-0.1941	(0.1685)	-0.2750 (0.1267)	
More HS	0.0798	(0.1117)	0.1983	(0.1251)	0.0898	(0.1071)	-0.0526 (0.0789)	
Edu missing	-0.4051	(0.1124)	-0.0166	(0.2444)	-0.3698	(0.1015)	-0.4219 (0.1408)	
Welfare	0.9096	(0.3287)	1.5383	(0.1878)	0.8364	(0.2907)	1.6941 (0.1116)	
Job missing	0.0758	(0.1509)	0.3408	(0.1635)	0.0237	(0.1191)	0.2625 (0.1248)	
Professional	0.0618	(0.1177)	0.0746	(0.1180)	0.0657	(0.1080)	0.0934 (0.1711)	
Other Jobs	-0.0107	(0.1019)	0.0237	(0.1581)	-0.0086	(0.0935)	0.0260 (0.1781)	
Num. of students at home	0.0704	(0.0722)	0.0206	(0.0736)	0.0670	(0.0667)	0.0335 (0.0645)	
latent (ρ_{t11})	-0.0198	(0.0456)	-	-	-0.0211	(0.0416)	-	
latent (ρ_{t12})	-0.0033	(0.0440)	-	-	-0.0008	(0.0434)	-	
latent (ρ_{t13})	-0.0115	(0.0435)	-	-	-0.0167	(0.0438)	-	
Contextual (WX)								
Age	-0.0063	(0.0054)	-0.0166	(0.0051)	-0.0059	(0.0049)	-0.0155 (0.0062)	
Male	-0.2012	(0.0591)	-0.0673	(0.0614)	-0.2092	(0.0604)	-0.1655 (0.0655)	
Black	0.0983	(0.0473)	0.0967	(0.0557)	0.1058	(0.0412)	0.2349 (0.0460)	
Asian	-0.1615	(0.2596)	-0.7789	(0.3334)	-0.2050	(0.2474)	-1.0356 (0.2677)	

Continued on Next Page

Table – Continued

Hispanic	0.1733	(0.1097)	0.1831	(0.1481)	0.1926	(0.1060)	0.1688	(0.1225)
Other race	0.0438	(0.1163)	-0.1074	(0.1291)	0.0685	(0.1021)	0.1648	(0.1002)
Both Parents	-0.0645	(0.0758)	0.0060	(0.0773)	-0.0714	(0.0658)	0.1145	(0.0788)
Less HS	-0.0236	(0.1151)	0.0759	(0.1195)	-0.0330	(0.0984)	0.0646	(0.0954)
More HS	0.1285	(0.0590)	0.2365	(0.0669)	0.1213	(0.0579)	0.2133	(0.0640)
Edu missing	0.2649	(0.1283)	0.4247	(0.1154)	0.2646	(0.1312)	0.4125	(0.1342)
Welfare	-0.1205	(0.2371)	0.4216	(0.2771)	-0.1087	(0.2261)	0.2538	(0.0802)
Job missing	0.0106	(0.1189)	0.0011	(0.1412)	0.0075	(0.1156)	-0.1563	(0.0948)
Professional	-0.0733	(0.0770)	-0.1230	(0.0940)	-0.0657	(0.0771)	-0.2740	(0.0645)
Other Jobs	-0.1072	(0.0776)	-0.0633	(0.0817)	-0.0912	(0.0830)	-0.1556	(0.0554)
Num. of students at home	0.0098	(0.0445)	-0.0098	(0.0407)	0.0073	(0.0400)	-0.0243	(0.0440)
latent (ρ_{t21})	0.0018	(0.0260)	-		-0.0014	(0.0230)	-	
latent (ρ_{t22})	0.0014	(0.0241)	-		0.0027	(0.0248)	-	
latent (ρ_{t23})	0.0028	(0.0245)	-		-0.0038	(0.0242)	-	
Group Fixed Effect	Yes		Yes		Yes		Yes	
$\sigma_{\epsilon_t}^{2(*)}$	7.1316	(3.6623)	7.1349	(3.5563)	7.1211	(3.6533)	7.0322	(3.5196)

Note: The full model contains the activity outcome equation with interactions and the network formation model, where the network formation model involves the latent characteristic variables and the incentive effect. In the second column, we remove the latent variables from the network formation model. In the third column, we remove the incentive effect from the network formation model. In the fourth column, we only estimate the activity outcome equation. The MCMC runs for 150,000 iterations and the first 50,000 runs are dropped for the burn-in. Values in parentheses are standard deviations of draws from MCMC. $\sigma_{\epsilon_t}^{2(*)}$ denote the average of $\sigma_{\epsilon_t}^2$ estimated from different groups and the value in the parenthesis is the average of standard deviations.

Table 4: Estimation result based on both GPA and Smoking

	Full		Without Latent		Without Incentive		Activity Alone	
Network Formation								
Higher sender age (γ_1)	-0.0583	(0.0440)	-0.0143	(0.0512)	-0.0512	(0.0282)	-	
Higher receiver age (γ_2)	0.0142	(0.0314)	0.0807	(0.0625)	0.0054	(0.0385)	-	
Constant (γ_{31})	-3.2972	(0.0729)	-4.5936	(0.1105)	-3.2913	(0.0274)	-	
Same age (γ_{32})	0.1564	(0.0508)	0.1350	(0.0470)	0.1874	(0.0464)	-	
Same sex (γ_{33})	0.3563	(0.0444)	0.3598	(0.0458)	0.3836	(0.0358)	-	
Same race (γ_{34})	0.6587	(0.0362)	0.6391	(0.0608)	0.6470	(0.0361)	-	
Diff. in smoking (γ_{35})	-0.0452	(0.0108)	-0.0351	(0.0088)	-0.0406	(0.0078)	-	
Latent dist. (γ_{41})	-0.3865	(0.0551)	-		-0.4448	(0.0387)	-	
Latent dist. (γ_{42})	-0.3279	(0.0561)	-		-0.3690	(0.0381)	-	
Latent dist. (γ_{43})	-0.3110	(0.0420)	-		-0.2800	(0.0588)	-	
Reciprocity (η_1)	1.2032	(0.0426)	1.3505	(0.0442)	1.2361	(0.0454)	-	
Expansiveness (η_2)	0.2051	(0.0217)	0.2430	(0.0244)	0.2021	(0.0118)	-	
Expansiveness (η_3)	-0.0237	(0.0023)	-0.0272	(0.0026)	-0.0241	(0.0015)	-	
Popularity (η_4)	-0.0458	(0.0082)	-0.0405	(0.0068)	-0.0387	(0.0056)	-	
Trans. triads (η_5)	0.5682	(0.0203)	0.5743	(0.0227)	0.5833	(0.0207)	-	
Three cycles (η_6)	-0.2201	(0.0170)	-0.2414	(0.0210)	-0.2269	(0.0185)	-	
Incentive from GPA (δ_c)	0.0028	(0.0024)	0.0023	(0.0013)	-		-	
Incentive from Smoking (δ_t)	0.7508	(0.0917)	1.0174	(0.2411)	-		-	
Activity Outcome								
Endogenous (λ_c)	0.0102	(0.0028)	0.0100	(0.0023)	0.0138	(0.0032)	0.0159	(0.0035)
Endogenous (λ_t)	0.0475	(0.0096)	0.0513	(0.0076)	0.0490	(0.0117)	0.0579	(0.0114)
GPA								
Own (X)								
Age	-0.1734	(0.0246)	-0.1711	(0.0290)	-0.1733	(0.0320)	-0.1659	(0.0107)
Male	-0.0964	(0.0429)	-0.0627	(0.0403)	-0.0765	(0.0299)	-0.0840	(0.0406)
Black	0.0727	(0.0383)	0.0338	(0.0670)	0.0298	(0.0590)	0.0420	(0.0614)
Asian	0.2132	(0.0951)	0.1732	(0.0726)	0.1407	(0.0631)	0.1533	(0.1153)
Hispanic	0.0883	(0.0638)	0.2272	(0.0831)	0.1520	(0.0676)	0.1388	(0.0996)
Other race	-0.0356	(0.0737)	-0.0592	(0.0722)	-0.0751	(0.0615)	-0.0407	(0.0952)
Both Parents	0.0446	(0.0415)	0.0625	(0.0413)	0.0302	(0.0378)	0.0345	(0.0486)
Less HS	-0.0660	(0.0638)	-0.0405	(0.0536)	-0.0922	(0.0600)	-0.1080	(0.0629)
More HS	0.0952	(0.0492)	0.1135	(0.0379)	0.1297	(0.0510)	0.1472	(0.0511)
Edu missing	0.0097	(0.0658)	-0.0844	(0.0544)	-0.0871	(0.0488)	-0.0802	(0.0755)
Welfare	0.0467	(0.1078)	-0.0149	(0.0735)	-0.0867	(0.0804)	-0.0245	(0.1188)
Job missing	-0.1914	(0.0609)	-0.1726	(0.0614)	-0.0843	(0.0534)	-0.0771	(0.0836)
Professional	0.0455	(0.0566)	0.0499	(0.0638)	0.0568	(0.0523)	0.0363	(0.0534)
Other Jobs	-0.0374	(0.0407)	0.0025	(0.0512)	-0.0127	(0.0467)	-0.0158	(0.0465)
Num. of students at home	-0.0133	(0.0273)	-0.0040	(0.0317)	-0.0070	(0.0285)	-0.0020	(0.0299)
latent (ρ_{c11})	0.0170	(0.0436)	-		-0.0366	(0.0832)	-	
latent (ρ_{c12})	0.0425	(0.0593)	-		-0.0337	(0.0779)	-	
latent (ρ_{c13})	0.0031	(0.0456)	-		-0.0066	(0.0636)	-	
Contextual (WX)								
Age	-0.0026	(0.0015)	-0.0013	(0.0014)	-0.0022	(0.0026)	-0.0036	(0.0026)
Male	-0.0347	(0.0162)	-0.0275	(0.0115)	-0.0065	(0.0265)	-0.0030	(0.0267)
Black	-0.0454	(0.0157)	-0.0352	(0.0120)	-0.0456	(0.0212)	-0.0376	(0.0198)
Asian	0.1219	(0.0540)	0.1145	(0.0952)	0.1514	(0.0878)	0.1636	(0.1153)

Continued on Next Page

Table – Continued

Hispanic	0.0280	(0.0255)	0.0348	(0.0276)	-0.0034	(0.0531)	-0.0012	(0.0605)
Other race	-0.0016	(0.0379)	0.0482	(0.0303)	-0.1117	(0.0606)	-0.0772	(0.0612)
Both Parents	0.0494	(0.0183)	0.0433	(0.0161)	0.0157	(0.0314)	0.0361	(0.0302)
Less HS	-0.0568	(0.0233)	-0.0438	(0.0199)	-0.0525	(0.0371)	-0.0487	(0.0455)
More HS	-0.0286	(0.0162)	-0.0232	(0.0126)	0.0159	(0.0351)	0.0234	(0.0337)
Edu missing	-0.0412	(0.0350)	-0.0427	(0.0271)	0.0269	(0.0608)	0.0320	(0.0578)
Welfare	-0.0991	(0.0401)	0.0424	(0.0548)	-0.1378	(0.1004)	-0.1075	(0.0830)
Job missing	-0.0340	(0.0291)	-0.0296	(0.0243)	0.0080	(0.0468)	-0.0024	(0.0464)
Professional	0.0173	(0.0258)	0.0148	(0.0160)	0.0276	(0.0398)	0.0207	(0.0418)
Other Jobs	0.0014	(0.0201)	0.0040	(0.0140)	0.0080	(0.0335)	0.0084	(0.0392)
Num. of students at home	0.0059	(0.0091)	-0.0002	(0.0076)	-0.0026	(0.0185)	-0.0023	(0.0184)
latent (ρ_{c21})	-0.0229	(0.0328)	-		-0.0064	(0.0348)	-	
latent (ρ_{c22})	-0.0506	(0.0218)	-		0.0011	(0.0250)	-	
latent (ρ_{c23})	0.0307	(0.0305)	-		0.0020	(0.0301)	-	
Smoking								
Own (X)								
Age	0.3641	(0.0401)	0.2377	(0.0857)	0.2143	(0.0580)	0.4047	(0.0561)
Male	0.1492	(0.1389)	0.1511	(0.1461)	0.1593	(0.1274)	0.0623	(0.1372)
Black	-1.7063	(0.1770)	-1.5487	(0.2821)	-2.0691	(0.2043)	-1.5954	(0.1716)
Asian	-1.4598	(0.3010)	-1.2536	(0.3758)	-1.5287	(0.2840)	-1.2423	(0.2593)
Hispanic	-0.8702	(0.2902)	-0.6784	(0.3572)	-0.6409	(0.1603)	-0.7279	(0.2111)
Other race	0.4511	(0.2432)	0.3531	(0.3237)	0.4825	(0.2015)	0.3955	(0.1511)
Both Parents	0.1191	(0.1342)	0.0309	(0.1696)	-0.0737	(0.1395)	0.0747	(0.1370)
Less HS	-0.1258	(0.2256)	0.0149	(0.2324)	-0.1363	(0.1611)	0.0950	(0.1831)
More HS	-0.1159	(0.1604)	-0.0690	(0.1806)	-0.2007	(0.1445)	-0.1125	(0.1664)
Edu missing	0.2666	(0.2154)	0.0443	(0.3024)	0.0983	(0.2287)	0.3707	(0.1904)
Welfare	0.9348	(0.3718)	0.8101	(0.3787)	1.2524	(0.2273)	1.3218	(0.2652)
Job missing	0.1638	(0.1971)	0.2669	(0.2278)	0.1838	(0.2106)	0.3462	(0.1885)
Professional	0.0008	(0.2035)	0.1032	(0.1975)	0.0449	(0.1513)	0.1776	(0.1714)
Other Jobs	-0.0284	(0.1639)	0.0650	(0.1891)	-0.0167	(0.1376)	0.1605	(0.1290)
Num. of students at home	-0.1177	(0.0876)	-0.1319	(0.0951)	-0.1540	(0.0983)	-0.1632	(0.0798)
latent (ρ_{t11})	-0.0751	(0.1712)	-		-0.0826	(0.1761)	-	
latent (ρ_{t12})	0.0429	(0.1257)	-		0.1133	(0.1756)	-	
latent (ρ_{t13})	-0.0199	(0.1618)	-		-0.0986	(0.1551)	-	
Contextual (WX)								
Age	-0.0234	(0.0088)	-0.0286	(0.0098)	-0.0295	(0.0099)	-0.0210	(0.0082)
Male	-0.0402	(0.0913)	0.0085	(0.1029)	0.0244	(0.0988)	-0.0143	(0.0821)
Black	0.1883	(0.0750)	0.1794	(0.0854)	0.2148	(0.0896)	0.1802	(0.0651)
Asian	-1.5693	(0.3854)	-1.4557	(0.3216)	-1.9430	(0.3103)	-1.5934	(0.3664)
Hispanic	0.1194	(0.1877)	0.1529	(0.2165)	0.1652	(0.1665)	0.1390	(0.1419)
Other race	-0.0665	(0.2261)	0.0153	(0.2515)	0.0630	(0.2010)	-0.0567	(0.1865)
Both Parents	-0.0925	(0.1137)	-0.0276	(0.1317)	-0.0778	(0.1252)	-0.0788	(0.0945)
Less HS	0.0404	(0.1596)	0.0367	(0.1772)	-0.0330	(0.1476)	-0.1256	(0.1428)
More HS	0.1414	(0.1261)	0.0997	(0.1163)	0.0159	(0.1241)	0.0699	(0.1039)
Edu missing	0.2520	(0.2223)	0.2548	(0.2206)	0.3726	(0.1756)	0.2434	(0.1740)
Welfare	0.2370	(0.2525)	0.3843	(0.5784)	0.3490	(0.2925)	0.1697	(0.1864)
Job missing	0.1301	(0.1545)	0.1551	(0.2274)	0.1848	(0.1498)	0.1964	(0.1819)
Professional	0.0213	(0.1527)	0.0525	(0.1627)	0.1390	(0.1267)	-0.0293	(0.1313)
Other Jobs	0.2417	(0.1244)	0.2234	(0.1327)	0.2748	(0.1197)	0.1722	(0.1122)

Continued on Next Page

Table – Continued

Num. of students at home	0.0124 (0.0601)	0.0151 (0.0636)	0.0026 (0.0636)	-0.0009 (0.0604)
latent (ρ_{t21})	-0.0006 (0.0899)	-	0.0726 (0.1125)	-
latent (ρ_{t22})	0.0514 (0.0825)	-	0.0752 (0.1100)	-
latent (ρ_{t23})	-0.0198 (0.0863)	-	0.0265 (0.1063)	-
Group Fixed Effect	Yes	Yes	Yes	Yes
$\sigma_{\epsilon_c}^{2(*)}$	0.7146 (0.3343)	0.7668 (0.3547)	0.7044 (0.3113)	0.7204 (0.3221)
$\sigma_{\epsilon_t}^{2(*)}$	9.6977 (5.2670)	9.9626 (5.4396)	9.7996 (5.7607)	9.1737 (4.6752)
$\sigma_{\epsilon_{tc}}^{(*)}$	-0.4980 (0.7091)	-0.5104 (0.6754)	-0.4986 (0.6626)	-0.4618 (0.6111)

Note: The full model contains the activity outcome equation with interactions and the network formation model, where the network formation model involves the latent characteristic variables and the incentive effect. In the second column, we remove the latent variables from the network formation model. In the third column, we remove the incentive effect from the network formation model. In the fourth column, we only estimate the activity outcome equation. The MCMC runs for 150,000 iterations and the first 50,000 runs are dropped for the burn-in. Values in parentheses are standard deviations of draws from MCMC. $\sigma_{\epsilon_c}^{2(*)}$, $\sigma_{\epsilon_t}^{2(*)}$, and $\sigma_{\epsilon_{tc}}^{(*)}$ denote the average of $\sigma_{\epsilon_c}^2$, $\sigma_{\epsilon_t}^2$, and $\sigma_{\epsilon_{tc}}$ estimated from different groups and the value in the parenthesis is the average of standard deviations.

5 Conclusion

Researchers are interested in network structures in order to analyze its impact on outcomes. As mentioned in Jackson (2010, section 5), if networks only serve as conduits for diffusion, e.g., diseases or ideas, their impact on outcomes is sort of mechanical and one need not worry about any feedback effects from outcomes. However, for studying the impact of a friendship network on outcomes, both the network structure and strategic interactions between the network and outcomes should be considered. This extra consideration should be reflected in a dynamic or static equilibrium model. In this paper, we propose a static equilibrium model which takes into account these features. The modeling approach in this paper is a complete information game, which assumes that students respond to incentives stemming from interaction benefits of activity outcomes with friends in making their friendship decisions in addition to unobserved individual characteristics in network formation and as factors for activities. The empirical results show that American high school students regard the interaction benefit from academic learning as a significant incentive for forming friendships, while the incentive effect of smoking together is not, even smoking behavior is important for the homophily effect. Another novelty of our approach to the social interaction literature is

to correct possible friendship selection biases in activity outcomes under interactions, which can be attributed to the specification of incentive effects, latent characteristic variables, or both.

Some issues that are not emphasized in this paper remain important for future extensions. The first is an alternative model of simultaneous non-cooperative network formation game and the problem of multiple game equilibria. We circumvent the latter problem in the present paper with a cooperative setting by assuming a benevolent social planner who manages overall network links and activities to maximize an aggregated utility, or individuals themselves coordinate their friendship formation processes and interactions for activities. Those assumptions in friendship formation may be appropriate for a school setting, but are questionable for some other circumstances. With multiple equilibria in games, one could either provide an equilibrium selection rule or characterize the estimation problem with moment inequalities. The second issue to consider is the dynamic evolution of networks and outcomes. The work of [Steglich et al. \(2010\)](#) provides an important reference in that direction. Last, we are interested in applying our modeling strategies to study the formation of other types of networks, e.g., criminal network, physician referral network, or academic co-author network, and economic activities in those networks in future research.

References

- Albert, James H. and Siddhartha Chib (1993) “Bayesian analysis of binary and polychotomous response data,” *Journal of the American Statistical Association*, Vol. 88, pp. 669–679.
- Amemiya, Takeshi (1974) “Multivariate regression and simultaneous equation models when the dependent variables are truncated normal,” *Econometrica*, Vol. 42, pp. 999–1012.
- Badev, Anton (2013) “Discrete games in endogenous networks: Theory and policy,” *working paper*.
- Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou (2006) “Who’s who in networks. wanted: the key player,” *Econometrica*, Vol. 74, pp. 1403–1417.

- Ballester, Coralio, Yves Zenou, and Antoni Calvó-Armengol (2010) “Delinquent networks,” *Journal of the European Economic Association*, Vol. 8, pp. 34–61.
- Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson (2013) “The diffusion of microfinance,” *Science*, Vol. 341, p. 1236498.
- Bartz, Kevin, Jun Liu, and Joseph Blitzstein (2009) “Monte Carlo maximum likelihood for exponential random graph models: From snowballs to umbrella densities,” *working paper*.
- Bayer, Patrick, Randi Hjalmarrsson, and David Pozen (2009) “Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections,” *The Quarterly Journal of Economics*, Vol. 124, pp. 105–147.
- Bayer, Patrick, Stephen L. Ross, and Giorgio Topa (2008) “Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes,” *Journal of Political Economy*, Vol. 116, pp. 1150–1196.
- Berger, James (1985) *Statistical Decision Theory and Bayesian Analysis*: New York: Springer-Verlag.
- Besag, Julian (1974) “Spatial interaction and the statistical analysis of lattice systems,” *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 36, pp. 192–236.
- Bhamidi, Shankar, Guy Bresler, Allan Sly et al. (2011) “Mixing time of exponential random graphs,” *The Annals of Applied Probability*, Vol. 21, pp. 2146–2170.
- 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.
- Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2009) “Identification of peer effects through social networks,” *Journal of Econometrics*, Vol. 150, pp. 41–55.
- Bramoullé, Yann and Bernard Fortin (2009) “The econometrics of social networks,” *Cahier de recherche/Working Paper*, Vol. 9, p. 13.

- Caimo, Alberto and Nial Friel (2011) “Bayesian inference for exponential random graph models,” *Social Networks*, Vol. 33, pp. 41–55.
- Calvo-Armengol, Antoni and Matthew O. Jackson (2004) “The effects of social networks on employment and inequality,” *The American Economic Review*, Vol. 94, pp. 426–454.
- Calvó-Armengol, Antoni and Matthew O. Jackson (2007) “Networks in labor markets: Wage and employment dynamics and inequality,” *Journal of Economic Theory*, Vol. 132, pp. 27–46.
- 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.
- Chandrasekhar, Arun G and Matthew O Jackson (2014) “Tractable and consistent random graph models,” *NBER working paper*.
- Chatterjee, Sourav, Persi Diaconis et al. (2013) “Estimating and understanding exponential random graph models,” *The Annals of Statistics*, Vol. 41, pp. 2428–2461.
- Chib, Siddhartha and Edward Greenberg (1996) “Markov chain Monte Carlo simulation methods in econometrics,” *Econometric Theory*, Vol. 12, pp. 409–431.
- Christakis, Nicholas A. and James H. Fowler (2007) “The spread of obesity in a large social network over 32 years,” *New England Journal of Medicine*, Vol. 357, pp. 370–379.
- Christakis, Nicholas A., James H. Fowler, Guido W. Imbens, and Karthik Kalyanaraman (2010) “An empirical model for strategic network formation,” *NBER working paper*.
- Clark, Andrew E and Youenn Lohéac (2007) “It wasn't me, it was them! Social influence in risky behavior by adolescents,” *Journal of Health Economics*, Vol. 26, pp. 763–784.
- Comola, Margherita (2007) “The network structure of informal arrangements: evidence from rural Tanzania,” *Barcelona, Universitat Pompeu Fabra, mimeo*.
- Conley, Timothy G. and Christopher R. Udry (2010) “Learning about a new technology: Pineapple in Ghana,” *The American Economic Review*, Vol. 100, pp. 35–69.

- Conley, Timothy and Christopher Udry (2001) “Social learning through networks: The adoption of new agricultural technologies in Ghana,” *American Journal of Agricultural Economics*, Vol. 83, pp. 668–673.
- Drèze, J. H. (1974) “Bayesian theory of identification in simultaneous equations models,” in S. E. Fienberg and A. Zellner eds. *Studies in Bayesian econometrics and statistics*, Vol. 1: North-Holland Publishing Co.
- 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,” *The American economic review*, Vol. 97, pp. 75–79.
- Fehr, Ernst and Klaus M Schmidt (1999) “A theory of fairness, competition, and cooperation,” *Quarterly Journal of Economics*, Vol. 114, pp. 817–868.
- Flashman, Jennifer (2012) “Academic achievement and its impact on friend dynamics,” *Sociology of Education*, Vol. 85, pp. 61–80.
- Fletcher, Jason M (2010) “Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects,” *Health Economics*, Vol. 19, pp. 466–484.
- Fosco, Constanza and Friederike Mengel (2011) “Cooperation through imitation and exclusion in networks,” *Journal of Economic Dynamics and Control*, Vol. 35, pp. 641–658.
- Fowler, James H. and Nicholas A. Christakis (2008) “Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; Trogdon, Nonnemaker, Pais,” *Journal of Health Economics*, Vol. 27, pp. 1400–1405.
- Frank, Ove and David Strauss (1986) “Markov graphs,” *Journal of the American Statistical Association*, Vol. 81, pp. 832–842.
- Fruehwirth, Jane Cooley (2014) “Can achievement peer effect estimates inform policy? A view from inside the black box,” *Review of Economics and Statistics*, Vol. 96, pp. 514–523.

- Fu, Feng, Christoph Hauert, Martin A Nowak, and Long Wang (2008) “Reputation-based partner choice promotes cooperation in social networks,” *Physical Review E*, Vol. 78, p. 026117.
- Gaviria, Alejandro and Steven Raphael (2001) “School-based peer effects and juvenile behavior,” *Review of Economics and Statistics*, Vol. 83, pp. 257–268.
- Geweke, John (1992) “Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments,” in *Bayesian Statistics*, Citeseer.
- Geyer, Charles J (1991) “Markov chain Monte Carlo maximum likelihood,” in *Computer Science and Statistics: Proceedings of the 23rd symposium on the Interface*: Interface Foundation of North America.
- Geyer, Charles J and Elizabeth A Thompson (1992) “Constrained Monte Carlo maximum likelihood for dependent data,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 657–699.
- 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. (2014) “An econometric model of link formation with degree heterogeneity,” *conditionally accepted by Econometrica*.
- Handcock, Mark S., Adrian E. Raftery, and Jeremy M. Tantrum (2007) “Model-based clustering for social networks,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 170, pp. 301–354.
- Hanushek, Eric A, John F Kain, Jacob M Markman, and Steven G Rivkin (2003) “Does peer ability affect student achievement?” *Journal of Applied Econometrics*, Vol. 18, pp. 527–544.
- Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock (2002) “Latent space approaches to social network analysis,” *Journal of the American Statistical Association*, Vol. 97, pp. 1090–1098.

- Hoxby, Caroline (2000) “Peer effects in the classroom: Learning from gender and race variation.”
- Hruschka, Daniel J and Joseph Henrich (2006) “Friendship, cliquishness, and the emergence of cooperation,” *Journal of Theoretical Biology*, Vol. 239, pp. 1–15.
- 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 and Hans Van Kippersluis (2016) “Smoking Initiation: Peers and Personality,” *CESR-Schaeffer Working Paper*.
- Jackson, Matthew O. (2010) “An overview of social networks and economic applications,” *The Handbook of Social Economics*, Vol. 1, pp. 511–85.
- Johnsson, Ida and Hyungsik Roger Moon (2016) “Estimation of peer effects in endogenous social networks: control function approach,” *working paper*.
- Kelejian, Harry H and Ingmar R Prucha (2010) “Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances,” *Journal of Econometrics*, Vol. 157, pp. 53–67.
- Kiuru, Noona, William J Burk, Brett Laursen, Katariina Salmela-Aro, and Jari-Erik Nurmi (2010) “Pressure to drink but not to smoke: Disentangling selection and socialization in adolescent peer networks and peer groups,” *Journal of Adolescence*, Vol. 33, pp. 801–812.
- Koop, Gary, M Hashem Pesaran, and Ron P Smith (2013) “On identification of Bayesian DSGE models,” *Journal of Business & Economic Statistics*, Vol. 31, pp. 300–314.
- Kovářík, Jaromír and Marco J van der Leij (2014) “Risk aversion and social networks,” *Review of Network Economics*, Vol. 13, pp. 121–155.
- Krauth, Brian V (2006) “Simulation-based estimation of peer effects,” *Journal of Econometrics*, Vol. 133, pp. 243–271.

- Lancaster, Tony (2004) *An introduction to modern Bayesian econometrics*: Blackwell Oxford.
- 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.
- Liang, Faming (2010) “A double Metropolis–Hastings sampler for spatial models with intractable normalizing constants,” *Journal of Statistical Computation and Simulation*, Vol. 80, pp. 1007–1022.
- Liang, Faming, Ick Hoon Jin, Qifan Song, and Jun S Liu (2016) “An adaptive exchange algorithm for sampling from distributions with intractable normalizing constants,” *Journal of the American Statistical Association*, Vol. 111, pp. 377–393.
- 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, Eleonora Patacchini, and Yves Zenou (2014) “Endogenous peer effects: local aggregate or local average?” *Journal of Economic Behavior & Organization*, Vol. 103, pp. 39–59.
- Lomi, Alessandro, Tom AB Snijders, Christian EG Steglich, and Vanina Jasmine Torló (2011) “Why are some more peer than others? Evidence from a longitudinal study of social networks and individual academic performance,” *Social Science Research*, Vol. 40, pp. 1506–1520.
- Lundborg, Petter (2006) “Having the wrong friends? Peer effects in adolescent substance use,” *Journal of Health Economics*, Vol. 25, pp. 214–233.
- Majolo, Bonaventura, Kaye Ames, Rachel Brumpton, Rebecca Garratt, Kate Hall, and Natasha Wilson (2006) “Human friendship favours cooperation in the Iterated Prisoner’s Dilemma,” *Behaviour*, Vol. 143, pp. 1383–1395.

- Manski, Charles F (1993) “Identification of endogenous social effects: The reflection problem,” *The Review of Economic Studies*, Vol. 60, pp. 531–542.
- Mas-Colell, Andreu, Michael Dennis Whinston, Jerry R Green et al. (1995) *Microeconomic theory*, Vol. 1: Oxford university press New York.
- Mele, Angelo (2016) “A structural model of Dense Network Formation,” *conditionally accepted in Econometrica*.
- Moffitt, Robert A et al. (2001) “Policy interventions, low-level equilibria, and social interactions,” *Social Dynamics*, Vol. 4, pp. 6–17.
- Mundlak, Yair (1978) “On the pooling of time series and cross section data,” *Econometrica*, Vol. 46, pp. 69–85.
- Murray, Iain, Zoubin Ghahramani, and David MacKay (2006) “MCMC for doubly-intractable distributions,” in *Proceedings of the 22nd Annual Conference on Uncertainty in Artificial Intelligence: UAI*.
- Patacchini, Eleonora and Yves Zenou (2008) “The strength of weak ties in crime,” *European Economic Review*, Vol. 52, pp. 209–236.
- (2012) “Juvenile delinquency and conformism,” *Journal of Law, Economics, and Organization*, Vol. 28, pp. 1–31.
- Peck, Joel R (1993) “Friendship and the evolution of co-operation,” *Journal of Theoretical Biology*, Vol. 162, pp. 195–228.
- Poirier, Dale J (1998) “Revising beliefs in nonidentified models,” *Econometric Theory*, Vol. 14, pp. 483–509.
- Powell, Lisa M, John A Tauras, and Hana Ross (2005) “The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior,” *Journal of Health Economics*, Vol. 24, pp. 950–968.

- Propp, James Gary and David Bruce Wilson (1996) “Exact sampling with coupled Markov chains and applications to statistical mechanics,” *Random structures and Algorithms*, Vol. 9, pp. 223–252.
- Rendon, Silvio R (2013) “Fixed and random effects in classical and Bayesian regression,” *Oxford Bulletin of Economics and Statistics*, Vol. 75, pp. 460–476.
- Sacerdote, Bruce (2001) “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” *The Quarterly Journal of Economics*, Vol. 116, pp. 681–704.
- Schaefer, David R, Steven A Haas et al. (2013) “Social networks and smoking exploring the effects of peer influence and smoker popularity through simulations,” *Health Education & Behavior*, Vol. 40, pp. 24S–32S.
- Sheng, Shuyang (2014) “A structural econometric analysis of network formation games,” *working paper*.
- Smith, Tony E and James P LeSage (2004) “A Bayesian probit model with spatial dependencies,” *Advances in Econometrics*, Vol. 18, pp. 127–160.
- Snijders, Tom A., Gerhard G. Van de Bunt, and Christian E. Steglich (2010) “Introduction to stochastic actor-based models for network dynamics,” *Social Networks*, Vol. 32, pp. 44–60.
- Snijders, Tom AB (2001) “The statistical evaluation of social network dynamics,” *Sociological Methodology*, Vol. 31, pp. 361–395.
- (2002) “Markov chain Monte Carlo estimation of exponential random graph models,” *Journal of Social Structure*, Vol. 3, pp. 1–40.
- Snijders, Tom AB, Philippa E. Pattison, Garry L. Robins, and Mark S. Handcock (2006) “New specifications for exponential random graph models,” *Sociological Methodology*, Vol. 36, pp. 99–153.

- Soetevent, Adriaan R and Peter Kooreman (2007) “A discrete-choice model with social interactions: with an application to high school teen behavior,” *Journal of Applied Econometrics*, Vol. 22, pp. 599–624.
- Steglich, Christian, Tom AB Snijders, and Michael Pearson (2010) “Dynamic networks and behavior: Separating selection from influence,” *Sociological Methodology*, Vol. 40, pp. 329–393.
- Strauss, David and Michael Ikeda (1990) “Pseudolikelihood estimation for social networks,” *Journal of the American Statistical Association*, Vol. 85, pp. 204–212.
- Tierney, Luke (1994) “Markov chains for exploring posterior distributions,” *The Annals of Statistics*, pp. 1701–1728.
- Udry, J Richard (2003) “The national longitudinal study of adolescent health (Add Health), waves I and II, 1994–1996; wave III, 2001–2002 [machine-readable data file and documentation],” *Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill*.
- Ule, Aljaz (2008) *Partner choice and cooperation in networks: Theory and experimental evidence*, Vol. 598: Springer Science & Business Media.
- 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.
- Weinberg, Bruce A. (2007) “Social interactions with endogenous associations,” *NBER working paper*.
- Zimmerman, David J (2003) “Peer effects in academic outcomes: Evidence from a natural experiment,” *Review of Economics and Statistics*, Vol. 85, pp. 9–23.

Appendix

A Tobit-type activity interactions

In certain cases, activity variables might be continuous, but subject to non-negative constraints. The observed outcomes may mix with zeros and positive continuous quantities. This is a Tobit-type variable left-censored at the value zero. To distinguish a Tobit-type activity variable from a continuous one, we replace $y_{i,cg}$ with $y_{i,tg}$ and impose the constraint, $y_{i,tg} \geq 0$, on the individual utility of Eq. (1). Under this constraint, the sub-game cooperative equilibrium activity vector for a given W_g will be solved from maximization of Eq. (2) subject to non-negative constraints, which is characterized as:

$$\begin{aligned} Y_{tg} &= \max \left(0, \ddot{Y}_{tg} \right) \\ \ddot{Y}_{tg} &= \lambda_t(W_g + W'_g)Y_{tg} + X_g\beta_{1t} + W_g X_g \beta_{2t} + l_g \alpha_{tg} + \epsilon_{tg}, \end{aligned} \quad (\text{A.1})$$

where \ddot{Y}_{tg} represents a vector of latent variables. The solution Y_{tg} must satisfy $Y_{tg} \geq \lambda_t(W_g + W'_g)Y_{tg} + X_g\beta_{1t} + W_g X_g \beta_{2t} + l_g \alpha_{tg} + \epsilon_{tg}$, $Y_{tg} \geq 0$, and $y_{i,tg} = \lambda_t \sum_{j=1}^{m_g} (w_{ij,g} + w_{ji,g})y_{j,tg} + x_{i,g}\beta_{1t} + \sum_{j=1}^{m_g} w_{ij,g}x_{j,g}\beta_{2t} + \alpha_{tg} + \epsilon_{i,tg}$ whenever $y_{i,tg} > 0$ for each i , $i = 1, \dots, m_g$. Under the conditions as in Amemiya (1974) for a general simultaneous Tobit equation system, the solution Y_{tg} is unique and can be obtained from a constrained optimization problem

$$\max_{y_{i,tg}, i=1, \dots, n} \sum_{i=1}^{m_g} u_{i,tg}(y_{i,tg}, Y_{-i,tg}, W_g) \quad \text{subject to } Y_{tg} \geq 0. \quad (\text{A.2})$$

For our case, the unique solution follows from that the aggregated utility function is a strictly concave function of Y_{tg} on a convex set. Furthermore, with a properly restricted parameter space on λ_t , namely, $\|\lambda_t(W_g + W'_g)\|_\infty < 1$, where $\|\cdot\|_\infty$ is the maximum row sum norm, we can show that the solution Y_{tg} can be conveniently obtained via a contraction mapping algorithm and Y_{tg} is stable. Denote $a \vee 0 = \max\{a, 0\}$ for a scalar a . Consider a mapping $F : R_+^{m_g} \rightarrow R_+^{m_g}$, where $R_+^{m_g} = \{Y_{tg} : Y_{tg} \in R^{m_g}, Y_{tg} \geq 0\}$, defined by

$$F(Y_{tg}) = (\lambda_t(W_g + W'_g)Y_{tg} + H_g) \vee 0 = \begin{pmatrix} (\lambda_t(w_{1..g} + w'_{.1,g})Y_{tg} + H_{1,g}) \vee 0 \\ \vdots \\ (\lambda_t(w_{m_g..g} + w'_{.m_g,g})Y_{tg} + H_{m_g,g}) \vee 0 \end{pmatrix},$$

where $H_g = X_g\beta_{1t} + W_gX_g\beta_{2t} + l_g\alpha_{tg} + \epsilon_{tg}$, $w_{i..g}$ is the i^{th} row of W_g , $w_{.i,g}$ is the i^{th} column of W_g , and $H_{i,g}$ is the i th row of H_g . For any \widetilde{Y}_{tg} and $\widetilde{\widetilde{Y}}_{tg}$ in $R_+^{m_g}$,

$$\begin{aligned}
& \| F(\widetilde{\widetilde{Y}}_{tg}) - F(\widetilde{Y}_{tg}) \|_\infty \\
&= \| ((\lambda_t(W_g + W'_g)\widetilde{\widetilde{Y}}_{tg} + H_g) \vee 0) - ((\lambda_t(W_g + W'_g)\widetilde{Y}_{tg} + H_g) \vee 0) \|_\infty \\
&= \max_{i=1,\dots,m_g} \left| ((\lambda_t(w_{i..g} + w'_{.i,g})\widetilde{\widetilde{Y}}_{tg} + H_{i,g}) \vee 0) - ((\lambda_t(w_{i..g} + w'_{.i,g})\widetilde{Y}_{tg} + H_{i,g}) \vee 0) \right| \\
&\leq \max_{i=1,\dots,m_g} \left| \lambda_t(w_{i..g} + w'_{.i,g})(\widetilde{\widetilde{Y}}_{tg} - \widetilde{Y}_{tg}) \right| = \| \lambda_t(W_g + W'_g)(\widetilde{\widetilde{Y}}_{tg} - \widetilde{Y}_{tg}) \|_\infty \\
&\leq \| \lambda_t(W_g + W'_g) \|_\infty \cdot \| \widetilde{\widetilde{Y}}_{tg} - \widetilde{Y}_{tg} \|_\infty .
\end{aligned}$$

Thus, if $\| \lambda_t(W_g + W'_g) \|_\infty < 1$, $F(\cdot)$ is a contraction mapping. The assumption $\| \lambda_t(W_g + W'_g) \|_\infty < 1$ implies the SAR process is stable in the cross section dimension because the series expansion of the spatial transformation $(I_{m_g} - \lambda_t(W_g + W'_g))^{-1} = \sum_{l=0}^{\infty} \lambda_t^l(W_g + W'_g)^l$ exists. As $F(\cdot)$ is a contraction mapping, there exist a unique fixed point Y_{tg} such that $F(Y_{tg}) = Y_{tg}$. The observed Y_{tg} is the unique solution of this Tobit equation, because $Y_{tg} = F(Y_{tg}) = (\lambda_t(W_g + W'_g)Y_{tg} + H_g) \vee 0$, which gives $Y_{tg} \geq 0$, $Y_{tg} \geq \lambda(W_g + W'_g)Y_{tg} + H_g$ and $y_{i,tg} = \lambda_t(w_{i..g} + w'_{.i,g})Y_{tg} + H_{i,g}$ whenever $y_{i,tg} > 0$ for any i in the group g . This contraction mapping feature also suggests a simple iterative algorithm to solve for Y_{tg} given values of λ_t , W_g and H_g .

B Likelihood function for a full model with Tobit activity outcomes

For the Tobit-type activity variable Y_{tg} , we can divide the m_g individuals in group g into two blocks, such that the first m_{g1} individuals have activity variables equal to zero and the remaining individuals who are arranged from $m_{g1} + 1$ to m_g have positive values. Eq. (A.1)

of Y_{tg} and the network W_g can be conformably decomposed into

$$\begin{aligned} & \begin{pmatrix} \ddot{Y}_{tg1} \\ Y_{tg2} \end{pmatrix} = \\ & \lambda_t \begin{pmatrix} W_{11,g} + W'_{11,g} & W_{12,g} + W'_{21,g} \\ W_{21,g} + W'_{12,g} & W_{22,g} + W'_{22,g} \end{pmatrix} \begin{pmatrix} Y_{tg1} \\ Y_{tg2} \end{pmatrix} + \begin{pmatrix} X_{1g} \\ X_{2g} \end{pmatrix} \beta_{1t} + \begin{pmatrix} W_{11,g} & W_{12,g} \\ W_{21,g} & W_{22,g} \end{pmatrix} \begin{pmatrix} X_{1g} \\ X_{2g} \end{pmatrix} \beta_{2t} \\ & + \begin{pmatrix} Z_{1g} \\ Z_{2g} \end{pmatrix} \rho_{1t} + \begin{pmatrix} W_{11,g} & W_{12,g} \\ W_{21,g} & W_{22,g} \end{pmatrix} \begin{pmatrix} Z_{1g} \\ Z_{2g} \end{pmatrix} \rho_{2t} + \begin{pmatrix} l_{g1} \\ l_{g2} \end{pmatrix} \alpha_{tg} + \begin{pmatrix} \xi_{tg1} \\ \xi_{tg2} \end{pmatrix}, \quad (\text{B.1}) \end{aligned}$$

where $Y_{tg2} > 0$ and $Y_{tg1} = 0$ with the corresponding latent $\ddot{Y}_{tg1} \leq 0$. Based on Eq. (B.1), the joint probability function of Y_{tg} and W_g can be written as

$$\begin{aligned} & P(Y_{tg}, W_g | \theta_{tg}, \alpha_{tg}, Z_g) \\ & = P(Y_{tg1} = 0, Y_{tg2}, W_g | \theta_{tg}, \alpha_{tg}, Z_g) \\ & = \int I(Y_{tg1} = 0, \ddot{Y}_{tg1}) \cdot P(\ddot{Y}_{tg1}, Y_{tg2}, W_g | \theta_{tg}, \alpha_{tg}, Z_g) \cdot d\ddot{Y}_{tg1} \\ & = \int_{-\infty}^{-\left(\lambda_t(W_{12,g} + W'_{21,g})Y_{tg2} + X_{1g}\beta_{1t} + (W_{11,g}X_{1g} + W_{12,g}X_{2g})\beta_{2t} + Z_{1g}\rho_{1t} + (W_{11,g}Z_{1g} + W_{12,g}Z_{2g})\rho_{2t} + l_{g1}\alpha_{tg}\right)} \\ & \quad \left| I_{m_g - m_{g1}} - \lambda_t(W_{22,g} + W'_{22,g}) \right| \cdot f(\xi_{tg1}, \xi_{tg2} | \theta_{tg}, \alpha_{tg}, Z_g) \cdot P(W_g | \xi_{tg1}, \xi_{tg2}, \theta_{tg}, \alpha_{tg}, Z_g) \cdot d\xi_{tg1}, \quad (\text{B.2}) \end{aligned}$$

where $I(Y_{tg1} = 0, \ddot{Y}_{tg1})$ is a dichotomous indicator which is equal to 1 when \ddot{Y}_{tg1} is negative, and equal to 0, otherwise; $\xi_{tg2} = (I_{m_g - m_{g1}} - \lambda_t(W_{22,g} + W'_{22,g}))Y_{tg2} - X_{2g}\beta_{1t} - (W_{21,g}X_{1g} + W_{22,g}X_{2g})\beta_{2t} - Z_{2g}\rho_{1t} - (W_{21,g}Z_{1g} + W_{22,g}Z_{2g})\rho_{2t} - l_{2g}\alpha_{tg}$ and $\theta_{tg} = (\gamma', \eta', \delta_t, \lambda_t, \beta'_t, \rho'_t, \sigma_{\xi_{tg}}^2)$.

Incentive effects can be from a total of \bar{d} activity variables which consist of mixed continuous and Tobit-type ones. For simplicity, we consider a model with one continuous and one Tobit-type activity variable, where the disturbances, $\xi_{i,tg}$ and $\xi_{i,cg}$, follow a bivariate normal distribution,

$$(\xi_{i,tg}, \xi_{i,cg}) \sim i.i.d. \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\xi_{tg}}^2 & \sigma_{\xi_{tcg}} \\ \sigma_{\xi_{ctg}} & \sigma_{\xi_{cg}}^2 \end{pmatrix} \right), \quad i = 1, \dots, m_g. \quad (\text{B.3})$$

From Eq. (B.3), conditional on ξ_{cg} , one has

$$\xi_{tg} = \sigma_{\xi_{tcg}} \sigma_{\xi_{cg}}^{-2} \xi_{cg} + u_g, \quad u_g \sim \mathcal{N}_{m_g}(0, \sigma_{ug}^2 I_{m_g}), \quad (\text{B.4})$$

where u_g is independent of ξ_{cg} and $\sigma_{ug}^2 = (\sigma_{\xi_{tg}}^2 - \sigma_{\xi_{tcg}} \sigma_{\xi_{cg}}^{-2} \sigma_{\xi_{cg}})$.

Let $\theta_{ctg} = (\gamma', \eta', \delta_c, \delta_t, \lambda_c, \lambda_t, \beta'_c, \beta'_t, \rho'_c, \rho'_t, \sigma_{\xi_{cg}}^2, \sigma_{\xi_{tg}}^2, \sigma_{\xi_{tcg}})$, the joint probability function of Y_{tg} , Y_{cg} , and W_g is

$$\begin{aligned}
& P(Y_{tg}, Y_{cg}, W_g | \theta_{ctg}, \alpha_{cg}, \alpha_{tg}, Z_g) \\
&= \int_{-\infty}^{-\left(\lambda_t(W_{12,g} + W'_{21,g})Y_{tg2} + X_{1g}\beta_{1t} + (W_{11,g}X_{1g} + W_{12,g}X_{2g})\beta_{2t} + Z_{1g}\rho_{1t} + (W_{11,g}Z_{1g} + W_{12,g}Z_{2g})\rho_{2t} + l_{g1}\alpha_{tg}\right)} \\
& \left| I_{m_g - m_{g1}} - \lambda_t(W_{22,g} + W'_{22,g}) \right| \cdot f(\xi_{tg} | \xi_{cg}, \theta_{ctg}, \alpha_{tg}, \alpha_{cg}, Z_g) \cdot |S_{cg}(W_g)| \cdot f(\xi_{cg} | \theta_{cg}, \alpha_{cg}, Z_g) \\
& \cdot P(W_g | \xi_{cg}, \xi_{tg}, \theta_{ctg}, \alpha_{tg}, \alpha_{cg}, Z_g) \cdot d\xi_{tg1}. \tag{B.5}
\end{aligned}$$

C Posterior analysis for a full model with continuous and Tobit-type activities

Here we discuss the posterior distribution of parameters in the model with one continuous and one Tobit-type activity variables. By Bayes' theorem, the joint posterior distribution of the parameters and latent variables in the model is

$$\begin{aligned}
& P\left(\{\theta_{ctg}\}, \{\alpha_{cg}\}, \{\alpha_{tg}\}, \{Z_g\}, \{\ddot{Y}_{tg1}\} | \{Y_{cg}\}, \{Y_{tg}\}, \{W_g\}\right) \\
& \propto \pi(\{\theta_{ctg}\}, \{\alpha_{cg}\}, \{\alpha_{tg}\}, \{Z_g\}) \cdot \\
& \prod_{g=1}^G \left\{ \left(\prod_{i=1}^{m_{g1}} I(y_{i,tg} = 0) \cdot I(\ddot{y}_{i,tg} \leq 0) \right) \cdot P\left(Y_{tg}, Y_{cg}, W_g, \ddot{Y}_{tg1} | \theta_{ctg}, \alpha_{cg}, \alpha_{tg}, Z_g\right) \right\}, \tag{C.1}
\end{aligned}$$

where $\pi(\cdot)$ represents the prior density function. The prior distributions for parameters and latent variables in the model are defined as follows:

- (i) Unobserved individual latent variables in both network formation model and economic activity equation,

$$z_{i,g} \sim \mathcal{N}_{\bar{\ell}}(0, I_{\bar{\ell}}), \quad i = 1, \dots, m_g; \quad g = 1, \dots, G.$$

- (ii) Coefficients of network formation model,

$$\phi = (\gamma', \eta', \delta_c, \delta_t) \sim \mathcal{N}_{2\bar{s} + \bar{q} + \bar{\ell} + \bar{h} + 2}(\phi_0, \Phi_0 I_{2\bar{s} + \bar{q} + \bar{\ell} + \bar{h} + 2}), \quad \phi \in O.$$

- (iii) Endogenous interaction parameters in the economic activity equation,

$$\lambda_c \sim U[-1/\tau_G, 1/\tau_G] \text{ and } \lambda_t \sim U[-1/\tau_G, 1/\tau_G].$$

(iv) Coefficients of own and contextual effects in the economic activity equation,

$$\beta_c \sim \mathcal{N}_{2k}(\beta_0, B_0 I_{2k}) \text{ and } \beta_t \sim \mathcal{N}_{2k}(\beta_0, B_0 I_{2k}).$$

(v) Coefficients of own and contextual correlated effects in the economic activity equation,

$$\rho_c \sim \mathcal{N}_{2\bar{\ell}}(\rho_0, R_0 I_{2\bar{\ell}}) \text{ and } \rho_t \sim \mathcal{N}_{2\bar{\ell}}(\rho_0, R_0 I_{2\bar{\ell}}).$$

(vi) Variances and covariance of disturbance in economic activity equation,

$$\sigma_g = (\sigma_{\xi_{cg}}^2, \sigma_{\xi_{tg}}^2, \sigma_{\xi_{ctg}}) \sim \mathcal{N}_3(\sigma_0, \Sigma_0 I_3), \sigma_g \in T_g, \quad g = 1, \dots, G.$$

(vii) Group fixed effects in the economic activity equation,

$$\alpha_{cg} \sim \mathcal{N}(\alpha_0, A_0) \text{ and } \alpha_{tg} \sim \mathcal{N}(\alpha_0, A_0), \quad g = 1, \dots, G.$$

The above prior distributions are largely inherited from the ones in Section 3.2. The only difference here is that we put $\sigma_{\xi_{cg}}^2$, $\sigma_{\xi_{tg}}^2$, and $\sigma_{\xi_{ctg}}$ into a group, σ_g , and specify a multivariate normal distribution truncated to the area, $T_g = \{\sigma_g \in \mathbb{R}^3 | \sigma_{\xi_{cg}}^2 > 0, \sigma_{\xi_{tg}}^2 > 0, \sigma_{\xi_{cg}}^2 \sigma_{\xi_{tg}}^2 - \sigma_{\xi_{ctg}}^2 \geq 0\}$, so that $\sigma_{\xi_{cg}}^2$, $\sigma_{\xi_{tg}}^2$, and $\sigma_{\xi_{ctg}}$ form a proper covariance matrix. In the case of correlated random group effects, we change the prior settings for α_{cg} and α_{tg} in (vii). We specify α_{cg} and α_{tg} as follows:

$$\begin{aligned} \alpha_{cg} &= \bar{X}_g \beta_{3c} + \bar{Z}_g \rho_{3c} + \zeta_{cg}, \quad \zeta_{cg} \sim \mathcal{N}(0, \sigma_{\alpha,c}^2), \\ \alpha_{tg} &= \bar{X}_g \beta_{3t} + \bar{Z}_g \rho_{3t} + \zeta_{tg}, \quad \zeta_{tg} \sim \mathcal{N}(0, \sigma_{\alpha,t}^2), \end{aligned}$$

where \bar{X}_g and \bar{Z}_g are, respectively, the group averages of X_g and Z_g . β_{3c} , ρ_{3c} , β_{3t} , ρ_{3t} , $\sigma_{\alpha,c}^2$, and $\sigma_{\alpha,t}^2$ are unknown parameters, and we specify the following prior distributions for them: $\beta_{3c} \sim \mathcal{N}_k(\beta_0, B_0 I_k)$, $\rho_{3c} \sim \mathcal{N}_{\bar{\ell}}(\rho_0, R_0 I_{\bar{\ell}})$, $\beta_{3t} \sim \mathcal{N}_k(\beta_0, B_0 I_k)$, $\rho_{3t} \sim \mathcal{N}_{\bar{\ell}}(\rho_0, R_0 I_{\bar{\ell}})$, $\sigma_{\alpha,c}^2 \sim \mathcal{IG}(\frac{\kappa_0}{2}, \frac{\nu_0}{2})$, and $\sigma_{\alpha,t}^2 \sim \mathcal{IG}(\frac{\kappa_0}{2}, \frac{\nu_0}{2})$

D Simulation experiment for parameter identification

In Bayesian analysis, data are conditionally uninformative for nonidentified parameters given identified ones (Drèze, 1974; Poirier, 1998). Let a $K \times 1$ vector $\theta = [\theta'_1, \theta_2]'$ denote unknown model parameters, where θ_1 ($(K-1) \times 1$) are identifiable and θ_2 (1×1) may not. Proposition 2 of Poirier (1998) states that in that case data are always conditionally (on θ_1) uninformative

for θ_2 , i.e., $f(\theta_2|\theta_1, y) = f(\theta_2|\theta_1)$, where $f(\theta_2|\theta_1)$ denotes the probability density function of θ_2 given θ_1 . Applying this proposition to our model, if $f(\lambda|\theta_1, y)$ and $f(\delta|\theta_1, y)$ (conditioning on other identified parameters θ_1), are different from $f(\lambda|\theta_1)$ and $f(\delta|\theta_1)$, then λ and δ are identifiable. Furthermore, Bayesian analysis from stationary models (see, e.g., Berger (1985), p.224) shows that under certain regularity conditions, the Bayesian asymptotic theory relating to the posterior is identical to the asymptotic distribution theory for the maximum likelihood estimator (MLE). One of the regularity conditions is that the parameter is identified. Therefore, identification will be further assured when the posterior distribution is concentrated at the MLE in the large sample case where the role of the prior vanishes.

In the simulation, we generate artificial networks from the exponential distribution in Eq. (10) and also generate continuous and Tobit-type activity variables from the activity outcome equations in Eqs. (5) and (A.1). Given artificial data y , we simulate draws from the conditional posterior distributions $f(\lambda|\theta_1, y)$ and $f(\delta|\theta_1, y)$ with identifiable parameters (θ_1) fixed at their true values and compare them with draws from $f(\lambda|\theta_1)$ and $f(\delta|\theta_1)$, which would be just $f(\lambda)$ and $f(\delta)$ by assuming independence among priors. We implement the simulation with different sample sizes and inspect whether the conditional posterior plots of λ and δ will become more concentrated at their true parameter values when there is more information from sample.

In the data generating process (DGP), the network size is fixed at 30 and the number of groups varies from 20, 60, to 100. In order to enhance the notion that identification of λ and δ is valid generally for different DGPs, we examine the results based on two different sets of DGP parameters. In the network interaction models (for both continuous and Tobit-type variables), the exogenous variable X_i is generated from $N(0, 36)$. The group effect α is generated from $N(0, 0.5)$. The disturbance term ϵ_i is generated from $N(0, 0.5)$. For simplicity, there are no unobserved latent variable Z , and the contextual variable WX is not included.

In the network formation model, exogenous effects for each link w_{ij} are captured by a constant term and a dyad-specific exogenous variable C_{ij} which is generated as follows: first draw two uniform random variables from $U(0, 1)$, which are denoted as U_1 and U_2 . If U_1 and U_2 are both larger than 0.7 or less than 0.3, then we set C_{ij} to one. Otherwise, we set it to

zero. The parameters for exogenous and network structure effects are set as: $\gamma_{31} = -3.2$; $\gamma_{32} = 0.4$; $\eta_1 = 0.4$; $\eta_2 = 0.2$; $\eta_3 = -0.03$; $\eta_4 = 0.03$; $\eta_5 = 0.3$; $\eta_6 = -0.2$. We consider two sets of parameters for the endogenous effect in activities and the incentive effect in network formation. In the first set (DGP-I), $(\lambda_c, \lambda_t, \delta_c, \delta_t)$ equals $(0.05, 0.05, 0.3, 0.3)$. In the second set (DGP-II), $(\lambda_c, \lambda_t, \delta_c, \delta_t)$ equals $(0.02, 0.02, 0.6, 0.6)$. These parameters are chosen to control the average number of friends at around 2 in a group of 30 people. Each artificial network W is simulated by the M-H algorithm from an empty network based on $P(W|\theta)$ in Eq. (10). The following step is implemented iteratively (Snijders, 2002): For each entry of W_g , $w_{ij,g}$, $i \neq j$, in turn, we propose $\tilde{w}_{ij,g} = 1 - w_{ij,g}$. With the acceptance probability

$$\alpha(\tilde{w}_{ij}, w_{ij}) = \min \left\{ \frac{P(\tilde{w}_{ij}, W_{-ij}|\theta)}{P(w_{ij}, W_{-ij}|\theta)}, 1 \right\} = \min \left\{ \frac{\exp(V(\tilde{w}_{ij}, W_{-ij}, \theta))}{\exp(V(w_{ij}, W_{-ij}, \theta))}, 1 \right\},$$

updating w_{ij} to \tilde{w}_{ij} . Note that the denominators of $P(\tilde{w}_{ij}, W_{-ij}|\theta)$ and $P(w_{ij}, W_{-ij}|\theta)$ are cancelled out because the two probabilities are evaluated at the same θ . Activity variables are simulated along with networks. The M-H algorithm runs through the whole network for a total of 10,000 iterations and realizations of the network and the activity variables from the last iteration are used as the data.

The total of 100,000 draws are simulated for λ and δ from the conditional posterior distributions (given other parameters fixed at the true values) using the double M-H algorithm discussed in Section 3.3. Figures A.1 and A.2 show the results of continuous and Tobit cases based on, respectively, the DGP-I and the DGP-II. In each figure, panel (a) gives the continuous case and panel (b) gives the Tobit case. The true values of the parameters are marked in the horizontal axis. One can see that the conditional posterior distributions are clearly different from the prior distributions. When the sample size increases, the posterior distributions of both λ and δ become sharply concentrated around the true values. The findings provide the evidence that the two key parameters in our model are identified from the Bayesian point of view.

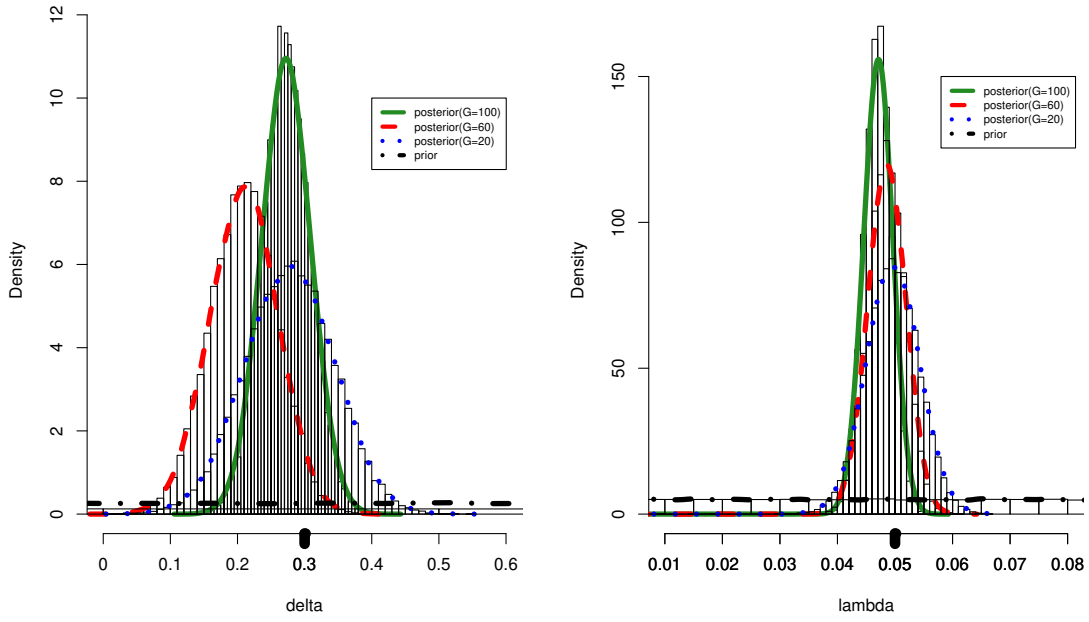
Finally, we use this simulation environment to illustrate the computational cost of our estimation algorithm. We take the data generating process for the continuous activity case from the DGP-I aforementioned and consider the different network sizes at $(20, 40, 60, 80, 100)$. The number of networks is less of a concern for computation because we can use parallel

computation to speed up. So we fix the number of networks at 20. We estimate the true model using the DMH algorithm and record the average CPU time for each MCMC iteration based on a machine with dual Xeon 2.60 GHz CPUs. The plot of computational time is in Figure A.3. We can see that the time increases exponentially with the network size.

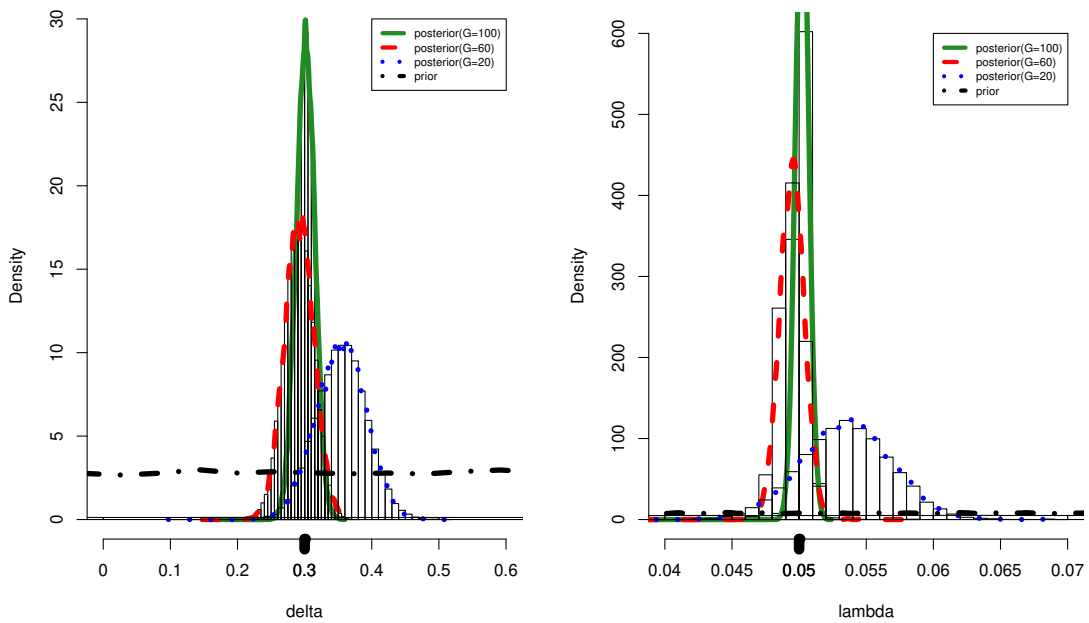
E Bayesian data imputation for missing dependent variables

The advantage of data implementation this is that we do not need to drop individuals with missing variables and could retrieve information provided by these missing observations. This can be important for observations of agents in a game if some agents play important roles in the game, e.g., star players. It is intuitive if stars were dropped, we might not be able to and obtain consistent and efficient estimates. The imputation of a missing independent variable can be much difficult if there are no specific stochastic processes assumed for their generation. However, missing dependent variables can be easily handled based on the specified stochastic processes. Let Y_{mis} and Y_{obs} denote, respectively, the missing and observed parts of Y . The Bayesian approach for missing data is implemented by the following two steps at every $(t + 1)^{th}$ MCMC sampling iteration: Imputation (I) step: $Y_{mis}^{(t+1)} \sim P(Y_{mis}|Y_{obs}, W, \theta^{(t)})$ and Posterior (P) step: $\theta^{(t+1)} \sim P(\theta|Y_{obs}, Y_{mis}^{(t+1)}, W)$. For the I step, since $P(Y_{mis}|Y_{obs}, W, \theta) = \frac{P(Y_{mis}, Y_{obs}, W|\theta)}{P(Y_{obs}, W|\theta)}$, it can be implemented by the double M-H algorithm in Section 3.3. Given Y_{mis} from the I step, the P step will follow the procedures described in Section 3.2.

F Appendix Tables and Figures

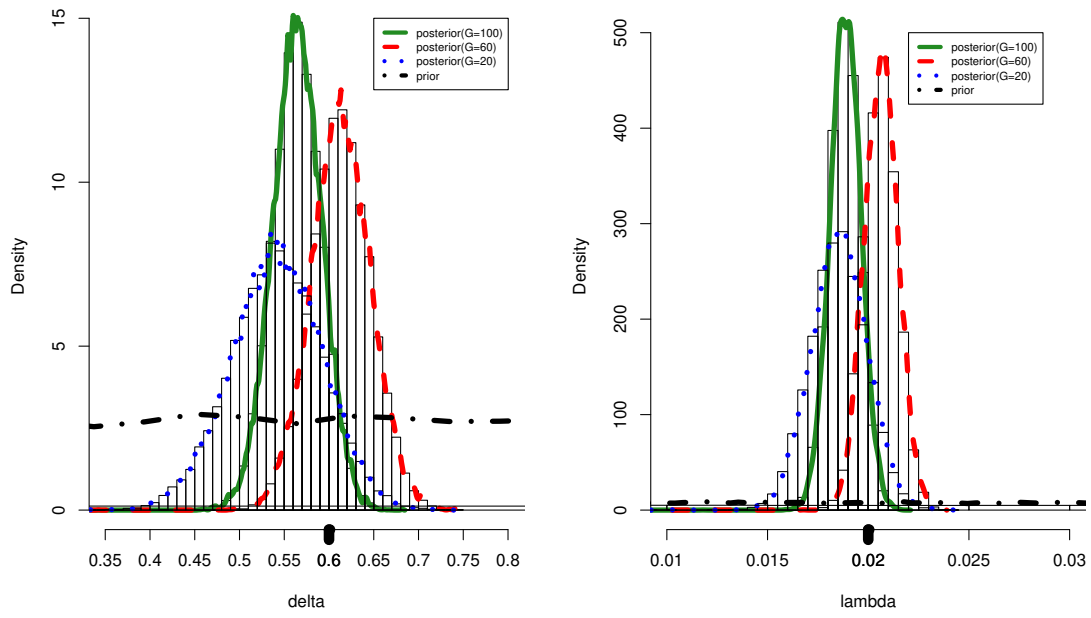


(a) Continuous case

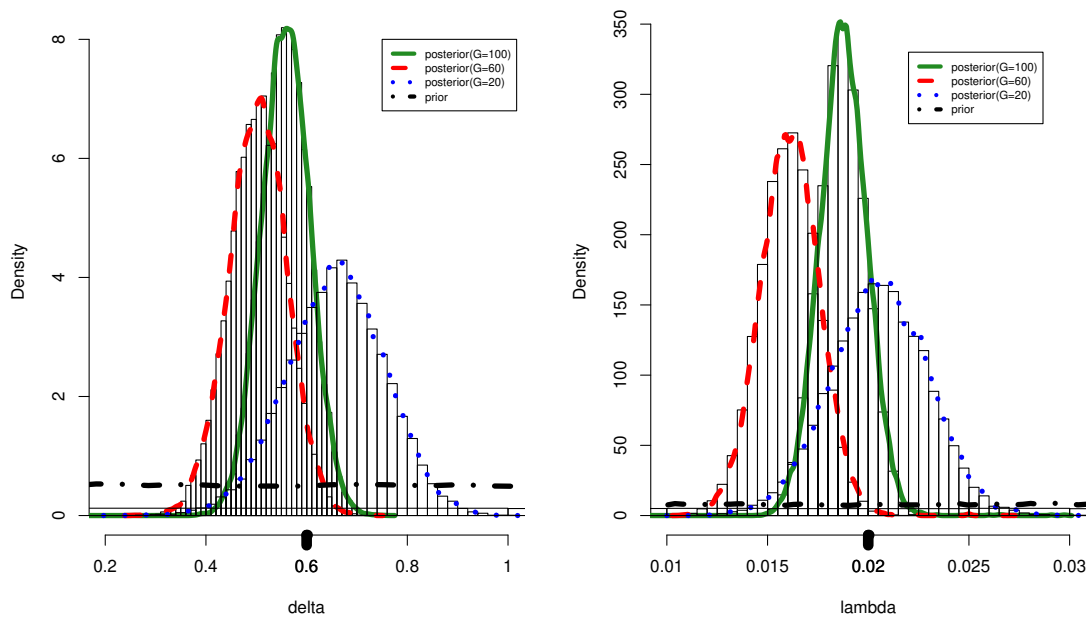


(b) Tobit case

Figure A.1: Simulation results based on DGP-I



(a) Continuous case



(b) Tobit case

Figure A.2: Simulation results based on DGP-II

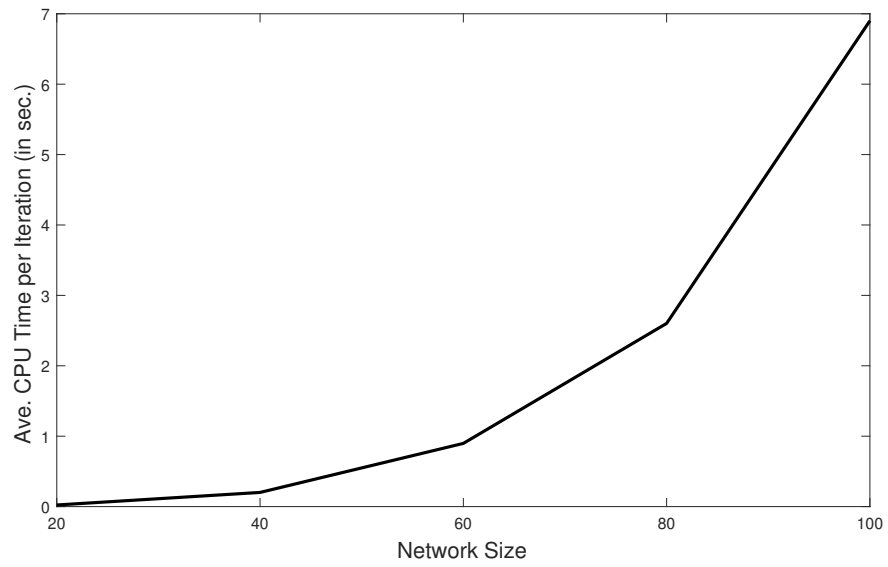


Figure A.3: The average computation time in seconds for a single MCMC iteration

Table A1: Estimation result with correlated random group effects

	Full		Without Latent		Without Incentive		Activity Alone	
Network Formation								
Higher sender age (γ_1)	0.0048	(0.0614)	-0.0020	(0.0478)	-0.0271	(0.0378)	-	
Higher receiver age (γ_2)	0.0365	(0.0455)	0.0929	(0.0408)	0.0199	(0.0401)	-	
Constant (γ_{31})	-3.2546	(0.0772)	-4.7160	(0.0389)	-3.3591	(0.2192)	-	
Same age (γ_{32})	0.1791	(0.0444)	0.1612	(0.0431)	0.2044	(0.0440)	-	
Same sex (γ_{33})	0.3394	(0.0291)	0.3692	(0.0474)	0.3780	(0.0392)	-	
Same race (γ_{34})	0.6582	(0.0401)	0.6320	(0.0467)	0.6635	(0.0503)	-	
Diff. in smoking (γ_{35})	-0.0486	(0.0093)	-0.0325	(0.0079)	-0.0416	(0.0092)	-	
Latent dist. (γ_{41})	-0.3834	(0.0549)	-		-0.4490	(0.0655)	-	
Latent dist. (γ_{42})	-0.3142	(0.0383)	-		-0.3439	(0.0610)	-	
Latent dist. (γ_{43})	-0.3190	(0.0750)	-		-0.2511	(0.0980)	-	
Reciprocity (η_1)	1.2330	(0.0488)	1.3725	(0.0246)	1.2192	(0.0507)	-	
Expansiveness (η_2)	0.1868	(0.0180)	0.2594	(0.0172)	0.2068	(0.0262)	-	
Expansiveness (η_3)	-0.0225	(0.0019)	-0.0286	(0.0019)	-0.0244	(0.0026)	-	
Popularity (η_4)	-0.0512	(0.0075)	-0.0336	(0.0048)	-0.0417	(0.0072)	-	
Trans. triads (η_5)	0.5910	(0.0255)	0.5647	(0.0124)	0.5844	(0.0198)	-	
Three cycles (η_6)	-0.2360	(0.0229)	-0.2347	(0.0125)	-0.2251	(0.0177)	-	
Incentive from GPA (δ_c)	0.0022	(0.0014)	0.0021	(0.0013)	-		-	
Incentive from Smoking (δ_t)	0.9020	(0.1620)	0.6092	(0.1436)	-		-	
Activity Outcome								
Endogenous (λ_c)	0.0106	(0.0027)	0.0108	(0.0041)	0.0142	(0.0032)	0.0162	(0.0034)
Endogenous (λ_t)	0.0467	(0.0100)	0.0545	(0.0123)	0.0449	(0.0107)	0.0570	(0.0115)
GPA								
Own (X)								
Age	-0.1932	(0.0193)	-0.1899	(0.0338)	-0.1940	(0.0403)	-0.2174	(0.0205)
Male	-0.1356	(0.0314)	-0.0728	(0.0489)	-0.1142	(0.0338)	-0.0758	(0.0434)
Black	0.0187	(0.0477)	0.0324	(0.0704)	0.0205	(0.0494)	0.0312	(0.0737)
Asian	0.0669	(0.0497)	0.4337	(0.1725)	0.1255	(0.0651)	0.3062	(0.2026)
Hispanic	0.0993	(0.0507)	0.1127	(0.1096)	0.1655	(0.0709)	0.1413	(0.0883)
Other race	-0.0835	(0.0390)	0.0190	(0.0669)	-0.0026	(0.0576)	-0.0084	(0.1105)
Both Parents	0.0454	(0.0341)	0.0603	(0.0495)	0.0474	(0.0403)	0.0270	(0.0519)
Less HS	-0.0991	(0.0447)	-0.0740	(0.0805)	-0.1434	(0.0518)	-0.0971	(0.0722)
More HS	0.1550	(0.0439)	0.1440	(0.0487)	0.1137	(0.0451)	0.1339	(0.0503)
Edu missing	-0.1225	(0.0343)	-0.0947	(0.0949)	-0.0766	(0.0694)	-0.0828	(0.0794)
Welfare	-0.1185	(0.0437)	0.0579	(0.1171)	0.0574	(0.0640)	0.0461	(0.1808)
Job missing	-0.1435	(0.0351)	-0.1207	(0.0798)	-0.0333	(0.0498)	-0.1101	(0.0803)
Professional	0.0239	(0.0423)	0.0582	(0.0651)	0.0698	(0.0486)	0.0206	(0.0647)
Other Jobs	-0.0312	(0.0229)	-0.0015	(0.0600)	-0.0066	(0.0348)	-0.0406	(0.0562)
Num. of students at home	0.0002	(0.0216)	-0.0147	(0.0312)	-0.0111	(0.0269)	-0.0078	(0.0284)
latent (ρ_{c11})	-0.0048	(0.0360)	-		0.2430	(0.0693)	-	
latent (ρ_{c12})	-0.0434	(0.0292)	-		-0.0028	(0.0469)	-	
latent (ρ_{c13})	0.0045	(0.0347)	-		0.0226	(0.0604)	-	
Contextual (WX)								
Age	-0.0031	(0.0016)	-0.0018	(0.0019)	-0.0014	(0.0027)	-0.0027	(0.0029)
Male	-0.0271	(0.0138)	-0.0308	(0.0168)	-0.0095	(0.0303)	-0.0168	(0.0307)
Black	-0.0354	(0.0151)	-0.0338	(0.0190)	-0.0495	(0.0208)	-0.0363	(0.0243)
Asian	0.0267	(0.0537)	-0.1162	(0.1229)	0.1507	(0.0522)	0.1958	(0.2200)

Continued on Next Page

Table – Continued

Hispanic	0.0180	(0.0297)	0.0528	(0.0360)	0.0371	(0.0466)	-0.0104	(0.0704)
Other race	0.0004	(0.0360)	0.0583	(0.0441)	-0.0833	(0.0539)	-0.0805	(0.0666)
Both Parents	0.0588	(0.0165)	0.0583	(0.0203)	0.0106	(0.0340)	0.0251	(0.0368)
Less HS	-0.0496	(0.0169)	-0.0446	(0.0265)	-0.0630	(0.0386)	-0.0432	(0.0475)
More HS	-0.0173	(0.0161)	-0.0261	(0.0202)	0.0312	(0.0326)	0.0295	(0.0318)
Edu missing	-0.0575	(0.0276)	-0.0713	(0.0379)	0.0064	(0.0628)	0.0211	(0.0607)
Welfare	-0.0528	(0.0393)	0.0090	(0.0855)	-0.1344	(0.0510)	-0.1594	(0.1751)
Job missing	-0.0508	(0.0211)	-0.0537	(0.0348)	-0.0065	(0.0450)	-0.0131	(0.0578)
Professional	-0.0002	(0.0199)	0.0078	(0.0271)	0.0174	(0.0329)	0.0099	(0.0444)
Other Jobs	-0.0122	(0.0171)	-0.0066	(0.0228)	-0.0057	(0.0307)	-0.0001	(0.0354)
Num. of students at home	0.0061	(0.0098)	0.0012	(0.0114)	0.0025	(0.0181)	0.0031	(0.0196)
latent (ρ_{c21})	-0.0269	(0.0213)	-		-0.0364	(0.0306)	-	
latent (ρ_{c22})	0.0547	(0.0243)	-		0.0158	(0.0256)	-	
latent (ρ_{c23})	-0.0452	(0.0239)	-		-0.0144	(0.0359)	-	
Correlated Random Effect (\bar{X})								
Age	0.3257	(0.6458)	0.3111	(0.6116)	0.3214	(0.6486)	0.3382	(0.6434)
Male	0.1967	(9.6754)	0.4265	(9.0640)	0.1688	(9.6482)	0.3577	(9.4387)
Black	-0.0922	(4.9512)	-0.1236	(4.6723)	-0.1716	(4.9416)	-0.1436	(4.9194)
Asian	0.3193	(18.7235)	0.3134	(18.4243)	0.7362	(18.6750)	0.5420	(18.6344)
Hispanic	0.1056	(9.7168)	-0.1482	(9.3593)	0.2323	(9.7102)	0.0720	(9.6599)
Other race	-0.0274	(14.5305)	0.1654	(13.9461)	-0.3759	(14.2762)	0.2466	(14.2024)
Both Parents	0.5370	(10.8596)	0.4546	(10.5223)	0.4652	(10.8335)	0.5256	(10.9213)
Less HS	0.1247	(12.8917)	0.1546	(12.4712)	0.1432	(12.8580)	0.0003	(12.8120)
More HS	0.4390	(8.7619)	0.1840	(8.3857)	0.2549	(8.7778)	0.1366	(8.6497)
Edu missing	0.1843	(15.6363)	0.2516	(14.9003)	0.3408	(15.4428)	0.1761	(15.2419)
Welfare	-0.0125	(18.3675)	-0.1758	(18.3475)	-0.1359	(18.4517)	0.2984	(18.3747)
Job missing	0.0691	(13.8986)	0.1937	(13.2609)	0.0224	(13.6692)	0.1137	(13.7176)
Professional	0.1912	(12.7137)	0.5938	(12.2890)	0.2481	(12.6612)	0.4122	(12.6191)
Other Jobs	-0.0949	(9.4387)	-0.0296	(8.8763)	0.0883	(9.3578)	0.0795	(9.2946)
Num. of students at home	0.0552	(5.5880)	0.0541	(5.2834)	0.1521	(5.5671)	0.0745	(5.4621)
latent (ρ_{t11})	0.0635	(5.5163)	-		-0.1827	(5.6083)	-	
latent (ρ_{t12})	0.0338	(5.5291)	-		-0.0490	(5.5158)	-	
latent (ρ_{t13})	0.0281	(5.4614)	-		-0.0379	(5.5604)	-	
Smoking								
Own (X)								
Age	0.4221	(0.0860)	0.3807	(0.0798)	0.5352	(0.0530)	0.4164	(0.0828)
Male	0.2383	(0.1308)	0.1508	(0.1517)	0.1303	(0.1247)	0.1299	(0.1254)
Black	-1.9280	(0.2679)	-1.6525	(0.2469)	-1.7920	(0.2543)	-1.6762	(0.2481)
Asian	-1.9015	(0.3318)	-2.0060	(0.4363)	-1.4356	(0.3264)	-1.8215	(0.2691)
Hispanic	-0.0401	(0.1506)	-0.7522	(0.3475)	-0.1200	(0.2954)	-0.8075	(0.3219)
Other race	0.4624	(0.2102)	0.1610	(0.2734)	0.2263	(0.2060)	0.1639	(0.3249)
Both Parents	-0.0773	(0.2025)	0.0963	(0.1648)	0.0911	(0.1784)	0.0777	(0.1615)
Less HS	-0.0970	(0.1915)	0.0281	(0.2397)	-0.0824	(0.1371)	-0.0105	(0.1875)
More HS	-0.1787	(0.1483)	-0.1338	(0.1701)	-0.0028	(0.1194)	-0.1037	(0.1476)
Edu missing	-0.3274	(0.2467)	-0.0058	(0.2638)	0.0503	(0.1632)	-0.0059	(0.1812)
Welfare	1.2654	(0.3151)	1.2816	(0.5558)	1.6113	(0.1931)	0.9937	(0.3211)
Job missing	0.2229	(0.1459)	0.2886	(0.2370)	0.3021	(0.1588)	0.1286	(0.2084)
Professional	-0.1035	(0.1522)	0.1265	(0.2052)	0.1295	(0.1571)	0.1629	(0.1814)
Other Jobs	-0.0727	(0.0889)	0.0995	(0.1787)	0.0845	(0.1460)	0.0827	(0.1668)

Continued on Next Page

Table – Continued

Num. of students at home	-0.0082	(0.1064)	-0.1450	(0.1053)	-0.1087	(0.0934)	-0.1324	(0.0897)
latent (ρ_{t11})	-0.0603	(0.1505)	-	-	-0.0045	(0.1129)	-	-
latent (ρ_{t12})	0.1276	(0.1265)	-	-	-0.1500	(0.1326)	-	-
latent (ρ_{t13})	0.0973	(0.1301)	-	-	0.1133	(0.1598)	-	-
Contextual (WX)								
Age	-0.0158	(0.0078)	-0.0250	(0.0089)	-0.0199	(0.0087)	-0.0240	(0.0082)
Male	-0.1266	(0.1240)	-0.0428	(0.1072)	-0.0418	(0.0964)	-0.0116	(0.0938)
Black	0.1686	(0.0778)	0.1745	(0.0839)	0.2081	(0.0730)	0.1857	(0.0772)
Asian	-1.1693	(0.5327)	-1.4833	(0.6703)	-1.8210	(0.2862)	-0.3909	(0.4563)
Hispanic	0.1043	(0.2275)	0.1525	(0.1769)	0.1823	(0.1498)	0.0967	(0.1749)
Other race	-0.0381	(0.1580)	0.1338	(0.2676)	0.2542	(0.1989)	0.0244	(0.2096)
Both Parents	-0.0578	(0.1225)	-0.0883	(0.1194)	-0.0246	(0.1178)	-0.0797	(0.1124)
Less HS	-0.0574	(0.1323)	-0.0529	(0.1778)	-0.1050	(0.1363)	0.0148	(0.1449)
More HS	0.2210	(0.1252)	0.0431	(0.1270)	0.0791	(0.1048)	0.0321	(0.1054)
Edu missing	0.5387	(0.1479)	0.3788	(0.2402)	0.3156	(0.2545)	0.2278	(0.2021)
Welfare	-0.0076	(0.2431)	0.6396	(0.4986)	0.1081	(0.1513)	0.5351	(0.3508)
Job missing	0.0141	(0.1541)	0.2258	(0.2094)	0.0124	(0.1642)	0.2133	(0.1709)
Professional	-0.1489	(0.1417)	0.1163	(0.1626)	-0.0870	(0.1368)	0.0680	(0.1317)
Other Jobs	0.0330	(0.1308)	0.3102	(0.1317)	0.0916	(0.1309)	0.2113	(0.1142)
Num. of students at home	0.0362	(0.0584)	-0.0085	(0.0696)	0.0045	(0.0576)	-0.0131	(0.0593)
latent (ρ_{t21})	0.0462	(0.0802)	-	-	-0.0086	(0.0866)	-	-
latent (ρ_{t22})	0.0375	(0.0731)	-	-	0.0382	(0.0921)	-	-
latent (ρ_{t23})	0.0171	(0.0717)	-	-	0.0691	(0.1028)	-	-
Correlated Random Effect (\bar{X})								
Age	-0.1652	(0.6071)	-0.1080	(0.5373)	-0.3250	(0.7489)	-0.1603	(0.5695)
Male	-1.9332	(8.8872)	-2.0111	(7.8687)	-1.4017	(10.9842)	-2.1007	(8.4227)
Black	0.6760	(4.6041)	0.2763	(4.0341)	0.4143	(5.6302)	0.2532	(4.2966)
Asian	0.2341	(18.2343)	0.4194	(17.8418)	-0.0043	(19.3088)	0.2590	(18.0800)
Hispanic	-0.9169	(9.1570)	-0.5190	(8.2714)	-1.3354	(10.7728)	-0.3832	(8.8383)
Other race	-0.5382	(13.6187)	-0.4558	(12.3635)	-0.5057	(15.4460)	-0.5624	(13.0836)
Both Parents	0.3329	(10.2479)	0.1654	(9.1941)	0.1205	(12.1814)	0.1857	(9.7085)
Less HS	-0.9616	(12.0032)	-1.5175	(11.0930)	-0.3752	(14.0601)	-1.0755	(11.6925)
More HS	-1.0461	(7.9894)	-0.6889	(7.1412)	-1.0351	(9.7780)	-0.4123	(7.5507)
Edu missing	-0.7423	(14.6718)	-0.7777	(13.3904)	-0.4994	(16.6094)	-0.4121	(14.1074)
Welfare	0.3628	(17.8178)	0.0864	(17.2431)	0.0106	(18.9248)	0.5849	(17.6998)
Job missing	-0.9863	(12.8182)	-1.7372	(11.6045)	-0.9351	(15.1333)	-1.5522	(12.2518)
Professional	-2.1732	(11.8763)	-2.9265	(10.7475)	-1.5813	(13.7070)	-2.8710	(11.4338)
Other Jobs	-0.3667	(8.8045)	-0.9255	(7.7338)	-0.2556	(10.5796)	-0.7250	(8.2415)
Num. of students at home	-0.8708	(5.1811)	-0.3703	(4.4959)	-0.4571	(6.5132)	-0.2821	(4.8715)
latent (ρ_{t11})	-0.0339	(5.0441)	-	-	0.1076	(6.6847)	-	-
latent (ρ_{t12})	-0.2307	(5.0791)	-	-	0.0487	(6.6196)	-	-
latent (ρ_{t13})	-0.1692	(5.0234)	-	-	-0.3445	(6.7278)	-	-
$\sigma_{\alpha,c}^2$	34.8945	(9.2216)	33.2820	(9.9289)	34.3369	(11.5099)	38.5704	(10.1055)
$\sigma_{\alpha,t}^2$	28.2333	(15.0682)	27.2746	(12.3443)	27.7732	(17.4457)	27.5373	(16.5197)
$\sigma_{\epsilon_c}^{2(*)}$	0.7145	(0.3142)	0.7689	(0.3501)	0.6686	(0.3238)	0.7619	(0.3564)
$\sigma_{\epsilon_t}^{2(*)}$	10.3129	(5.9717)	9.8392	(5.3859)	9.9881	(5.6195)	9.8911	(5.4676)
$\sigma_{\epsilon_{tc}}^{(*)}$	-0.5313	(0.6983)	-0.5246	(0.7401)	-0.5102	(0.7255)	-0.5107	(0.6744)

Continued on Next Page

Table – Continued

Note: The full model contains the activity outcome equation with interactions and the network formation model, where the network formation model involves the latent characteristic variables and the incentive effect. In the second column, we remove the latent variables from the network formation model. In the third column, we remove the incentive effect from the network formation model. In the fourth column, we only estimate the activity outcome equation. The MCMC runs for 150,000 iterations and the first 50,000 runs are dropped for the burn-in. Values in parentheses are standard deviations of draws from MCMC. $\sigma_{\epsilon c}^{2(*)}$, $\sigma_{\epsilon t}^{2(*)}$, and $\sigma_{\epsilon tc}^{(*)}$ denote the average of $\sigma_{\epsilon c}^2$, $\sigma_{\epsilon t}^2$, and $\sigma_{\epsilon tc}$ estimated from different groups and the value in the parenthesis is the average of standard deviations.