Multivariate Choice and Identification of Social Interactions*

Ethan Cohen-Cole†  Xiaodong Liu†  Yves Zenou§


Abstract

We investigate the identification and estimation of peer effects in a network where individuals make a multitude of interdependent choices. First, we generalize the theoretical social-interaction model to allow for multiple choices and provide a micro-foundation for the proposed econometric simultaneous equations network model. Second, we provide a set of identification conditions of peer effects in contexts of multiple choices. Third, we consider the empirical salience of the proposed model, and show that considering more than one activity changed the significance and magnitude of the estimated peer effects. Therefore, we believe that empirical peer effects research should be carefully constructed to ensure that the model reflects the full gamut of choices that individuals make.

Key words: Social networks, identification, peer effects.


---

*We thank the editor, the associate editor, and two anonymous referees for valuable comments and suggestions. This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu). No direct support was received from grant P01-HD31921 for this analysis.
†Econ One Research. E-mail: ecohencole@gmail.com.
‡Corresponding author. University of Colorado Boulder. E-mail: xiaodong.liu@colorado.edu.
§Stockholm University, Research Institute of Industrial Economics (IFN) and GAINS. E-mail: yves.zenou@ne.su.se.
1 Introduction

Peer decisions and/or peer characteristics have been shown to be important in predicting outcomes for individuals, ranging from education and crime to participation in the labor market (Ioannides and Loury, 2004; Sacerdote, 2011; Patacchini and Zenou, 2012). Most of this literature has, however, considered peer effects on choices regarding one specific activity. For example, Calvó-Armengol et al. (2009) show that peers’ grades affect a student’s own grades. Fletcher (2012) finds that a 10% increase in the proportion of classmates who drink increases the likelihood an individual drinks by five percentage points. Cutler and Glaeser (2010) find evidence for peer effects in smoking, etc.

In reality, individuals make a multitude of choices in different activities, many of which may depend on each other. As a result, peers can have multiple and sometimes opposite influences on their friends. For example, if a student’s friends smoke, drink but also perform well at school, how do these impact this student’s choice across the three activities? This joint decision problem is what we study in the current paper. Our purpose is to help understand the impact of making more than one choice at a time. In particular, we discuss the process of decision making with more than one choice in the context of peer influences and social networks. To the best of our knowledge, this is the first paper that analyzes this issue, at least from an econometric viewpoint.¹

Even with the considerable gains made to date in understanding how interactions influence individual decisions, a significant gap remains. To our knowledge, the existing literature on peer effects has entirely focused on choices in a single activity. In analyzing the decision to study, this literature ignores the related and complex decisions to smoke, drink, play sports or watch TV. While evaluating decisions in isolation makes the empirical challenges of social networks more tractable, it requires a host of unstated, and largely implausible, exclusion restrictions. Few people make important decisions about their lives completely in isolation of other factors; however, this is precisely the assumption maintained by the literature.

The inclusion of more than one choice in a individual optimization problem introduces at least two complexities. The first is a standard simultaneity problem well known in econometric simultaneous equations models. Studying, playing sports, and participating in other extracurriculars all require important time commitments. In this respect, these activities are substitutes. The presence of

¹Belhaj and Derozan (2014) develop a more restrictive network model where only two activities are considered and are assumed to be substitutes. They only analyze the theoretical implications of this model without looking at the econometric implications.
substitutability may mean that a student, who is in a varsity team, may be less responsive to the social influence of studying than another student who does not play sports. Estimation of the social influence of studying without considering the decision of playing sports may incorrectly attribute the student’s low responsiveness to the studying influence as a lack of social influence rather than the presence of other activities.

The second feature relates to the interdependence of social spillovers themselves and the link between the social effects of one action and choices in the other actions. For example, a student’s study effort may be influenced by how hard her friends study. The same is true about the decision to participate in sports. Furthermore, it also appears reasonable that the social influence of a student’s friends on studying will also have impact on her decision to play sports. Therefore, estimation of a single-activity social interaction model may suffer the risk of confounding these different social spillover effects.

The current social network literature focuses on choices in a single activity and the influence of peers on that choice (Ballester et al., 2006; Bramoullé and Kranton, 2007; Bramoullé et al., 2014; and Jackson and Zenou, 2014). The spillover effects occur as an optimization process in which individuals enjoy utility by making similar choices as their peers. Choices in other activities are assumed to be exogenous to this decision making process. We relax this strong assumption, so that the expectations that a student forms over her peers’ choices in a certain activity will depend on peers’ choices in other related activities as well. Thus, in addition to the complementarity/substitutability effects associated with multiple decisions, there are transmission channels between choices that pass through the social network as well.

Our purpose in this paper is threefold. First, we provide a structural model that helps characterize choices in multiple activities in a social interaction setting. As is common in this literature, our model has the feature that individuals enjoy utility as a function of the choices of peers. Furthermore, our model allows individuals to make choices in multiple activities that have an arbitrary degree of complementarity or substitutability. The model is general enough to encompass any arbitrary combinations of choices; that is, we make no assumption regarding the orderings of choice bundles. This generality is essential because combining sets of choices in a social interactions context into bundles dramatically restricts the set of possible actions available to individuals. It is easy to construct examples of preference reversals in the bundled goods setting that comply with standard choice axioms in the general setting here. We develop the general theoretical model in Section 2.
Second, we investigate the identification of peer effects in contexts of the simultaneous equations network model developed in Section 3. First, as in the single-activity social interaction model, we may also have the reflection problem (Manski 1993) that emerges from the coexistence of endogenous and contextual effects. Second, the social spillovers across choices in different activities add a new level of difficulty. Finally, the simultaneity problem that is endemic in simultaneous equations models requires treatment as well. To better understand the layers of complexity added by these identification issues, in Section 4, we study the identification of models with increasing degrees of interdependence in decision making.

Third, we show the empirical salience of this model for education in Section 5. In the empirical example, we show that considering more than a single activity changed the significance and magnitude of the estimated peer effects. Hence, empirical peer effects research should be carefully constructed to ensure that results reflect the full gamut of choices that individuals make. Finally, section 6 concludes.

2 Theoretical Model

Suppose a finite set of individuals \( \{1, \cdots, n\} \) is connected by a network. We keep track of social connections in the network through an adjacency matrix \( G = [g_{ij}]_{i,j=1,\cdots,n} \), where \( g_{ij} = g_{ij}^*/\sum_{j=1}^n g_{ij}^* \) with \( g_{ij}^* = 1 \) if individuals \( i \) and \( j \) are connected and \( g_{ij}^* = 0 \) otherwise. We set \( g_{ii} = 0 \). Observe that \( G \) is a row-normalized matrix where each row of \( G \) sums to one. We define the peers of individual \( i \) as the set of individuals connected to individual \( i \), i.e. \( \{ j : g_{ij}^* = 1 \} \). An example is given in Figure 1 for a tree network with four individuals.

![Figure 1: an example network and the corresponding adjacency matrix.](image)

In the network game, individuals choose effort levels of \( m \) activities to maximize their utility.  

---

2 One reason social networks matter in social interactions analysis is because they facilitate identification by solving the reflection problem. This was originally recognized in Cohen-Cole (2006) and was systematically explored in Bramoulle et al. (2009).

3 For ease of presentation, we assume that the network is undirected and no agent is isolated so that \( \sum_{j=1}^n g_{ij}^* \neq 0 \) for all \( i \). The results of the paper also hold for directed networks.
The utility of individual $i$ is a linear-quadratic function of the effort levels $y_{1i}, \cdots , y_{mi}$ given by

$$u(y_{1i}, \cdots , y_{mi}) = \sum_{k=1}^{m} \pi_i^k y_{ki} + \frac{1}{2} \left( \sum_{k=1}^{m} \sum_{l \neq i}^{m} \phi_{lk}^* y_{li} y_{ki} + \sum_{k=1}^{m} \phi_{ik}^* y_{ki}^2 \right)$$

where $\phi_{ik}^* = \phi_{ki}^*$. As in the single-activity linear-quadratic utility function (Ballester et al., 2006), the utility given by (1) has two components: payoff and cost. The marginal payoff of individual $i$’s effort in activity $k$ depends on (exogenous) attributes of individual $i$ (i.e. $\pi_i^k$) and the average effort of her peers in the same and related activities.\(^4\) The parameter $\lambda_{ik}^*$ captures the strategic substitutability or complementarity (depending on the sign of $\lambda_{ik}^*$) between individual $i$’s own effort in activity $k$ and her peers’ average effort in activity $l$. The marginal cost of individual $i$’s effort in activity $k$ depends on individual $i$’s effort in the same and related activities. The parameter $\phi_{ik}^*$ measures the substitutability or complementarity (depending on the sign of $\phi_{ik}^*$) of an individual’s effort levels in activities $k$ and $l$.\(^5\)

Given the network structure and effort levels of the peers, individual $i$ chooses effort levels $y_{1i}, \cdots , y_{mi}$ to maximize the utility (1). From the first order condition of utility maximization, we have the equilibrium best response function as

$$y_{ki} = \pi_{ki} + \sum_{l=1, l \neq k}^{m} \phi_{lk}^* y_{li} + \sum_{l=1}^{m} \lambda_{lk}^* \sum_{j=1}^{n} g_{ij} y_{lj}$$

for $k = 1, \cdots , m$, where $\pi_{ki} = \pi_i^k / \phi_{ki}^*$, $\phi_{lk} = -\phi_{lk}^*/\phi_{ki}^*$, and $\lambda_{lk} = \lambda_{ik}^*/\phi_{ki}^*$. In matrix form, the equilibrium best response function is

$$y_k = \pi_k + \sum_{l=1, l \neq k}^{m} \phi_{lk}^* y_l + \sum_{l=1}^{m} \lambda_{lk} \mathbf{G} y_l$$

for $k = 1, \cdots , m$, where $y_k = (y_{k1}, \cdots , y_{kn})'$ and $\pi_k = (\pi_{k1}, \cdots , \pi_{kn})'$. Let $\mathbf{Y} = (y_1, \cdots , y_m)$ and $\mathbf{\Pi} = (\pi_1, \cdots , \pi_m)$. It follows from (3) that

$$\mathbf{Y} = \mathbf{\Pi} + \mathbf{Y} \Phi + \mathbf{G} \mathbf{Y} \Lambda$$

\(^4\)As $\sum_{j=1}^{n} g_{ij} = 1$ for all $i$, i.e., the adjacency matrix $\mathbf{G}$ is row-normalized, the utility of an individual depends on the average effort of peers (the local-average peer effect) but not the total effort of the peers (the local-aggregate peer effect). An interpretation of the local-average peer effect is that individuals conform to the average behavior of their peers and that deviating from this social norm can be costly (Patacchini and Zenou, 2012; Liu, Patacchini and Zenou, 2014).

\(^5\)If $\phi_{kl}^* > 0$, then $\phi_{kl}^*$ partially captures the time (or total effort) constraint of an individual.
where \( \Phi = [\phi_{lk}]_{l,k=1,\ldots,m} \) (we set \( \phi_{kk} = 0 \) for \( k = 1,\ldots,m \)) and \( \Lambda = [\lambda_{lk}]_{l,k=1,\ldots,m} \). Let \( y = \text{vec}(Y) = (y_1', \ldots, y_m')' \) and \( \pi = \text{vec}(\Pi) = (\pi_1', \ldots, \pi_m') \). Then (4) can be written as

\[
y = \pi + (\Phi' \otimes I_n) y + (A' \otimes G) y.
\]

(5)

If \( (I_{mn} - \Phi' \otimes I_n - \Lambda' \otimes G) \) is invertible, then the network game with the utility (1) has a unique Nash equilibrium in pure strategies with the equilibrium effort given by

\[
y = (I_{mn} - \Phi' \otimes I_n - \Lambda' \otimes G)^{-1} \pi.
\]

(6)

Observe that when \( m = 1 \), the equilibrium effort given by (6) reduces to \( y_1 = (I_n - \lambda_{11} G)^{-1} \pi_1 \), which is the equilibrium effort of the single-activity network game studied in Ballester et al. (2006).

This theoretical model provides the economic fundamentals to understand an individual’s behavior involving multiple activities. It also facilitates the interpretation of empirical results from the econometric model described below.

3 Multivariate Choice Network Model

3.1 The econometric model

Consider a data set containing \( n \) individuals, partitioned into \( \bar{r} \) networks such that there are \( n_r \) individuals in the \( r \)-th network (\( r = 1, \ldots, \bar{r} \)) and \( \sum_{r=1}^{\bar{r}} n_r = n \). Links between individuals in network \( r \) are captured by an \( n_r \times n_r \) zero-diagonal row-normalized adjacency matrix \( G_r = [g_{ij;r}] \) as defined in the previous section.

Our specification of the econometric model follows closely from the equilibrium best response function (2) of the theoretical model. For individual \( i \) in network \( r \), let \( x_{i;r} \) be a \( p_x \times 1 \) vector of observable individual characteristics. Then, for activity \( k \) (\( k = 1, \ldots, m \)), \( \pi_{ki;r} \) in (2) can be written as

\[
\pi_{ki;r} = x_{i;r}' \beta_k + \sum_{j=1}^{n_r} g_{ij;r} x_{j;r}' \gamma_k + \alpha_{k;r} + \epsilon_{ki;r},
\]

where \( \beta_k \) and \( \gamma_k \) are \( p_x \times 1 \) vectors of unknown parameters, \( \alpha_{k;r} \) is an unknown parameter that captures unobservable network heterogeneity, and \( \epsilon_{ki;r} \) is an disturbance term that captures unobserved

---

6 For an \( n \times m \) matrix \( A = [a_{ij}] \), the vectorization of \( A \) is given by \( \text{vec}(A) = (a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{nm})' \). Note that if \( A, B, C \) are conformable matrices, then \( \text{vec}(ABC) = (C' \otimes A)\text{vec}(B) \), where \( \otimes \) denotes the Kronecker product.

7 For simplicity, we assume no individual is isolated such that \( n_r > 1 \).
servable individual heterogeneity. Thus, the econometric simultaneous equations social interaction model corresponding to (2) is then given by

\[ y_{ki,r} = \sum_{l=1, l \neq k}^{m} \phi_{lk} y_{li,r} + \sum_{l=1}^{m} \lambda_{lk} \sum_{j=1}^{n} g_{ij,r} y_{ij,r} + x'_{i,r} \beta_k + \sum_{j=1}^{n} g_{ij,r} x'_{j,r} \gamma_k + \alpha_{k,r} + \epsilon_{ki,r}, \]

or in matrix form,

\[ y_{k,r} = \sum_{l=1, l \neq k}^{m} \phi_{lk} y_{l,r} + \sum_{l=1}^{m} \lambda_{lk} G_r y_{l,r} + X_r \beta_k + G_r X_r \gamma_k + \alpha_{k,r} l_n + \epsilon_{k,r}, \]

for \( k = 1, \ldots, m \) and \( r = 1, \ldots, \tilde{r} \), where \( y_{k,r} = (y_{k1,r}, \ldots, y_{kn_r,r})' \) is an \( n_r \times 1 \) vector of observations on effort levels of activity \( k \), \( X_r = (x_{1,r}, \ldots, x_{n_r,r})' \) is an \( n_r \times p_x \) matrix of observations on exogenous individual characteristics, \( l_n \) is an \( n_r \times 1 \) vector of ones, and \( \epsilon_{k,r} = (\epsilon_{k1,r}, \ldots, \epsilon_{kn_r,r})' \) is an \( n_r \times 1 \) vector of disturbances.

Let \( \text{diag}(A_s)_{k=1}^{\tilde{s}} \) denote a “generalized” block diagonal matrix with diagonal blocks being \( n_s \times m_s \) matrices \( A_s \)'s for \( s = 1, \ldots, \tilde{s} \). For all \( \tilde{r} \) networks in the sample, let \( y_k = (y'_{k1}, \ldots, y'_{k\tilde{r}})' \), \( X = (X'_{1}, \ldots, X'_{\tilde{r}})' \), \( \epsilon_k = (\epsilon'_{k1}, \ldots, \epsilon'_{k\tilde{r}})' \), \( \alpha_k = (\alpha_{k1}, \ldots, \alpha_{k\tilde{r}})' \), \( L = \text{diag}(l_n)_{r=1}^{\tilde{r}} \), and \( G = \text{diag}(G_r)_{r=1}^{\tilde{r}} \). Then,

\[ y_k = \sum_{l=1, l \neq k}^{m} \phi_{lk} y_l + \sum_{l=1}^{m} \lambda_{lk} G y_l + X \beta_k + GX \gamma_k + L \alpha_k + \epsilon, \]  

for \( k = 1, \ldots, m \). We assume \( (\epsilon_1, \ldots, \epsilon_m) = \mathbf{V} \Sigma_* \) where \( \mathbf{V} = [v_{ik}] \) is an \( n \times m \) matrix of i.i.d. innovations with zero mean and unit variance and \( \Sigma_* \) is a nonsingular \( m \times m \) matrix. Observe that \( \epsilon = (\epsilon_1', \ldots, \epsilon_m')' = \text{vec}(\mathbf{V} \Sigma_*) = (\Sigma_*' \otimes I_n)\text{vec}(\mathbf{V}) \). Hence \( E(\epsilon) = 0 \) and \( E(\epsilon \epsilon') = (\Sigma_*' \otimes I_n)(\Sigma_* \otimes I_n) = \Sigma \otimes I_n \), where \( \Sigma = \Sigma'_* \Sigma_* \). Thus, the disturbances of the same individual are allowed to be correlated across different activities.

3.2 Identification challenges

As in most models in the social interaction literature, a host of identification issues arises in model (7). In particular, model (7) not only suffers from the reflection problem as single-activity peer effect models but also has the simultaneity issue that is endemic to simultaneous equations models. Our main interest in this paper is to study the identification of the following effects in model (7).
3.2.1 The endogenous effect and contextual effect

The *endogenous effect*, where an individual’s choice depends on the choices of her peers in the same activity, is captured by the coefficient $\lambda_{kk}$. The *contextual effect*, where an individual’s choice depends on the exogenous characteristics of her peers, is captured by $\gamma_k$. The *reflection problem* (Manski, 1993) is well known and emerges from the coexistence of these two effects. In Manski’s linear-in-means model, individuals are affected by all members of their group and by no one outside the group, and thus the simultaneity in behavior of individuals in the same group introduces a perfect collinearity between the endogenous effect and the contextual effect. Hence, these two effects cannot be separately identified in the linear-in-means model.

On the other hand, in a social network, usually individuals are not impacted evenly by the full population in the network. Instead, they are influenced by their (direct) connections or peers. Thus, the structure of social networks can be exploited to identify peer effects. Bramoullé et al. (2009) show that the endogenous and contextual effects can be identified if intransitivities exist in a network so that $I, G, G^2$ are linearly independent. Intuitively, if individuals $i, j$ are connected and $j, k$ are connected, it does not necessarily imply that $i, k$ are also connected. Because of intransitivities, the characteristics of an individual’s indirect connections are not collinear with her own characteristics and the characteristics of her direct connections. Therefore, the characteristics of an individual’s indirect connections can be used as IVs to identify the endogenous effect from the contextual effect.

Suppose $\phi_{lk} = 0$ and $\lambda_{lk} = 0$ for $l \neq k$. Then, model (7) becomes a system of ‘seemingly unrelated’ equations with only endogenous and contextual effects. In this case, the identification condition given by Bramoullé et al. (2009) would apply because this condition is derived based on the mean of the reduced-form equation and is thus not affected by the cross-equation correlation in the disturbances.

3.2.2 The simultaneity effect and cross-activity peer effect

The standard economic *simultaneity effect* can be seen in the coefficient $\phi_{lk}$ of model (7). Indeed, an individual’s choice in a certain activity may depend on her own choices in related activities. In the absence of peer effects, this simultaneity in decision making leads to a well known identification problem for a simultaneous equations model. The usual remedy for this identification problem is to impose exclusion restrictions on the model coefficients.

A central component of our model is that we allow peers’ choices in certain activities to influence
an individual’s decisions on related activities (e.g. the time an individual’s friends spend on watching TV may have impact on her decision on study effort). We model this effect through the cross-activity peer effect that is represented by the coefficient $\lambda_{lk}$ for $l \neq k$.

The simultaneity effect and cross-activity peer effect introduce additional layers of complication to the identification. In the following, we show that the simultaneous equations social interaction model can be identified if $\phi_{lk} = 0$ or $\lambda_{lk} = 0$, for all $l, k = 1, \ldots, m$ such that $l \neq k$, by further exploiting the set of exclusion restrictions from the intransitivities that exist in a social network. However, for a general simultaneous equations social interaction model given by (7), the intransitivities in the network structure would not be enough to identify the above four types of effects. Hence, to achieve identification, we may need to impose some exclusion restrictions on the model coefficients.

### 3.2.3 The network correlated effect and cross-activity correlated effect

The structure of the general simultaneous equations model is flexible enough to allow us to incorporate two types of correlated effects.

First, individuals in the same network may behave similarly, because they have similar unobserved individual characteristics and face similar institutional environments. We call this type of correlated effect the network correlated effect. The network correlated effect is captured in our general simultaneous equations model by the vector of network fixed effect parameters $\alpha_k$. The network fixed effect can be interpreted as originating from a two-step link formation model, where individuals self-select into different networks in a first step with selection bias due to network-specific characteristics and, then, in a second step, link formation takes place within networks based on observable individual characteristics only. Therefore, network fixed effects serve as a (partial) remedy for the bias that originates from the possible sorting of individuals into networks.

Second, choices of the same individual in related activities may be correlated. We call this type of correlated effect the cross-activity correlated effect, which is captured by the off-diagonal parameters in the covariance matrix $\Sigma$. Since our identification results are based on the mean of reduced-form equations, they will not be affected by the cross-equation correlation in the error term. However, for estimation efficiency, it is important to consider the correlation structure of error terms. Following Kelejian and Prucha (2004), we adopt a 3SLS estimator for the estimation of our model. The details of the estimator are given in Appendix A.
4 Identification of Social Interaction Effects

In model (7), we allow network fixed effects to depend on $G$ and $X$ by treating $\alpha_k = (\alpha_{k,1}, \cdots, \alpha_{k,r})'$ as a vector of unknown parameters. To avoid the “incidental parameters” problem (Neyman and Scott, 1948) when the number of network $r$ is large, we transform (7) with a projector $J = \text{diag}\{J_r\}_{r=1}^R$, where $J_r = I_{n_r} - \frac{1}{n_r}l_{n_r}l_{n_r}'$. This transformation is analogous to the “within” transformation for fixed effect panel data models. As $JL = 0$, the transformed model is

$$Jy_k = \sum_{l=1, l \neq k}^m \phi_{lk} Jy_l + \sum_{l=1}^m \lambda_{lk} JGy_l + JX\beta_k + JGX\gamma_k + Je_k,$$

for $k = 1, \cdots, m$.

Under the regularity assumptions listed in Appendix A, the transformed model (8) can be estimated by the generalized spatial 2SLS (GS2SLS) and 3SLS (GS3SLS) estimators in Kelejian and Prucha (2004) as detailed in Appendix A. In particular, Assumption 6 in Appendix A provides a sufficient condition for the identification of (8). Let $Z_k$ denote a matrix of regressors that appear in the $k$-th equation. Let $Q$ denote an IV matrix based on $G, X$ and their functions. Assumption 6 implies rank conditions that $E(JZ_k)$ has full column rank and that the column rank of $JQ$ is at least as high as that of $E(JZ_k)$, for large enough $n$.

To get more insights into this identification condition, we consider a simple multivariate choice network model involving only two activities, i.e., $m = 2$. For the ease of presentation, we assume that $X$ is a column vector with $p_x = 1$ in the following discussion. In this case, (7) can be written as

$$y_1 = \phi_{21} y_2 + \lambda_{11} Gy_1 + \lambda_{21} Gy_2 + X\beta_1 + GX\gamma_1 + L\alpha_1 + \epsilon_1$$

$$y_2 = \phi_{12} y_1 + \lambda_{22} Gy_2 + \lambda_{12} Gy_1 + X\beta_2 + GX\gamma_2 + L\alpha_2 + \epsilon_2.$$  

with reduced-form equations

$$y_1 = S^{-1}[X(\phi_{21}\beta_2 + \beta_1) + GX(\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \phi_{21}\gamma_2 + \gamma_1) + G^2X(\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1)] + L\alpha_1^* + S^{-1}u_1$$

\textbf{Footnotes:}

9 For example, an IV matrix can be $Q = [X, GX, \cdots, G^pX]$ for some $p > 1$. For ease of presentation, we assume $G$ and $X$ are nonstochastic in this paper.

10 A sufficient condition for $S$ to be invertible is $|\phi_1\phi_2| + |\lambda_{11} + \lambda_{22} + \phi_1\lambda_{12} + \phi_2\lambda_{21}| + |\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21}| < 1$.  

10
\[ y_2 = S^{-1} [X(\phi_{12}\beta_1 + \beta_2) + GX(\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \phi_{12}\gamma_1 + \gamma_2) + G^2X(\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2)] + L\alpha^*_2 + S^{-1}u_2 \]  

(12)

where

\[ S = (1 - \phi_{12}\phi_{21})I - (\lambda_{11} + \lambda_{22} + \phi_{12}\lambda_{21} + \phi_{21}\lambda_{12})G + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})G^2 \]  

(13)

\[ \alpha^*_1 = \frac{(1 - \lambda_{22})\alpha_1 + (\phi_{21} + \lambda_{21})\alpha_2}{1 - \lambda_{11} - \lambda_{22} - \phi_{12}\lambda_{21} - \phi_{21}\lambda_{12} - \phi_{12}\phi_{21} - \lambda_{12}\lambda_{21} + \lambda_{11}\lambda_{22}} \]

\[ \alpha^*_2 = \frac{(1 - \lambda_{11})\alpha_2 + (\phi_{12} + \lambda_{12})\alpha_1}{1 - \lambda_{11} - \lambda_{22} - \phi_{12}\lambda_{21} - \phi_{21}\lambda_{12} - \phi_{12}\phi_{21} - \lambda_{12}\lambda_{21} + \lambda_{11}\lambda_{22}} \]

and

\[ u_1 = S^{-1}[(I - \lambda_{22}G)e_1 + (\phi_{21}I + \lambda_{21}G)e_2] \]  

(14)

\[ u_2 = S^{-1}[(I - \lambda_{11}G)e_2 + (\phi_{12}I + \lambda_{12}G)e_1]. \]  

(15)

The within transformation with \( J \) of (9) and (10) gives

\[ Jy_1 = \phi_{21}Jy_2 + \lambda_{11}JGy_1 + \lambda_{21}JGy_2 + JX\beta_1 + JGX\gamma_1 + J\epsilon_1 \]  

(16)

\[ Jy_2 = \phi_{12}Jy_1 + \lambda_{22}JGy_2 + \lambda_{12}JGy_1 + JX\beta_2 + JGX\gamma_2 + J\epsilon_2. \]  

(17)

In this section, we inspect the rank of the regressor matrix for equations (16) and (17) to understand the identification condition for the multivariate choice network model with increasing levels of interdependence in decision making.

4.1 The ‘seemingly unrelated’ regressions (SUR) network model

First, we shut down the simultaneity effect and cross-activity peer effect in the utility function (1) by assuming that \( \phi_{lk} = \lambda_{lk}^* = 0 \) for \( l \neq k \). This exclusion restriction implies that, in the econometric model, \( \phi_{lk} = \lambda_{lk} = 0 \) for \( l \neq k \), such that (16) and (17) are in the form of

\[ Jy_k = \lambda_{kk}JGy_k + JX\beta_k + JGX\gamma_k + J\epsilon_k \]  

(18)

with the reduced-form equation

\[ Jy_k = J(I - \lambda_{kk}G)^{-1}X\beta_k + J(I - \lambda_{kk}G)^{-1}GX\gamma_k + J(I - \lambda_{kk}G)^{-1}\epsilon_k \]  

(19)
if \((I - \lambda_{kk}G)\) is assumed to be invertible. In this case, an individual’s decision in a certain activity is still allowed to be correlated with her decisions in other activities through correlation of the error terms.

Let \(JZ_k = [JGy_k, JX, JGX]\) be the within-transformed regressor matrix of (18). Observe that \(JGJ = JG\) and \((I - \lambda_{kk}G)^{-1} = I + \lambda_{kk}G + \lambda_{kk}^2 G^2 + \cdots\) if \(|\lambda_{kk}| < 1\). If we premultiply both sides of (19) by \(JG\) and take the expected value, we have

\[
E(JGy_k) = JGX \beta_k + J(I + \lambda_{kk}G + \lambda_{kk}^2 G^2 + \cdots)G^2X(\lambda_{kk}\beta_k + \gamma_k).
\]

If \(\lambda_{kk}\beta_k + \gamma_k = 0\) in the data generating process (DGP), then \(E(JGy_k) = JGX \beta_k\). In this case, \(E(JZ_k) = [E(JGy_k), JX, JGX]\) does not have full column rank due to the perfect collinearity of \(E(JGy_k)\) and \(JGX\), and thus Assumption 6 in Appendix A is violated. A special case of \(\lambda_{kk}\beta_k + \gamma_k = 0\) is \(\beta_k = \gamma_k = 0\). In this case, there is no relevant exogenous variable in (18) that can be used as an IV for the endogenous regressor \(Gy_k\).

On the other hand, if \(\lambda_{kk}\beta_k + \gamma_k \neq 0\) in the DGP, \(JG^2X\) can be used as an IV to identify the endogenous peer effect as long as \(JG^2X\) is linearly independent of the exogenous regressors \(JX\) and \(JGX\) of the model. The following proposition shows that the rank condition of \(E(JZ_k)\) relies on some topological properties of the network.

**Proposition 1** Suppose \(\lambda_{kk}\beta_k + \gamma_k \neq 0\) in the DGP. Then \(E(JZ_k)\) of (18) has full column rank if \(J, JG, JG^2\) are linearly independent.

Bramoullé et al. (2009) have shown that the single-activity network model with the network correlated effect can be identified if \(I, G, G^2, G^3\) are linearly independent. As shown by Lemma B.1 in Appendix B, linear independence of \(I, G, G^2, G^3\) implies linear independence of \(J, JG, JG^2\), and hence the identification condition derived by Bramoullé et al. (2009) applies to the SUR model. This is not surprising since the identification condition derived by Bramoullé et al. (2009) is based on the mean of the reduced-form equation (19), which is not affected by the correlation structure of the error terms in the SUR model.

### 4.2 The simultaneous equations network model with simultaneity effects

Next, we shut down only the cross-activity peer effect in the utility function (1) by assuming that \(\lambda_{lk} = 0\) for \(l \neq k\). This exclusion restriction implies that \(\lambda_{12} = \lambda_{21} = 0\) in (16) and (17) so that the
within-transformed model can be written as

\begin{align}
Jy_1 & = \phi_{21} Jy_2 + \lambda_{11} JGy_1 + JX \beta_1 + JGX \gamma_1 + J\epsilon_1 \\
Jy_2 & = \phi_{12} Jy_1 + \lambda_{22} JGy_2 + JX \beta_2 + JGX \gamma_2 + J\epsilon_2
\end{align}

(20)

(21)

with the reduced-form equations

\begin{align}
E(Jy_1) & = JS_1^{-1} [X(\phi_{21}\beta_2 + \beta_1) + GX(\phi_{21}\gamma_2 + \gamma_1 - \lambda_{22}\beta_1) - \lambda_{22}G^2 X \gamma_1] \\
E(Jy_2) & = JS_1^{-1} [X(\phi_{12}\beta_1 + \beta_2) + GX(\phi_{12}\gamma_1 + \gamma_2 - \lambda_{11}\beta_2) - \lambda_{11}G^2 X \gamma_2]
\end{align}

(22)

(23)

where \( S_1 = (1 - \phi_{12}\phi_{21})I - (\lambda_{11} + \lambda_{22})G + \lambda_{11}\lambda_{22}G^2 \). Due to the symmetry of (20) and (21), we focus on the identification of (20) in the following discussion. Let \( JZ_1 = [Jy_2, JGy_1, JX, JGX] \).

The following proposition gives a sufficient condition for \( E(JZ_1) \) to have full column rank. Let

\[
A_1 = \begin{bmatrix}
\phi_{12}\beta_1 + \beta_2 & 0 & 1 - \phi_{12}\phi_{21} & 0 \\
\phi_{12}\gamma_1 + \gamma_2 - \lambda_{11}\beta_2 & \beta_1 + \phi_{21}\beta_2 & -(\lambda_{11} + \lambda_{22}) & 1 - \phi_{12}\phi_{21} \\
-\lambda_{11}\gamma_2 & \phi_{21}\gamma_2 + \gamma_1 - \lambda_{22}\beta_1 & \lambda_{11}\lambda_{22} & -(\lambda_{11} + \lambda_{22}) \\
0 & -\lambda_{22}\gamma_1 & 0 & \lambda_{11}\lambda_{22}
\end{bmatrix}.
\]

(24)

**Proposition 2** Suppose the parameter matrix \( A_1 \) given by (24) has full rank in the DGP. Then \( E(JZ_1) \) of model (20) has full column rank if \( J, JG, JG^2, JG^3 \) are linearly independent.

Similar to the condition that \( \lambda_{kk}\beta_k + \gamma_k \neq 0 \) for the identification of SUR models, the rank condition of \( A_1 \) is necessary for \( E(JZ_1) \) to have full column rank. Suppose \( \lambda_{11} = \lambda_{22} = 0 \) in the DGP and thus \( A_1 \) does not have full rank. Then, it follows from (23) that

\[
E(Jy_2) = \frac{1}{1 - \phi_{12}\phi_{21}} [JX(\phi_{12}\beta_1 + \beta_2) + JGX(\phi_{12}\gamma_1 + \gamma_2)].
\]

In this case, \( E(Jy_2) \) is a linear combination of \([JX, JGX]\) and hence \( E(JZ_1) \) does not have full rank. This is not surprising as the true model is a “standard” simultaneous equations model without endogenous peer effects if \( \lambda_{11} = \lambda_{22} = 0 \) in the DGP. It is well known that simultaneous equations models cannot be identified without imposing exclusion restrictions on coefficients of exogenous regressors.
On the other hand, if $A_1$ has full rank, then it follows from the reduced-form equations (22), (23) and a series expansion of $S_1^{-1}$ that $E(Jy_1)$ and $E(Jy_2)$ can be written as linear combinations of $[JX, JGX, JG^2X, JG^3X, \cdots]$. Therefore, if $J, JG, JG^2, JG^3$ are linearly independent, then $[JG^2X, JG^3X]$ can be used as IVs to identify the simultaneity effect and endogenous peer effect in (20) and (21), without imposing exclusion restrictions on coefficients of exogenous regressors.

### 4.3 The simultaneous equations network model with cross-activity peer effects

Now, we shut down only the simultaneity effect in the utility function (1) by assuming that $\phi_{lk} = 0$ for $l \neq k$. This exclusion restriction implies that $\phi_{12} = \phi_{21} = 0$ in (16) and (17) so that the within-transformed model can be written as

\[
Jy_1 = \lambda_{11}JGy_1 + \lambda_{21}JGy_2 + JX\beta_1 + JGX\gamma_1 + Je_1 \quad (25)
\]

\[
Jy_2 = \lambda_{22}JGy_2 + \lambda_{12}JGy_1 + JX\beta_2 + JGX\gamma_2 + Je_2 \quad (26)
\]

with the reduced-form equations

\[
E(Jy_1) = JS_2^{-1}[X\beta_1 + GX(\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \gamma_1) + G^2X(\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1)] \quad (27)
\]

\[
E(Jy_2) = JS_2^{-1}[X\beta_2 + GX(\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \gamma_2) + G^2X(\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2)] \quad (28)
\]

where $S_2 = I - (\lambda_{11} + \lambda_{22})G + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})G^2$. In the following discussion, we focus on the identification of (25). The identification of (26) is analogous. Let $JZ_1 = [JGy_1, JGy_2, JX, JGX]$.

The following proposition gives a sufficient condition for $E(JZ_1)$ to have full column rank. Let

\[
A_2 = \begin{bmatrix}
0 & 0 & 1 & 0 \\
\beta_1 & \beta_2 & -(\lambda_{11} + \lambda_{22}) & 1 \\
\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \gamma_1 & \lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \gamma_2 & \lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21} & -(\lambda_{11} + \lambda_{22}) \\
\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1 & \lambda_{12}\gamma_1 - \lambda_{11}\gamma_2 & 0 & \lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21}
\end{bmatrix} \quad (29)
\]

**Proposition 3** Suppose the parameter matrix $A_2$ given by (29) has full rank in the DGP. Then $E(JZ_1)$ of model (25) has full column rank if $J, JG, JG^2, JG^3$ are linearly independent.

Similar to the rank condition of $A_1$ for the identification of (20), the rank condition of $A_2$ is necessary for $E(JZ_1)$ of (25) to have a full column rank. Suppose $\lambda_{11} = \lambda_{22}$ and $\lambda_{12} = \lambda_{21} = 0$ in
the DGP and thus the true model is a SUR model with identical endogenous peer effects. In this case, \( A_2 \) does not have full rank. As \( JGJ = JG \), if we premultiply both sides of (27) and (28) by \( JG \), then

\[
\begin{align*}
E(JGy_1) &= JG(I - \lambda_{11}G)^{-1}(X\beta_1 + GX\gamma_1) \\
E(JGy_2) &= JG(I - \lambda_{11}G)^{-1}(X\beta_2 + GX\gamma_2).
\end{align*}
\]

Hence, (25) cannot be identified due to the perfect collinearity of \( E(JGy_1) \) and \( E(JGy_2) \).

On the other hand, if \( A_2 \) has full rank, then it follows from the reduced-form equations (27), (28) and a series expansion of \( S_2^{-1} \) that \( E(JGy_1) \) and \( E(JGy_2) \) can be written as linear combinations of \([JGX, JG^2X, JG^3X, \cdots] \). Therefore, if \( J, JG, JG^2, JG^3 \) are linearly independent, then \([JG^2X, JG^3X] \) can be used as IVs to identify the endogenous peer effect and cross-activity peer effect in (25) and (26).

### 4.4 The general simultaneous equations network model

Finally, we consider the identification of the general simultaneous equations network model (16) and (17) with both simultaneity and cross-activity effects. Let \( Z_1 = [y_2, Gy_1, Gy_2, X, GX] \) and \( Z_2 = [y_1, Gy_2, Gy_1, X, GX] \). The following proposition shows that, without imposing any exclusion restrictions, (16) and (17) cannot be identified using an IV-based estimation method.

**Proposition 4** \( E(JZ_1) \) and \( E(JZ_2) \) of (16) and (17) do not have full column rank.

To gain some insight on this non-identification result, we inspect the reduced-form equations of the general simultaneous equations network model. If we premultiply the reduced-form equations (11) and (12) by \( JS \), we have

\[
\begin{align*}
JSy_1 &= JX(\phi_1\beta_2 + \beta_1) + JGX(\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \gamma_1 + \phi_1\gamma_2) + JG^2X(\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1) + Ju_1 \\
JSy_2 &= JX(\phi_2\beta_1 + \beta_2) + JGX(\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \gamma_2 + \phi_2\gamma_1) + JG^2X(\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2) + Ju_2
\end{align*}
\]

where \( u_1 \) and \( u_2 \) are given by (14) and (15) respectively. Substitution of (13) into (30) and (31) gives

\[
Jy_1 = \psi_1JGy_1 + \psi_2JG^2y_1 + JX\delta_{11} + JGX\delta_{12} + JG^2X\delta_{13} + \frac{1}{1 - \phi_1\phi_2}Ju_1
\]
\[ Jy_2 = \psi_1 JGy_2 + \psi_2 JG^2 y_2 + JX\delta_{21} + JGX\delta_{22} + JG^2 X\delta_{23} + \frac{1}{1 - \phi_1 \phi_2} Ju_2 \]  

(33)

where

\[ \psi_1 = \frac{\lambda_{11} + \lambda_{22} + \phi_1 \lambda_{12} + \phi_2 \lambda_{21}}{1 - \phi_1 \phi_2}, \quad \psi_2 = \frac{\lambda_{11} \lambda_{22} - \lambda_{12} \lambda_{21}}{1 - \phi_1 \phi_2}, \]

and

\[ \delta_{11} = \frac{\phi_1 \beta_2 + \beta_1}{1 - \phi_1 \phi_2}, \quad \delta_{12} = \frac{\lambda_{21} \beta_2 - \lambda_{22} \beta_1 + \gamma_1 + \phi_1 \gamma_2}{1 - \phi_1 \phi_2}, \quad \delta_{13} = \frac{\lambda_{21} \gamma_2 - \lambda_{22} \gamma_1}{1 - \phi_1 \phi_2}, \]

\[ \delta_{21} = \frac{\phi_2 \beta_1 + \beta_2}{1 - \phi_1 \phi_2}, \quad \delta_{22} = \frac{\lambda_{12} \beta_1 - \lambda_{11} \beta_2 + \gamma_2 + \phi_2 \gamma_1}{1 - \phi_1 \phi_2}, \quad \delta_{23} = \frac{\lambda_{12} \gamma_2 - \lambda_{11} \gamma_1}{1 - \phi_1 \phi_2}. \]

Let \( y = (y_1, y_2)', \ J = I_2 \otimes J, \ G = I_2 \otimes G, \ X = I_2 \otimes X, \) and \( u = \frac{1}{1 - \phi_1 \phi_2} (u_1', u_2'). \) Then, (32) and (34) can be written more compactly as

\[ \tilde{J} y = \psi_1 \tilde{J} G y + \psi_2 \tilde{J} G^2 y + \tilde{J} X \delta_1 + \tilde{J} G X \delta_2 + \tilde{J} G^2 X \delta_3 + \tilde{J} v \]  

(34)

where \( \delta_1 = (\delta_{11}, \delta_{21})', \ \delta_2 = (\delta_{12}, \delta_{22})', \) and \( \delta_3 = (\delta_{13}, \delta_{23})'. \) The reduced-form equation (34) can be considered as a (within-transformed) single-equation spatial Durbin model (LeSage and Pace, 2009) with high order spatial lags in the dependent variable and regressors. Under some regularity conditions, the reduced-form parameters \( \psi_1, \psi_2, \delta_1, \delta_2, \delta_3 \) in (34) can be identified. However, it is not possible to recover all the structural parameters in (16) and (17) from the reduced-form parameters without imposing some exclusion restrictions on the structural parameters.

Therefore, for a general simultaneous equations network model with both simultaneity and cross-activity peer effects, exploiting the exclusion restrictions from the intransitivities that exist in the network structure is not sufficient for the identification. One possibility to achieve identification is to impose exclusion restrictions on the exogenous regressors. To clarify ideas, we consider the following model

\[ y_1 = \phi_{21} y_2 + \lambda_{11} G y_1 + \lambda_{21} G y_2 + X_1 \beta_1 + GX_1 \gamma_1 + L \alpha_1 + \epsilon_1 \]  

(35)

\[ y_2 = \phi_{12} y_1 + \lambda_{22} G y_2 + \lambda_{12} G y_1 + X_2 \beta_2 + GX_2 \gamma_2 + L \alpha_1 + \epsilon_2 \]  

(36)

where \( X_1 \) and \( X_2 \) are assumed to be column vectors for simplicity. Let \( Z_1 = [y_2, Gy_1, Gy_2, X_1, GX_1] \) and \( Z_2 = [y_1, Gy_2, Gy_1, X_2, GX_2]. \) The following proposition gives a sufficient condition for the
rank condition of $E(JZ_1)$. The sufficient condition for the rank condition of $E(JZ_2)$ can be analogously derived. Let

$$A_3 = [B', C']'$$

(37)

where

$$B = [b_{ij}] = \begin{bmatrix}
\phi_{12}\beta_1 & 0 & 0 & 1 - \phi_{12}\phi_{21} & 0 \\
\lambda_{12}\beta_1 + \phi_{12}\gamma_1 & \beta_1 & \phi_{12}\beta_1 & b_{24} & 1 - \phi_{12}\phi_{21} \\
\lambda_{12}\gamma_1 & \gamma_1 - \lambda_{22}\beta_1 & \lambda_{12}\beta_1 + \phi_{12}\gamma_1 & \lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21} & b_{35} \\
0 & -\lambda_{22}\gamma_1 & \lambda_{12}\gamma_1 & 0 & (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})
\end{bmatrix}$$

with $b_{24} = b_{35} = - (\lambda_{11} + \lambda_{22} + \lambda_{12}\phi_{21} + \lambda_{21}\phi_{12})$, and

$$C = [c_{ij}] = \begin{bmatrix}
\beta_2 & 0 & 0 & 0 & 0 \\
\gamma_2 - \lambda_{11}\beta_2 & \phi_{21}\beta_2 & \beta_2 & 0 & 0 \\
-\lambda_{11}\gamma_2 & \lambda_{21}\beta_2 + \phi_{21}\gamma_2 & \gamma_2 - \lambda_{11}\beta_2 & 0 & 0 \\
0 & \lambda_{21}\gamma_2 & -\lambda_{11}\gamma_2 & 0 & 0
\end{bmatrix}.$$ 

Let

$$A'_3 = [B'', C'']'$$

(38)

where $B''$ is a $3 \times 5$ matrix with the $(i,j)$-th entry being $b_{ij} + \rho_i b_{4j}$ and $C''$ is a $3 \times 5$ matrix with the $(i,j)$-th entry being $c_{ij} + \rho_i c_{4j}$ for some constant scalars $\rho_1, \rho_2, \rho_3$.

**Proposition 5** Suppose the column rank of $[X_1, X_2]$ is higher than that of $X_1$ and $X_2$. If (i) $J, JG, JG^2, JG^3$ are linearly independent and the parameter matrix $A_3$ given by (37) has full rank in the DGP, or (ii) $J, JG, JG^2$ are linearly independent, $JG^3 = \rho_1 J + \rho_2 JG + \rho_3 JG^2$ for some constant scalars $\rho_1, \rho_2, \rho_3$, and the parameter matrix $A'_3$ given by (38) has full rank in the DGP, then $E(JZ_1)$ of (35) has full column rank.

5  **Empirical Application**

5.1  **Data**

To illustrate the empirical salience of the proposed model, we study the peer effect on a student’s choice between studying and watching TV, using a unique and now widely used data set provided
by the National Longitudinal Survey of Adolescent Health (Add Health). The data set collected national representative information on 7th-12th graders in both public and private schools in the United States. It provides detailed data on students’ social, economic, psychological and physical well-being with information on friendship networks.

In this empirical example, we consider the estimation of (16) and (17) where $y_1$ and $y_2$ measure, respectively, study effort and time spent on watching TV. To be more specific, $y_1$ is based on the survey question “In general, how hard do you try to do your school work well?”, coded as 0 if the response is “I never try at all”, 1 if “I don’t try very hard”, 2 if “I try hard enough / but not as hard as I could”, and 3 if “I try very hard to do my best”. $y_2$ is based on the survey question “Outside of school hours, about how much time do you spend watching television or video cassettes on an average school day?”, coded as 0 if the response is “none”, 1 if “less than 1 hour”, 2 if “1-2 hours”, 3 if “3-4 hours”, and 4 if “more than 4 hours”.

The adjacency matrix $G = [g_{ij}]$, where $g_{ij} = g_{ij}^* / \sum_{j=1}^{n} g_{ij}^*$, is constructed based on the friend-nomination information provided by the Add Health data. In the Add Health survey questionnaire, students were asked to identify their 10 best friends from a school roster. Thus, $g_{ij}^* = 1$ if student $i$ nominates student $j$ as a friend and $g_{ij}^* = 0$ otherwise. After removing isolated students (i.e. students who nominated no friends and were not nominated by any students), the sample consists of 64,858 students distributed over 1183 networks, with network size ranging from 2 to 484. The mean and the standard deviation of network sizes are 54.83 and 89.23. A list of the variables used in the empirical example is given in Table 1.

5.2 The SUR network model

First, we consider the SUR model with network fixed effects. After a within transformation to eliminate the fixed effect parameters, we estimate the model by the GS2SLS and GS3SLS estimators with the IV matrix $Q = [X, GX, G^2 X]$. The GS3SLS estimator can be considered as a generalized least squares version of the GS2SLS estimator, which takes into account the cross-equation correlation of the disturbances. The details of the estimators are given in Appendix A.

Table 2 reports the estimation results. The average study effort of friends has a positive impact

---

11 Less than 0.5% of the students in the sample nominated 10 friends and thus the bound on the number of friend-nominations is not binding.
on one’s own study effort. This result is in line with most papers showing positive peer effects in education (see, e.g. Calvó-Armengol, et al., 2009; DiGiorgi et al., 2010; Bifulco, et al., 2011). Also, the time friends spend watching TV has a positive and statistically significant impact on one’s own time spent watching TV.

5.3 The simultaneous equations network model

We consider three specifications of the simultaneous equations network model. Model 1 given by (20) and (21) does not consider the cross-activity peer effect. Model 2 given by (25) and (26) does not consider the simultaneity effect. Model 3 given by (35) and (36) is the most general one with all the social interaction effects. As discussed in Section 4.4, identification of Model 3 requires two IVs (or exclusion restrictions), one for each activity. We instrument the study effort of a student with her health condition, and instrument the TV watching time by a dummy variable that takes value 1 if the parents let the student make her own decision on what TV program to watch (the variable “TV program” in Table 1). Table 3 reports the GS3SLS estimation results for these models.

The estimation results of the three models are largely consistent. The estimates of simultaneity effects are negative and statistically significant, suggesting a substitutability effect between study effort and TV watching. That is, the more a student watches TV, the less effort she puts into her study. The estimates of cross-activity peer effects are negative. Thus, quite naturally, a student studies harder if her friends watches less TV and vice versa. The estimates of the endogenous peer effects are positive and statistically significant for both activities, which means a student studies harder if her friends study harder and a student watches more TV if her friends watch more TV. Interestingly, in the SUR model without simultaneity and cross-activity peer effects, we find that the endogenous effect on study effort is not statistically significant (see Table 2). This highlights the importance of considering more than one activity in the network model.

19
6 Conclusion

In this paper, we investigate the impact of peers on own outcomes where individuals embedded in a network are involved in more than one activity. We develop a general simultaneous equations network model that captures the different social interaction effects. In addition to endogenous, contextual and correlated effects that exist in a single-activity network model, we introduce the simultaneity effect, where an individual’s choice on a certain activity may be affected by her choices on related activities; the cross-activity peer effect, where an individual’s choice on a certain activity may be affected by peers’ choices on related activities; and the cross-activity correlated effect, where the unobservables that affect an individual’s choices are correlated across related activities. We consider identification of network models with some or all of the above effects. We show that the general simultaneous equations network model with all the effects cannot be identified without any exclusion restrictions. We then study the impact of peer effects on education and recreation activities (e.g. TV watching) and show that the estimated simultaneity effect, endogenous effect and cross-activity peer effect all have non-trivial impacts on adolescent behavior.

We believe that the methodology developed in this paper is important because, in real-world situations, individuals often make decisions involving more than one activity. In terms of policy implications, this implies that the social planner could use more than one instrument to make an effective policy. For example, most policies aiming at reducing crime focus on the deterrence effect of punishment and the social influence of punishment (Patacchini and Zenou, 2012). Using the model developed in this paper, one could characterize the social interdependence of crime and education and develop a more effective policy that uses both punishment and education to reduce crime.

References


APPENDIX

A GS2SLS and GS3SLS Estimation

Let $Y = (y_1, \ldots, y_m)$, $\bar{Y} = GY$, $\bar{X} = GX$, and $E = (\epsilon_1, \ldots, \epsilon_m)$. Then, from (7), we have

$$Y = Y\Phi + \bar{Y}\Lambda + XB + \bar{X}\Gamma + LA + E,$$

(39)

where $\Phi = [\phi_{ik}]$, $\Lambda = [\lambda_{ik}]$, $B = (\beta_1, \ldots, \beta_m)$, $\Gamma = (\gamma_1, \ldots, \gamma_m)$ and $A = (\alpha_1, \ldots, \alpha_m)$ are parameter matrices of dimension $m \times m$, $m \times m$, $p_x \times m$, $p_x \times m$ and $r \times m$ respectively. To eliminate the incidental parameter matrix $A$, we transform (39) with the projector $J = \text{diag}\{J_r\}_{r=1}^m$, where $J_r = I_{n_r} - \frac{1}{n_r}l_n l_n'$. As $JL = 0$, the transformed model is

$$JY = JY\Phi + J\bar{Y}\Lambda + JXB + J\bar{X}\Gamma + JE.$$ 

(40)

As discussed in Section 4, in order to achieve identification, we need to impose some exclusion restrictions on (40). Let $\phi_k$, $\lambda_k$, $\beta_k$ and $\gamma_k$ be vectors of nonzero elements of the $k$-th columns of $\Phi$, $\Lambda$, $B$ and $\Gamma$. Let $Y_k$, $\bar{Y}_k$, $X_k$ and $\bar{X}_k$ be matrices of the corresponding columns of $Y$, $\bar{Y}$, $X$ and $\bar{X}$ that appear in the $k$-th equation. Then, from (40), we have

$$Jy_k = JZ_k \theta_k + J\epsilon_k,$$

(41)

for $k = 1, \ldots, m$, where $Z_k = (Y_k, \bar{Y}_k, X_k, \bar{X}_k)$ and $\theta_k = (\phi_k', \lambda_k', \beta_k', \gamma_k')'$. The transformed model (41) can be estimated by the GS2SLS and GS3SLS estimators in Kelejian and Prucha (2004). Let $Q$ denote a (nonstochastic) IV matrix. Following Kelejian and Prucha (2004), we make the following assumptions.

Assumption 1 The diagonal elements of $G$ are zeros.

Assumption 2 $(I_{mn} - \Phi' \otimes I_n - \Lambda' \otimes G)$ is nonsingular.

Assumption 3 The row and column sums of $G$ and $(I_{mn} - \Phi' \otimes I_n - \Lambda' \otimes G)^{-1}$ are bounded uniformly in absolute value.
Assumption 4 \( X \) has full column rank (for a sufficiently large). In addition, the elements of \( X \) are uniformly bounded in absolute value.

Assumption 5 \( \epsilon \equiv (\epsilon_1', \cdots, \epsilon_m')' = (\Sigma_i' \otimes I_n)\text{vec}(V) \), where \( \Sigma_i \) is a nonsingular \( m \times m \) matrix and \( V = [v_{ik}] \) is an \( n \times m \) matrix of i.i.d. innovations with zero mean, unit variance and finite fourth moments. Furthermore, the diagonal elements of \( \Sigma \equiv \Sigma_i' \Sigma_i \) are bounded by some finite constant.

Assumption 6 The elements of the IV matrix \( Q \) are uniformly bounded in absolute value. Furthermore, \( H_{QQ} = \lim_{n \to \infty} \frac{1}{n} Q' J Q \) is a finite nonsingular matrix and \( H_{QZ,k} = \lim_{n \to \infty} \frac{1}{n} Q' J E(Z_k) \) is a finite matrix with full column rank for \( k = 1, \cdots, m \).

Let \( P = J Q (Q' J Q)^{-1} Q' J \) and \( \tilde{Z}_k = P Z_k \). The GS2SLS estimator for \( \theta_k \) can be written as

\[
\hat{\theta}_{gs2sls,k} = (\tilde{Z}_k' Z_k)^{-1} \tilde{Z}_k' y_k = (Z_k' P Z_k)^{-1} Z_k' P y_k.
\]

Under Assumptions 1-6, the asymptotic distribution of \( \hat{\theta}_{gs2sls,k} \) is

\[
\sqrt{n}(\hat{\theta}_{gs2sls,k} - \theta_k) \xrightarrow{d} N(0, \sigma_{kk}(H_{QZ,k}^{-1} H_{QZ,k})^{-1}).
\]

Let \( \tilde{\epsilon}_k = J y_k - J Z_k \hat{\theta}_{gs2sls,k} \) and \( \sigma_{kk} = \tilde{\epsilon}_k' \tilde{\epsilon}_k / (n - \bar{r}) \). The small sample distribution of \( \hat{\theta}_{gs2sls} \) can be approximated by \( N(\theta_k, \sigma_{kk}(Z_k' P Z_k)^{-1}) \).

The GS2SLS estimator is not efficient as it does not take into account the cross-equation correlation in the disturbances. To utilize the full system information, Kelejian and Prucha (2004) propose a GS3SLS estimator. Let \( y = (y_1', \cdots, y_m')' \), \( z = \text{diag}\{Z_k\}_{k=1}^m \), \( \epsilon = (\epsilon_1', \cdots, \epsilon_m')' \), and \( \theta = (\theta_1', \cdots, \theta_m')' \). Then, for the system of \( m \) equations,

\[
y^* = Z^* \theta + \epsilon^*,
\]

where \( y^* = (I_m \otimes J) y \), \( Z^* = (I_m \otimes J) Z \), and \( \epsilon^* = (I_m \otimes J) \epsilon \).

Let \( \tilde{Z}^* = (I_m \otimes P) Z^* \). As \( E(\epsilon^* \epsilon^*) = (I_m \otimes J) E(\epsilon \epsilon') (I_m \otimes J) = (I_m \otimes J)(\Sigma \otimes I_n)(I_m \otimes J) = \Sigma \otimes J \), the (infeasible) GS3SLS is given by

\[
\tilde{\theta}_{gs3sls} = [\tilde{Z}^* (\Sigma^{-1} \otimes J^-) Z^*]^{-1} \tilde{Z}^* (\Sigma^{-1} \otimes J^-) y^* = [Z' (\Sigma^{-1} \otimes P) Z]^{-1} Z' (\Sigma^{-1} \otimes P) y,
\]
where $J^-$ denotes the generalized inverse of $J$. To estimate $\Sigma$, let $\hat{\epsilon}_k^* = Jy_k - JZ_k\hat{\theta}_g2sls,k$. Then, $\Sigma$ can be estimated by $\hat{\Sigma} = [\hat{\sigma}_{kl}]$ with its $(k, l)$-th element being $\hat{\sigma}_{kl} = \hat{\epsilon}_k^*\hat{\epsilon}_l^*/(n - \bar{r})$ for $k, l = 1, \ldots, m$. Hence, the feasible GS3SLS is given by

$$\hat{\theta}_{g3sls} = [Z'(\hat{\Sigma}^{-1} \otimes P)Z]^{-1}Z'(\hat{\Sigma}^{-1} \otimes P)y.$$ 

Let $H_{QZ} = \text{diag}\{H_{QZ,k}\}_{k=1}^m$. Under Assumptions 1-6, the asymptotic distribution of $\hat{\theta}_{g3sls,k}$ is

$$\sqrt{n}(\hat{\theta}_{g3sls,k} - \theta_k) \overset{d}{\rightarrow} N(0, [H'_{QZ}(\Sigma^{-1} \otimes H_{QQ}^{-1})H_{QZ}]^{-1}).$$

The small sample distribution of $\hat{\theta}_{g3sls}$ can be approximated by $N(\theta, [Z'(\hat{\Sigma}^{-1} \otimes P)Z]^{-1})$.

**B Proofs**

**Proof of Proposition 1.** $E(JZ_k) = [E(JGy_k), JX, JG]\) has full column rank if and only if

$$E(JGy_k)d_1 + JXd_2 + JGXd_3 = 0 \quad (42)$$

implies that $d_1 = d_2 = d_3 = 0$. From (19), we have

$$E[JG(I - \lambda_{kk}G)y_k] = JGX\beta_k + JG^2X\gamma_k. \quad (43)$$

Since $J(I - \lambda_{kk}G)J = J(I - \lambda_{kk}G)$ and $(I - \lambda_{kk}G)G = G(I - \lambda_{kk}G)$, if we premultiply (42) by $J(I - \lambda_{kk}G)$, we have

$$E[JG(I - \lambda_{kk}G)y_k]d_1 + J(I - \lambda_{kk}G)Xd_2 + J(I - \lambda_{kk}G)GXd_3 = 0. \quad (44)$$

If we substitute (43) into (44) and rearrange terms, we have

$$JXd_2 + JGX(\beta_kd_1 - \lambda_{kk}d_2 + d_3) + JG^2X(\gamma_kd_1 - \lambda_{kk}d_3) = 0. \quad (45)$$

As $J, JG, JG^2$ are linearly independent, (45) implies $d_2 = 0$, $\beta_kd_1 - \lambda_{kk}d_2 + d_3 = 0$ and $\gamma_kd_1 - \lambda_{kk}d_3 = 0$, which in turn implies that $d_1 = d_2 = d_3 = 0$ if $\lambda_{kk}\beta_k + \gamma_k \neq 0$. ■
Lemma B.1 \( J, JG, JG^2 \) are linearly independent if \( I, G, G^2, G^3 \) are linearly independent.

Proof of Lemma B.1. \( J, JG, JG^2 \) are linearly independent if and only if

\[
d_1J + d_2JG + d_3JG^2 = 0
\]  \hspace{1cm} (46)

implies that \( d_1 = d_2 = d_3 = 0 \). (46) holds if and only if

\[
d_1I + d_2G + d_3G^2 = lc'
\]  \hspace{1cm} (47)

for some \( n \times 1 \) vector of constants \( c \). As \( Gl = l \), if we premultiply (47), we have

\[
d_1G + d_2G^2 + d_3G^3 = lc'.
\]  \hspace{1cm} (48)

From (47) and (48), we have

\[
d_1I + (d_2 - d_1)G + (d_3 - d_2)G^2 - d_3G^3 = 0.
\]  \hspace{1cm} (49)

If \( I, G, G^2, G^3 \) are linearly independent, (49) implies \( d_1 = d_2 - d_1 = d_3 - d_2 = d_3 = 0 \), which in turn implies that \( d_1 = d_2 = d_3 = 0 \). \( \blacksquare \)

Proof of Proposition 2. \( E(JZ_1) = [E(Jy_2), E(JGY_1), JX, JGX] \) has a full column rank if and only if

\[
E(Jy_2)d_1 + E(JGY_1)d_2 + JXd_3 + JGXd_4 = 0
\]  \hspace{1cm} (50)

implies that \( d_1 = d_2 = d_3 = d_4 = 0 \). Observe that \( JGJ = JG, JS_1J = JS_1 \) and \( S_1G = GS_1 \). If we premultiply both sides of (22) and (23) by \( JS_1 \), we have

\[
E(JS_1y_1) = JX(\phi_{21}\beta_2 + \beta_1) + JGX(\phi_{21}\gamma_2 + \gamma_1 - \lambda_{22}\beta_1) - JG^2X\lambda_{22}\gamma_1
\]  \hspace{1cm} (51)

\[
E(JS_1y_2) = JX(\phi_{12}\beta_1 + \beta_2) + JGX(\phi_{12}\gamma_1 + \gamma_2 - \lambda_{11}\beta_2) - JG^2X\lambda_{11}\gamma_2.
\]  \hspace{1cm} (52)

If we premultiply (50) by \( JS_1 \), then it follows from (51) and (52) that

\[
JX\eta_1 + JGX\eta_2 + JG^2X\eta_3 + JG^3X\eta_4 = 0,
\]

26
where

\[
\begin{align*}
\eta_1 &= (\phi_{12}\beta_1 + \beta_2)d_1 + (1 - \phi_{12}\phi_{21})d_3 \\
\eta_2 &= (\phi_{12}\gamma_1 + \gamma_2 - \lambda_{11}\beta_2)d_1 + (\beta_1 + \phi_{21}\beta_2)d_2 - (\lambda_{11} + \lambda_{22})d_3 + (1 - \phi_{12}\phi_{21})d_4 \\
\eta_3 &= -\lambda_{11}\gamma_2d_1 + (\phi_{21}\gamma_2 + \gamma_1 - \lambda_{22}\beta_1)d_2 + \lambda_{11}\lambda_{22}d_3 - (\lambda_{11} + \lambda_{22})d_4 \\
\eta_4 &= -\lambda_{22}\gamma_1d_2 + \lambda_{11}\lambda_{22}d_4
\end{align*}
\]

If $J, JG, JG^2, JG^3$ are linearly independent, then $\eta_1 = \eta_2 = \eta_3 = \eta_4 = 0$, which in turn implies that $d_1 = d_2 = d_3 = d_4 = 0$ if $A_1$ given by (24) has full rank.

Proof of Proposition 3. $E(JZ_1) = [E(JGy_1), E(JGy_2), JX, JGX]$ has a full column rank if and only if

\[
E(JGy_1)d_1 + E(JGy_2)d_2 + JXd_3 + JGXd_4 = 0 \quad (53)
\]

implies that $d_1 = d_2 = d_3 = d_4 = 0$. Observe that $JGJ = JG, JS_2J = JS_2$ and $S_2G = GS_2$. If we premultiply both sides of (27) and (28) by $JS_2$, we have

\[
\begin{align*}
E(JS_2y_1) &= JX\beta_1 + JGX(\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \gamma_1) + JG^2X(\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1) \\
E(JS_2y_2) &= JX\beta_2 + JGX(\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \gamma_2) + JG^2X(\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2).
\end{align*} \quad (54) (55)
\]

If we premultiply (53) by $JS_2$, then it follows by (54) and (55) that

\[
JX\eta_1 + JGX\eta_2 + JG^2X\eta_3 + JG^3X\eta_4 = 0,
\]

where

\[
\begin{align*}
\eta_1 &= d_3 \\
\eta_2 &= \beta_1d_1 + \beta_2d_2 - (\lambda_{11} + \lambda_{22})d_3 + d_4 \\
\eta_3 &= (\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \gamma_1)d_1 + (\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \gamma_2)d_2 + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})d_3 - (\lambda_{11} + \lambda_{22})d_4 \\
\eta_4 &= (\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1)d_1 + (\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2)d_2 + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})d_4.
\end{align*}
\]

If $J, JG, JG^2, JG^3$ are linearly independent, then $\eta_1 = \eta_2 = \eta_3 = \eta_4 = 0$, which in turn implies
that $d_1 = d_2 = d_3 = d_4 = 0$ if $A_2$ given by (29) has full rank. ■

**Proof of Proposition 4.** $E(JZ_1) = [E(Jy_2), E(JGy_1), E(JGy_2), JX, JG]$ has full column rank if and only if

\[
E(Jy_2)d_1 + E(JGy_1)d_2 + E(JGy_2)d_3 + JXd_4 + JGd_5 = 0 \tag{56}
\]

implies that $d_1 = d_2 = d_3 = d_4 = d_5 = 0$. Observe that $JG = G$, $JS = S$ and $SG = GS$. If we premultiply both sides of (11) and (12) by $JS$, we have

\[
E(JSy_1) = JX(\phi_{21}\beta_2 + \beta_1) + JGX(\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \phi_{21}\gamma_2 + \gamma_1) + JG^2X(\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1) \tag{57}
\]

\[
E(JSy_2) = JX(\phi_{12}\beta_1 + \beta_2) + JGX(\lambda_{12}\beta_2 - \lambda_{11}\beta_2 + \phi_{12}\gamma_1 + \gamma_2) + JG^2X(\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2) \tag{58}
\]

If we premultiply (56) by $JS$, then it follows by (57) and (58) that

\[
JX\eta_1 + JGX\eta_2 + JG^2X\eta_3 + JG^3X\eta_4 = 0,
\]

where

\[
\begin{align*}
\eta_1 &= (\phi_{12}\beta_1 + \beta_2)d_1 + (1 - \phi_{12}\phi_{21})d_4 \\
\eta_2 &= (\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \phi_{12}\gamma_1 + \gamma_2)d_1 + (\beta_1 + \phi_{21}\beta_2)d_2 + (\phi_{12}\beta_1 + \beta_2)d_3 \\
&\quad - (\lambda_{11} + \lambda_{22} + \lambda_{12}\phi_{21} + \lambda_{21}\phi_{12})d_4 + (1 - \phi_{12}\phi_{21})d_5 \\
\eta_3 &= (\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2)d_1 + (\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \phi_{21}\gamma_2 + \gamma_1)d_2 + (\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \phi_{12}\gamma_1 + \gamma_2)d_3 \\
&\quad + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})d_4 - (\lambda_{11} + \lambda_{22} + \lambda_{12}\phi_{21} + \lambda_{21}\phi_{12})d_5 \\
\eta_4 &= (\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1)d_2 + (\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2)d_3 + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})d_5
\end{align*}
\]

If $J$, $JG$, $JG^2$, $JG^3$ are linearly independent, then $\eta_1 = \eta_2 = \eta_3 = \eta_4 = 0$, which in turn implies that $d_2 = (\lambda_{12} + \lambda_{11}\phi_{12})d_1/(\phi_{12}\phi_{21} - 1)$, $d_3 = (\lambda_{22} + \lambda_{21}\phi_{12})d_1/(\phi_{12}\phi_{21} - 1)$, $d_4 = (\phi_{12}\beta_1 + \beta_2)d_1/(\phi_{12}\phi_{21} - 1)$ and $d_5 = (\phi_{12}\gamma_1 + \gamma_2)d_1/(\phi_{12}\phi_{21} - 1)$. Therefore, $E(JZ_1)$ does not have full column rank. Similarly, $E(JZ_2)$ does not have full column rank. ■

**Proof of Proposition 5.** $E(JZ_1) = [E(Jy_2), E(JGy_1), E(JGy_2), JX_1, JG_1]$ has a full
column rank if and only if

\[ E(Jy_2)d_1 + E(JGy_1)d_2 + E(JGy_2)d_3 + JX_1d_4 + JGX_1d_5 = 0 \] (59)

implies that \( d_1 = d_2 = d_3 = d_4 = d_5 = 0 \). From (35) and (36), the reduced-form equations of the within-transformed model are

\[ E(JSy_1) = JX_2\phi_{21}\beta_2 + JGX_2(\lambda_{21}\beta_2 + \phi_{21}\gamma_2) + JG^2X_2\lambda_{21}\gamma_2 + JX_1\beta_1 + JGX_1(\gamma_1 - \lambda_{22}\beta_1) - JG^2X_1\lambda_{22}\gamma_1 \] (60)

\[ E(JSy_2) = JX_1\phi_{12}\beta_1 + JGX_1(\lambda_{12}\beta_1 + \phi_{12}\gamma_1) + JG^2X_1\lambda_{12}\gamma_1 + JX_2\beta_2 + JGX_2(\gamma_2 - \lambda_{11}\beta_2) - JG^2X_2\lambda_{11}\gamma_2 \] (61)

Observe that \( JGJ = JG, \ JSJ = JS \) and \( SG = GS \). If we premultiply (59) by \( JS \), it follows by (60) and (61) that

\[
X_1\eta_1 + GX_1\eta_2 + G^2X_1\eta_3 + G^3X_1\eta_4 \\
+ X_2\xi_1 + GX_2\xi_2 + G^2X_2\xi_3 + G^3X_2\xi_4 = 0,
\]

where

\[
\begin{align*}
\eta_1 &= \phi_{12}\beta_1d_1 + (1 - \phi_{12}\phi_{21})d_4 \\
\eta_2 &= (\lambda_{12}\beta_1 + \phi_{12}\gamma_1)d_1 + \beta_1d_2 + \phi_{12}\beta_1d_3 - (\lambda_{11} + \lambda_{22} + \lambda_{12}\phi_{21} + \lambda_{21}\phi_{12})d_4 + (1 - \phi_{12}\phi_{21})d_5 \\
\eta_3 &= \lambda_{12}\gamma_1d_1 + (\gamma_1 - \lambda_{22}\beta_1)d_2 + (\lambda_{12}\beta_1 + \phi_{12}\gamma_1)d_3 + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})d_4 - (\lambda_{11} + \lambda_{22} + \lambda_{12}\phi_{21} + \lambda_{21}\phi_{12})d_5 \\
\eta_4 &= -\lambda_{22}\gamma_1d_2 + \lambda_{12}\gamma_1d_3 + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})d_5
\end{align*}
\]

and

\[
\begin{align*}
\xi_1 &= \beta_2d_1 \\
\xi_2 &= (\gamma_2 - \lambda_{11}\beta_2)d_1 + \phi_{21}\beta_2d_2 + \beta_2d_3 \\
\xi_3 &= -\lambda_{11}\gamma_2d_1 + (\lambda_{21}\beta_2 + \phi_{21}\gamma_2)d_2 + (\gamma_2 - \lambda_{11}\beta_2)d_3 \\
\xi_4 &= \lambda_{21}\gamma_2d_2 - \lambda_{11}\gamma_2d_3.
\end{align*}
\]
We consider two cases:

Case 1, if $J, JG, JG^2, JG^3$ are linearly independent, then

$$\eta_1 = \eta_2 = \eta_3 = \eta_4 = \xi_1 = \xi_2 = \xi_3 = \xi_4 = 0.$$ 

In this case, $d_1 = d_2 = d_3 = d_4 = 0$ if $A_3$ given by (37) has full rank.

Case 2, if $J, JG, JG^2$ are linearly independent and $JG^3 = \rho_1 J + \rho_2 JG + \rho_3 JG^2$, then

$$\eta_1 + \rho_1 \eta_4 = \eta_2 + \rho_2 \eta_4 = \eta_3 + \rho_3 \eta_4 = \xi_1 + \rho_1 \xi_4 = \xi_2 + \rho_2 \xi_4 = \xi_3 + \rho_3 \xi_4 = 0.$$ 

In this case, $d_1 = d_2 = d_3 = d_4 = 0$ if $A_3^*$ given by (38) has full rank. ■
### Table 1: List of Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Effort</td>
<td>In the text</td>
<td>2.25</td>
<td>0.66</td>
</tr>
<tr>
<td>TV</td>
<td>In the text</td>
<td>2.29</td>
<td>1.08</td>
</tr>
<tr>
<td>Age</td>
<td>Age</td>
<td>14.97</td>
<td>1.69</td>
</tr>
<tr>
<td>Female</td>
<td>1 if the respondent is female</td>
<td>0.53</td>
<td>0.50</td>
</tr>
<tr>
<td>White</td>
<td>1 if the respondent is white</td>
<td>0.57</td>
<td>0.50</td>
</tr>
<tr>
<td>Born in the U.S</td>
<td>1 if the respondent is born in the U.S.</td>
<td>0.92</td>
<td>0.27</td>
</tr>
<tr>
<td>Years in school</td>
<td>The number of years in the current school</td>
<td>2.55</td>
<td>1.44</td>
</tr>
<tr>
<td>Live with both parents</td>
<td>1 if the respondent lives with both parents</td>
<td>0.74</td>
<td>0.44</td>
</tr>
<tr>
<td>Parental care</td>
<td>1 if the respondent thinks parents care about him/her very much</td>
<td>0.84</td>
<td>0.36</td>
</tr>
<tr>
<td>Neighborhood safety</td>
<td>1 if the respondent feels safe in his/her neighborhood</td>
<td>0.71</td>
<td>0.45</td>
</tr>
<tr>
<td>School safety</td>
<td>1 if the respondent feels safe in his/her school</td>
<td>0.60</td>
<td>0.49</td>
</tr>
<tr>
<td>Parent job:</td>
<td>1 if the parent’s job is a doctor, lawyer, scientist, teacher, librarian,</td>
<td>0.28</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>manager, executive, or director</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other jobs</td>
<td>1 if the parent’s job is not listed above</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>(Stay home)</td>
<td>1 if the parent is a homemaker, retired, or does not work</td>
<td>0.09</td>
<td>0.29</td>
</tr>
<tr>
<td>Missing</td>
<td>1 if the parent’s job information is missing</td>
<td>0.09</td>
<td>0.29</td>
</tr>
<tr>
<td>Parent education:</td>
<td>1 if the parent’s education is less than high school (HS)</td>
<td>0.11</td>
<td>0.32</td>
</tr>
<tr>
<td>(Less than HS)</td>
<td>1 if the parent’s education is HS or higher but no college degree</td>
<td>0.45</td>
<td>0.50</td>
</tr>
<tr>
<td>HS grad</td>
<td>1 if the parent’s education is college or higher</td>
<td>0.32</td>
<td>0.47</td>
</tr>
<tr>
<td>College grad</td>
<td>1 if the parent’s education information is missing</td>
<td>0.12</td>
<td>0.32</td>
</tr>
<tr>
<td>Missing</td>
<td>1 if the respondent’s health condition is excellent</td>
<td>0.31</td>
<td>0.46</td>
</tr>
<tr>
<td>TV program</td>
<td>1 if the respondent is allowed to choose TV programs to watch</td>
<td>0.15</td>
<td>0.35</td>
</tr>
</tbody>
</table>

The variable in the parentheses is the reference category.

If both parents are in the household, the education and job of the father is considered.
<table>
<thead>
<tr>
<th></th>
<th>Study effort</th>
<th>TV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GS2SLS</td>
<td>GS3SLS</td>
</tr>
<tr>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endogenous peer effect</td>
<td>0.0771</td>
<td>0.0712</td>
</tr>
<tr>
<td></td>
<td>(0.0530)</td>
<td>(0.0530)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0077*</td>
<td>-0.0077*</td>
</tr>
<tr>
<td></td>
<td>(0.0042)</td>
<td>(0.0042)</td>
</tr>
<tr>
<td>Female</td>
<td>0.2098***</td>
<td>0.2098***</td>
</tr>
<tr>
<td></td>
<td>(0.0057)</td>
<td>(0.0057)</td>
</tr>
<tr>
<td>White</td>
<td>-0.0226***</td>
<td>-0.0226***</td>
</tr>
<tr>
<td></td>
<td>(0.0067)</td>
<td>(0.0067)</td>
</tr>
<tr>
<td>Born in the U.S.</td>
<td>-0.0715***</td>
<td>-0.0718***</td>
</tr>
<tr>
<td></td>
<td>(0.0108)</td>
<td>(0.0108)</td>
</tr>
<tr>
<td>Years in school</td>
<td>-0.0136***</td>
<td>-0.0136***</td>
</tr>
<tr>
<td></td>
<td>(0.0027)</td>
<td>(0.0027)</td>
</tr>
<tr>
<td>Live with both parents</td>
<td>0.0310***</td>
<td>0.0311***</td>
</tr>
<tr>
<td></td>
<td>(0.0063)</td>
<td>(0.0063)</td>
</tr>
<tr>
<td>Parental care</td>
<td>0.1839***</td>
<td>0.1840***</td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.0074)</td>
</tr>
<tr>
<td>Neighborhood safety</td>
<td>0.0262***</td>
<td>0.0262***</td>
</tr>
<tr>
<td></td>
<td>(0.0064)</td>
<td>(0.0064)</td>
</tr>
<tr>
<td>School safety</td>
<td>0.0674***</td>
<td>0.0674***</td>
</tr>
<tr>
<td></td>
<td>(0.0059)</td>
<td>(0.0059)</td>
</tr>
<tr>
<td>Parent job: professional</td>
<td>-0.0089</td>
<td>-0.0090</td>
</tr>
<tr>
<td></td>
<td>(0.0100)</td>
<td>(0.0100)</td>
</tr>
<tr>
<td>Parent job: other jobs</td>
<td>-0.0133</td>
<td>-0.0134</td>
</tr>
<tr>
<td></td>
<td>(0.0089)</td>
<td>(0.0089)</td>
</tr>
<tr>
<td>Parent job: missing</td>
<td>0.0732***</td>
<td>0.0731***</td>
</tr>
<tr>
<td></td>
<td>(0.0124)</td>
<td>(0.0124)</td>
</tr>
<tr>
<td>Parent education: HS grad</td>
<td>0.0250***</td>
<td>0.0250***</td>
</tr>
<tr>
<td></td>
<td>(0.0085)</td>
<td>(0.0085)</td>
</tr>
<tr>
<td>Parent education: college grad</td>
<td>0.0551***</td>
<td>0.0552***</td>
</tr>
<tr>
<td></td>
<td>(0.0095)</td>
<td>(0.0095)</td>
</tr>
<tr>
<td>Parent education: missing</td>
<td>0.0743***</td>
<td>0.0744***</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
<td>(0.0111)</td>
</tr>
<tr>
<td>Health</td>
<td>0.1738***</td>
<td>0.1739***</td>
</tr>
<tr>
<td></td>
<td>(0.0055)</td>
<td>(0.0055)</td>
</tr>
<tr>
<td>TV program</td>
<td>-0.0148***</td>
<td>-0.0149***</td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.0074)</td>
</tr>
</tbody>
</table>

**Selected contextual effects G**

<table>
<thead>
<tr>
<th></th>
<th>Study effort</th>
<th>TV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GS2SLS</td>
<td>GS3SLS</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0023</td>
<td>-0.0017</td>
</tr>
<tr>
<td></td>
<td>(0.0055)</td>
<td>(0.0055)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.0347**</td>
<td>-0.0333**</td>
</tr>
<tr>
<td></td>
<td>(0.0156)</td>
<td>(0.0156)</td>
</tr>
<tr>
<td>White</td>
<td>-0.0340***</td>
<td>-0.0344***</td>
</tr>
<tr>
<td></td>
<td>(0.0101)</td>
<td>(0.0101)</td>
</tr>
<tr>
<td>Born in the U.S.</td>
<td>-0.0835***</td>
<td>-0.0831***</td>
</tr>
<tr>
<td></td>
<td>(0.0157)</td>
<td>(0.0157)</td>
</tr>
<tr>
<td>Live with both parents</td>
<td>0.0449***</td>
<td>0.0453***</td>
</tr>
<tr>
<td></td>
<td>(0.0114)</td>
<td>(0.0114)</td>
</tr>
</tbody>
</table>

To save space, only selected contextual effects that are statistically significant are reported.
Standard errors in parentheses. Statistical significance: ***p<0.01; **p<0.05; *p<0.1.
Table 3: GS3SLS Estimation of Simultaneous Equations Network Models

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
<th>Model 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study effort</td>
<td>TV</td>
<td>Study effort</td>
<td>TV</td>
<td>Study effort</td>
<td>TV</td>
</tr>
<tr>
<td>Simultaneity effect</td>
<td>-0.2211***</td>
<td>-0.5502***</td>
<td>(0.0502)</td>
<td>(0.1935)</td>
<td>-0.1316*</td>
<td>-0.5210***</td>
</tr>
<tr>
<td></td>
<td>0.1344***</td>
<td>0.3161***</td>
<td>(0.0527)</td>
<td>(0.0419)</td>
<td>0.1939***</td>
<td>0.3061***</td>
</tr>
<tr>
<td>Endogenous peer effect</td>
<td>-0.1140***</td>
<td>-0.3323***</td>
<td>(0.0291)</td>
<td>(0.1014)</td>
<td>-0.0716*</td>
<td>-0.0332</td>
</tr>
<tr>
<td>Cross-activity peer effect</td>
<td>0.0263***</td>
<td>0.0260**</td>
<td>(0.0196)</td>
<td>(0.0125)</td>
<td>0.0218*</td>
<td>-0.0095</td>
</tr>
</tbody>
</table>

X

<table>
<thead>
<tr>
<th></th>
<th>Study effort</th>
<th>TV</th>
<th>Study effort</th>
<th>TV</th>
<th>Study effort</th>
<th>TV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.0178***</td>
<td>-0.0461***</td>
<td>(0.0049)</td>
<td>(0.0071)</td>
<td>-0.0144***</td>
<td>-0.0466***</td>
</tr>
<tr>
<td>Female</td>
<td>0.1876***</td>
<td>0.0180</td>
<td>(0.0076)</td>
<td>(0.0416)</td>
<td>0.1958***</td>
<td>0.0121</td>
</tr>
<tr>
<td>White</td>
<td>-0.0711***</td>
<td>-0.2134***</td>
<td>(0.0130)</td>
<td>(0.0123)</td>
<td>-0.0551***</td>
<td>-0.2133***</td>
</tr>
<tr>
<td>Born in the U.S.</td>
<td>-0.0413***</td>
<td>0.0746***</td>
<td>(0.0127)</td>
<td>(0.0222)</td>
<td>-0.0477***</td>
<td>0.0751***</td>
</tr>
<tr>
<td>Years in school</td>
<td>-0.0131***</td>
<td>-0.0041</td>
<td>(0.0028)</td>
<td>(0.0052)</td>
<td>-0.0135***</td>
<td>-0.0038</td>
</tr>
<tr>
<td>Live with both parents</td>
<td>0.0123</td>
<td>-0.0558***</td>
<td>(0.0076)</td>
<td>(0.0118)</td>
<td>0.0170**</td>
<td>-0.0563***</td>
</tr>
<tr>
<td>Parental care</td>
<td>0.1952***</td>
<td>0.1518***</td>
<td>(0.0079)</td>
<td>(0.0378)</td>
<td>0.1907***</td>
<td>0.1468***</td>
</tr>
<tr>
<td>Neighborhood safety</td>
<td>0.0201***</td>
<td>-0.0076</td>
<td>(0.0066)</td>
<td>(0.0115)</td>
<td>0.0212***</td>
<td>-0.0084</td>
</tr>
<tr>
<td>School safety</td>
<td>0.0482***</td>
<td>-0.0488***</td>
<td>(0.0074)</td>
<td>(0.0378)</td>
<td>0.0558***</td>
<td>-0.0505***</td>
</tr>
<tr>
<td>Parent job: professional</td>
<td>-0.0186*</td>
<td>-0.0483***</td>
<td>(0.0104)</td>
<td>(0.0163)</td>
<td>-0.0147</td>
<td>-0.0484***</td>
</tr>
<tr>
<td>Parent job: other jobs</td>
<td>-0.0073</td>
<td>0.0102</td>
<td>(0.0091)</td>
<td>(0.0147)</td>
<td>-0.0073</td>
<td>0.0103</td>
</tr>
<tr>
<td>Parent job: missing</td>
<td>0.0804***</td>
<td>0.0614***</td>
<td>(0.0126)</td>
<td>(0.0246)</td>
<td>0.0803***</td>
<td>0.0593***</td>
</tr>
<tr>
<td>Parent education: HS grad</td>
<td>0.0188***</td>
<td>-0.0143</td>
<td>(0.0087)</td>
<td>(0.0146)</td>
<td>0.0214***</td>
<td>-0.0149</td>
</tr>
<tr>
<td>Parent education: college grad</td>
<td>0.0252***</td>
<td>-0.0920***</td>
<td>(0.0117)</td>
<td>(0.0186)</td>
<td>0.0341***</td>
<td>-0.0932***</td>
</tr>
<tr>
<td>Parent education: missing</td>
<td>0.0642***</td>
<td>-0.0005</td>
<td>(0.0115)</td>
<td>(0.0232)</td>
<td>0.0673***</td>
<td>-0.0020</td>
</tr>
<tr>
<td>Health</td>
<td>0.1520***</td>
<td>0.0025</td>
<td>(0.0074)</td>
<td>(0.0349)</td>
<td>0.1598***</td>
<td>0.0803</td>
</tr>
<tr>
<td>TV program</td>
<td>-0.0065</td>
<td>0.0272**</td>
<td>(0.0078)</td>
<td>(0.0124)</td>
<td>0.0311***</td>
<td>0.0311***</td>
</tr>
</tbody>
</table>

Selected contextual effects GX

<table>
<thead>
<tr>
<th></th>
<th>Study effort</th>
<th>TV</th>
<th>Study effort</th>
<th>TV</th>
<th>Study effort</th>
<th>TV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.0064</td>
<td>-0.0310***</td>
<td>(0.0054)</td>
<td>(0.0064)</td>
<td>-0.0037</td>
<td>-0.0265***</td>
</tr>
<tr>
<td>Female</td>
<td>-0.0520***</td>
<td>-0.0274*</td>
<td>(0.0157)</td>
<td>(0.0146)</td>
<td>-0.0650***</td>
<td>-0.0213</td>
</tr>
<tr>
<td>White</td>
<td>-0.0807***</td>
<td>-0.1179***</td>
<td>(0.0151)</td>
<td>(0.0261)</td>
<td>-0.0855***</td>
<td>-0.1233***</td>
</tr>
<tr>
<td>Born in the U.S.</td>
<td>-0.0415**</td>
<td>0.0447</td>
<td>(0.0188)</td>
<td>(0.0307)</td>
<td>0.0374**</td>
<td>0.0531*</td>
</tr>
<tr>
<td>Live with both parents</td>
<td>0.0263***</td>
<td>-0.0103</td>
<td>(0.0121)</td>
<td>(0.0196)</td>
<td>0.0218*</td>
<td>-0.0095</td>
</tr>
</tbody>
</table>

To save space, only selected contextual effects that are statistically significant are reported. Standard errors in parentheses. Statistical significance: ***p<0.01; **p<0.05; *p<0.1.