

Multivariate Choices and Identification of Social Interactions*

Ethan Cohen-Cole[†] Xiaodong Liu[‡] Yves Zenou[§]

April 2017

Abstract

This paper considers the identification of social interaction effects in the context of multivariate choices. First, we generalize the theoretical social interaction model to allow individuals to make interdependent choices in different activities. Based on the theoretical model, we propose a simultaneous-equation network model and discuss the identification of social interaction effects in the econometric model. We also provide an empirical example to show the empirical salience of this model. Using the Add Health data, we find that a student's academic performance is not only affected by academic performance of her peers but also affected by screen-related activities of her peers.

Key words: Identification, peer effects, simultaneous equations, social networks

JEL Classification: C31, C36, I21, Z13.

*We thank the co-editor Thierry Magnac and three 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.

[§]Monash University and IFN. E-mail: yves.zenou@monash.edu.

1 Introduction

Peer choices and/or peer characteristics have been shown to be important in predicting individual outcomes, ranging from education and crime to participation in the labor market (see, e.g., Ioannides and Loury, 2004; Sacerdote, 2011; Patacchini and Zenou, 2012). Most of this literature has, however, focused on peer effects on choices regarding one specific activity.

In reality, individuals make a multitude of choices in different activities, many of which may depend on each other. As a result, an individual may have different and sometimes opposite influences on her friend. For example, if a student is very active in extracurricular activities but also studies very hard, how would these choices affect the study effort of her friends? The peer effects of interdependent choices is what we study in this paper. Our purpose is to help understand the decision making process involving multiple activities in the context of peer influences and social networks.

The contribution of this paper is threefold. First, we provide a microfoundation that helps characterize the decision making process in multiple activities in a social interaction setting. The theoretical model we consider has two important features. First, as is common in this literature (see, e.g., Ballester et al., 2006; Bramoullé and Kranton, 2007; Bramoullé et al., 2014; Jackson and Zenou, 2015), our model has the feature that individuals enjoy utility as a function of peers' choices. Second, our model allows individuals to make choices in multiple activities that have an arbitrary degree of complementarity or substitutability.¹ The model is general enough to encompass arbitrary combinations of choices without making assumptions regarding the orderings of choice bundles. This generality is essential because combining sets of choices into bundles in a social interaction context dramatically

¹Belhaj and Deroïan (2014) and Chen et al. (2015) develop a network model where two activities are considered. Both papers only analyze the theoretical implications of their respective models without addressing econometric issues.

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 considered here.

Second, we investigate the identification of peer effects in the context of multivariate choices. The econometric model implied by the best response function of the theoretical model extends the simultaneous-equation spatial autoregressive model introduced by Kelejian and Prucha (2004) to allow for network fixed effects. As single-activity social interaction models (e.g., Bramoullé et al., 2009; Lee et al., 2010), our model includes the *within-activity peer effect* (also known as the *endogenous peer effect* in single-activity social interaction models) where an individual’s choice in an activity may depend on the choices of her peers in the same activity; the *contextual effect*, where an individual’s choice may depend on the exogenous characteristics of her peers; and the *correlated effect*, where individuals in the same network tend to behave similarly because they have similar unobserved individual characteristics and/or face similar institutional environments. The *reflection problem* (Manski, 1993) is well known and emerges from the coexistence of these effects. Furthermore, an individual’s choice in a certain activity may depend on her own choices in related activities. This is the usual simultaneity effect that is endemic in simultaneous-equation models. To distinguish it from other types of simultaneity effects in our model, we call it the *self-simultaneity effect*. Finally, our model includes a new type of social interaction effect, the *cross-activity peer effect*, where an individual’s choice in an activity may depend on the choices of her peers in related activities. Following Bramoullé et al. (2009), we provide identification conditions for the different social interaction and simultaneity effects based on the topology of underlying networks.

Third, we test the empirical salience of this model. Using a representative sample of U.S. teenagers in the National Longitudinal Study of Adolescent Health (Add Health)

data, we find that a student’s academic performance is *positively* affected by academic performance of her peers and *negatively* affected by screen-related activities of her peers.

The rest of the paper is organized as follows. We develop a general theoretical model in Section 2. Section 3 presents the econometric model implied by the theoretical model and Section 4 discusses the identification of peer effects. Section 5 provides an empirical example. Finally, Section 6 concludes. The proofs are collected in the web appendix.

2 Theoretical Model

Suppose a finite set of individuals $\{1, \dots, n\}$ is connected by a network. We keep track of social connections in the network through an adjacency matrix $\mathbf{G} = [g_{ij}]$. Let n_i denote the number of direct connections of individual i . For ease of presentation, we assume $n_i > 0$ for all i . The (i, j) -th element of \mathbf{G} is given by $g_{ij} = 1/n_i$ if individuals i and j are connected and $g_{ij} = 0$ otherwise. We set $g_{ii} = 0$. We define the peers of individual i as the set of individuals connected to individual i , i.e. $\{j : g_{ij} > 0\}$. An example is given in Figure 1 for a star-shaped network with four individuals.

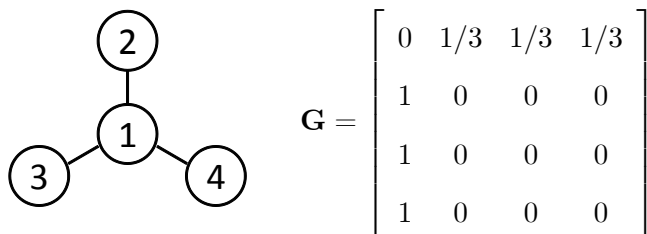


Figure 1: an example of \mathbf{G} for a star-shaped network.

In the network game, individuals choose their effort levels in two activities, denoted by $\mathbf{y}_1 = (y_{11}, \dots, y_{n1})'$ and $\mathbf{y}_2 = (y_{12}, \dots, y_{n2})'$, to maximize their utility. The utility of

individual i is a linear-quadratic function of the effort levels \mathbf{y}_1 and \mathbf{y}_2 given by

$$\begin{aligned}
 U_i(\mathbf{y}_1, \mathbf{y}_2) = & \underbrace{\varpi_{i1}y_{i1} + \varpi_{i2}y_{i2} + \sum_{j=1}^n g_{ij}(\varrho_{11}y_{j1}y_{i1} + \varrho_{12}y_{j1}y_{i2} + \varrho_{21}y_{j2}y_{i1} + \varrho_{22}y_{j2}y_{i2})}_{\text{payoff}} \\
 & - \underbrace{\frac{1}{2}(\varphi_{11}y_{i1}^2 + 2\varphi_{12}y_{i1}y_{i2} + \varphi_{22}y_{i2}^2)}_{\text{cost}}. \tag{2.1}
 \end{aligned}$$

As in the single-activity linear-quadratic utility function considered in Ballester et al. (2006), the utility given by (2.1) has two components: payoff and cost. The marginal payoff of individual i 's effort in activity k (for $k = 1, 2$) depends on (exogenous) attributes of individual i given by ϖ_{ik} and the average effort of her peers in the same and related activities given by $\sum_{l=1}^2 \varrho_{lk} \sum_{j=1}^n g_{ij}y_{jl}$. The parameter ϱ_{lk} (for $k, l = 1, 2$) captures the strategic substitutability or complementarity (depending on the sign of ϱ_{lk}) 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 both activities. The parameter φ_{12} measures the substitutability or complementarity (depending on the sign of φ_{12}) of an individual's effort levels in these two activities.

Given the network structure and effort levels of the peers, individual i chooses effort levels y_{i1} and y_{i2} to maximize the utility (2.1). From the first order condition, the equilibrium best response function is

$$y_{ik} = \phi_{lk}y_{il} + \lambda_{kk} \sum_{j=1}^n g_{ij}y_{jk} + \lambda_{lk} \sum_{j=1}^n g_{ij}y_{jl} + \pi_{ik}, \quad \text{for } k = 1, 2 \text{ and } l = 3 - k,$$

where $\phi_{lk} = -\varphi_{12}/\varphi_{kk}$, $\lambda_{kk} = \varrho_{kk}/\varphi_{kk}$, $\lambda_{lk} = \varrho_{lk}/\varphi_{kk}$, and $\pi_{ik} = \varpi_{ik}/\varphi_{kk}$. In matrix form, the equilibrium best response function is given by

$$\mathbf{y}_k = \phi_{lk}\mathbf{y}_l + \lambda_{kk}\mathbf{G}\mathbf{y}_k + \lambda_{lk}\mathbf{G}\mathbf{y}_l + \boldsymbol{\pi}_k, \quad \text{for } k = 1, 2 \text{ and } l = 3 - k, \tag{2.2}$$

where $\boldsymbol{\pi}_k = (\pi_{1k}, \dots, \pi_{nk})'$. Let

$$\mathbf{S} = (1 - \phi_{12}\phi_{21})\mathbf{I}_n - (\lambda_{11} + \lambda_{22} + \phi_{21}\lambda_{12} + \phi_{12}\lambda_{21})\mathbf{G} + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\mathbf{G}^2. \quad (2.3)$$

If $\phi_{12}\phi_{21} \neq 1$ and

$$\max\{|\lambda_{11} + \phi_{21}\lambda_{12}| + |\lambda_{21} + \phi_{21}\lambda_{22}|, |\lambda_{22} + \phi_{12}\lambda_{21}| + |\lambda_{12} + \phi_{12}\lambda_{11}|\} < |1 - \phi_{12}\phi_{21}|, \quad (2.4)$$

then \mathbf{S} defined in (2.3) is nonsingular and the network game with the utility (2.1) has a unique Nash equilibrium in pure strategies with the equilibrium efforts given by

$$\mathbf{y}_k = \mathbf{S}^{-1}[(\mathbf{I}_n - \lambda_{ll}\mathbf{G})\boldsymbol{\pi}_k + (\phi_{lk}\mathbf{I}_n + \lambda_{lk}\mathbf{G})\boldsymbol{\pi}_l], \quad \text{for } k = 1, 2 \text{ and } l = 3 - k.$$

This theoretical model provides a microfoundation to understand an individual's behavior involving multiple activities and motivates the econometric model considered in the following section. However, it is worth noting that the best response function that the econometric model is based on can be derived from theoretical models with other underlying utility functions (see web Appendix A). Hence, the usefulness of the proposed econometric model is not limited to the specific structural model considered here.

3 Simultaneous-Equation Network Model

3.1 The econometric model

Consider a data set containing \bar{r} networks, with n_r individuals in the r -th network ($r = 1, \dots, \bar{r}$) 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 $\mathbf{G}_r = [g_{ij,r}]$ as defined in the previous section.

Our specification of the econometric model follows closely from the equilibrium best response function of the theoretical model. For the r -th network, the best response functions (2.2) can be written as

$$\mathbf{y}_{k,r} = \phi_{lk}\mathbf{y}_{l,r} + \lambda_{kk}\mathbf{G}_r\mathbf{y}_{k,r} + \lambda_{lk}\mathbf{G}_r\mathbf{y}_{l,r} + \boldsymbol{\pi}_{k,r}, \quad \text{for } k = 1, 2 \text{ and } l = 3 - k. \quad (3.1)$$

Let $\boldsymbol{\pi}_{k,r} = \mathbf{X}_r\boldsymbol{\beta}_k + \mathbf{G}_r\mathbf{X}_r\boldsymbol{\gamma}_k + \alpha_{k,r}\boldsymbol{\iota}_{n_r} + \boldsymbol{\epsilon}_{k,r}$, for $k = 1, 2$, where \mathbf{X}_r is an $n_r \times p$ matrix of observations on p exogenous individual characteristics, $\boldsymbol{\iota}_{n_r}$ is an $n_r \times 1$ vector of ones, and $\boldsymbol{\epsilon}_{k,r}$ is an $n_r \times 1$ vector of disturbances. Then, substitution of $\boldsymbol{\pi}_{k,r}$ into the best response functions (3.1) gives the simultaneous-equation network model

$$\mathbf{y}_{k,r} = \phi_{lk}\mathbf{y}_{l,r} + \lambda_{kk}\mathbf{G}_r\mathbf{y}_{k,r} + \lambda_{lk}\mathbf{G}_r\mathbf{y}_{l,r} + \mathbf{X}_r\boldsymbol{\beta}_k + \mathbf{G}_r\mathbf{X}_r\boldsymbol{\gamma}_k + \alpha_{k,r}\boldsymbol{\iota}_{n_r} + \boldsymbol{\epsilon}_{k,r}, \quad (3.2)$$

for $k = 1, 2$, $l = 3 - k$, and $r = 1, \dots, \bar{r}$.

Let $\text{diag}\{\mathbf{D}_s\}_{s=1}^{\bar{s}}$ denote a “generalized” block diagonal matrix with diagonal blocks being (possibly non-square) matrices \mathbf{D}_s 's for $s = 1, \dots, \bar{s}$. For all \bar{r} networks in the sample, the simultaneous-equation network model can be written as

$$\mathbf{y}_k = \phi_{lk}\mathbf{y}_l + \lambda_{kk}\mathbf{G}\mathbf{y}_k + \lambda_{lk}\mathbf{G}\mathbf{y}_l + \mathbf{X}\boldsymbol{\beta}_k + \mathbf{G}\mathbf{X}\boldsymbol{\gamma}_k + \mathbf{L}\boldsymbol{\alpha}_k + \boldsymbol{\epsilon}_k, \quad (3.3)$$

where $\mathbf{y}_k = (\mathbf{y}'_{k,1}, \dots, \mathbf{y}'_{k,\bar{r}})'$, $\mathbf{X} = (\mathbf{X}'_1, \dots, \mathbf{X}'_{\bar{r}})'$, $\boldsymbol{\epsilon}_k = (\boldsymbol{\epsilon}'_{k,1}, \dots, \boldsymbol{\epsilon}'_{k,\bar{r}})'$, $\mathbf{G} = \text{diag}\{\mathbf{G}_r\}_{r=1}^{\bar{r}}$, $\mathbf{L} = \text{diag}\{\boldsymbol{\iota}_{n_r}\}_{r=1}^{\bar{r}}$, and $\boldsymbol{\alpha}_k = (\alpha_{k,1}, \dots, \alpha_{k,\bar{r}})'$, for $k = 1, 2$ and $l = 3 - k$.

In model (3.3), we allow network fixed effects captured by $\boldsymbol{\alpha}_k$ to depend on \mathbf{G} and \mathbf{X} by treating $\boldsymbol{\alpha}_k$ as vectors of unknown parameters. To avoid the “incidental parameters” problem (Neyman and Scott, 1948) when the number of network \bar{r} is large, we transform (3.3) with a projector $\mathbf{J} = \text{diag}\{\mathbf{J}_r\}_{r=1}^{\bar{r}}$, where $\mathbf{J}_r = \mathbf{I}_{n_r} - \frac{1}{n_r}\boldsymbol{\iota}_{n_r}\boldsymbol{\iota}'_{n_r}$. This transformation is

analogous to the within transformation for fixed effect panel data models. As $\mathbf{JL} = \mathbf{0}$, the within-transformed model is

$$\mathbf{Jy}_k = \phi_{lk}\mathbf{Jy}_l + \lambda_{kk}\mathbf{JGy}_k + \lambda_{lk}\mathbf{JGy}_l + \mathbf{JX}\boldsymbol{\beta}_k + \mathbf{JGX}\boldsymbol{\gamma}_k + \mathbf{J}\boldsymbol{\epsilon}_k, \quad (3.4)$$

for $k = 1, 2$ and $l = 3 - k$.

Following Bramoullé et al. (2009), we assume that we observe an independently distributed sample of size \bar{r} from a population of networks with a fixed and known structure (i.e. \mathbf{G} is nonstochastic). We consider the identification of the parameters in the within-transformed model (3.4) given the moment conditions $E(\boldsymbol{\epsilon}_k|\mathbf{X}) = \mathbf{0}$ for $k = 1, 2$. It is worth noting that we do not impose any restrictions on the variance and covariance matrices of $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ given by $E(\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_l'|\mathbf{X})$, for $k, l = 1, 2$, except that they are finite and the diagonal elements of $E(\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k'|\mathbf{X})$ are bounded away from zero.

3.2 Identification challenges

As in most models in the social interaction literature (see, e.g., Blume et al., 2011; Ioannides, 2012), a host of identification issues arises in the simultaneous-equation network model (3.3). In particular, model (3.3) not only suffers from the reflection problem as single-activity social interaction models but also has the simultaneity issue that is endemic to simultaneous-equation models. Our main interest in this paper is to study the identification of the following effects in this model.

The within-activity peer effect and contextual effect The *reflection problem* (Manski, 1993) is well known and emerges from the coexistence of the *within-activity peer effect* λ_{kk} (aka. the *endogenous peer effect* in single-activity social interaction models) and the *contextual effect* $\boldsymbol{\gamma}_k$. In Manski's linear-in-means model, individuals are assumed to

be 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 within-activity peer effect and the contextual effect. Hence, these two effects cannot be identified in the linear-in-means model from the moment conditions $E(\epsilon_k|\mathbf{X}) = \mathbf{0}$.

In most social networks, individuals are not impacted evenly by all members 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. This was originally recognized in Cohen-Cole (2006) and systematically explored in Bramoullé et al. (2009). Bramoullé et al. (2009) show that these two effects can be identified if intransitivities exist in a network so that $\mathbf{I}_{n_r}, \mathbf{G}_r, \mathbf{G}_r^2$ are linearly independent. Intuitively, if individuals i and j are connected and j and k are connected, it does not necessarily imply that i and 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 instruments to identify the endogenous within-activity peer effect from the exogenous contextual effect.

The cross-activity peer effect and self-simultaneity effect A central component of our model is that we allow an individual's behavior in a certain activity to be affected by her own and her peers' choices in other activities, by introducing the *self-simultaneity effect* ϕ_{lk} , and the *cross-activity peer effect* λ_{lk} , for $l \neq k$. These two effects bring additional layers of complication to the identification. In this paper, we show that the self-simultaneity effect, the within-activity and cross-activity peer effects, and the contextual effect cannot be separately identified solely relying on intransitivities of network connections. In order to achieve identification, we need to impose exclusion restrictions on the model coefficients as in a classical simultaneous-equation model (see, e.g., Schmidt, 1976).

The correlated effect Finally, in our model, the *correlated effect* is captured by the network fixed effect parameter α_k . The network fixed effect can be motivated by a two-step link formation model, where, in the first step, individuals self-select into different networks based on network-specific characteristics and, in the second step, link formation takes place within networks based on observable individual characteristics. Thus, network fixed effects serve as a partial remedy for the bias that originates from the possible sorting of individuals into networks. In our identification strategy, the correlated effect is eliminated by the within transformation.

4 Identification of Social Interaction Effects

Identification of the simultaneous-equation network model (3.4) via the moment conditions $E(\mathbf{J}\epsilon_1|\mathbf{X}) = E(\mathbf{J}\epsilon_2|\mathbf{X}) = \mathbf{0}$ requires $E(\mathbf{J}\mathbf{Z}_1|\mathbf{X})$ and $E(\mathbf{J}\mathbf{Z}_2|\mathbf{X})$ to have full column rank, where $\mathbf{Z}_1 = [\mathbf{y}_2, \mathbf{G}\mathbf{y}_1, \mathbf{G}\mathbf{y}_2, \mathbf{X}, \mathbf{G}\mathbf{X}]$ and $\mathbf{Z}_2 = [\mathbf{y}_1, \mathbf{G}\mathbf{y}_2, \mathbf{G}\mathbf{y}_1, \mathbf{X}, \mathbf{G}\mathbf{X}]$. However, in general, this rank condition is not satisfied.

Proposition 1. *For the simultaneous-equation network model (3.4), $E(\mathbf{J}\mathbf{Z}_1|\mathbf{X})$ and $E(\mathbf{J}\mathbf{Z}_2|\mathbf{X})$ do not have full column rank.*

Therefore, to achieve identification, we need to impose exclusion restrictions. Model (3.4) has a *pseudo* reduced form

$$\mathbf{J}\mathbf{y}_k = \lambda_{kk}^* \mathbf{J}\mathbf{G}\mathbf{y}_k + \lambda_{lk}^* \mathbf{J}\mathbf{G}\mathbf{y}_l + \mathbf{J}\mathbf{X}\beta_k^* + \mathbf{J}\mathbf{G}\mathbf{X}\gamma_k^* + \mathbf{J}\epsilon_k^*, \quad (4.1)$$

for $k = 1, 2$ and $l = 3 - k$, where

$$\begin{aligned} \lambda_{kk}^* &= (1 - \phi_{12}\phi_{21})^{-1}(\lambda_{kk} + \phi_{lk}\lambda_{kl}), & \lambda_{lk}^* &= (1 - \phi_{12}\phi_{21})^{-1}(\lambda_{lk} + \phi_{lk}\lambda_{ll}), \\ \beta_k^* &= (1 - \phi_{12}\phi_{21})^{-1}(\beta_k + \phi_{lk}\beta_l), & \gamma_k^* &= (1 - \phi_{12}\phi_{21})^{-1}(\gamma_k + \phi_{lk}\gamma_l), \end{aligned} \quad (4.2)$$

and $\mathbf{J}\boldsymbol{\epsilon}_k^* = (1 - \phi_{12}\phi_{21})^{-1}(\mathbf{J}\boldsymbol{\epsilon}_k + \phi_{lk}\mathbf{J}\boldsymbol{\epsilon}_l)$. Our identification strategy takes two steps as in Yang and Lee (2017). In the first step, we show that the pseudo reduced form parameters can be identified by exploiting intransitivities of network connections. In the second step, we show that the structural parameters in model (3.4) can be identified from the pseudo reduced form parameters by imposing exclusion restrictions as in a classical linear simultaneous-equation model.

4.1 Identification of pseudo reduced form parameters

The pseudo reduced form (4.1) has the same specification as a simultaneous-equation network model without self-simultaneity effects, i.e. $\phi_{12} = \phi_{21} = 0$. By a similar argument as in Bramoullé et al. (2009), this model is identified if the network topology satisfies Assumption 1 (see the proof of Proposition 2). Let $\beta_{k,h}^*$ and $\gamma_{k,h}^*$ denote the h -th element of $\boldsymbol{\beta}_k^*$ and $\boldsymbol{\gamma}_k^*$ respectively, for $k = 1, 2$.

Assumption 1. (i) *In the data generating process, for some $h \in \{1, \dots, p\}$,*

$$\begin{aligned} & (\lambda_{11}^*\beta_{1,h}^* + \lambda_{21}^*\beta_{2,h}^* + \gamma_{1,h}^*)[(\lambda_{12}^*\lambda_{21}^* - \lambda_{11}^*\lambda_{22}^*)\beta_{2,h}^* + \lambda_{12}^*\gamma_{1,h}^* - \lambda_{11}^*\gamma_{2,h}^*] \\ & \neq (\lambda_{22}^*\beta_{2,h}^* + \lambda_{12}^*\beta_{1,h}^* + \gamma_{2,h}^*)[(\lambda_{12}^*\lambda_{21}^* - \lambda_{11}^*\lambda_{22}^*)\beta_{1,h}^* + \lambda_{21}^*\gamma_{2,h}^* - \lambda_{22}^*\gamma_{1,h}^*]. \end{aligned}$$

(ii) *The matrices $\mathbf{I}_n, \mathbf{G}, \mathbf{G}^2, \mathbf{G}^3, \mathbf{G}^4$ are linearly independent.*

Remark 1. The moment condition $E(\mathbf{J}\boldsymbol{\epsilon}_k|\mathbf{X}) = \mathbf{0}$ implies $E(\mathbf{J}\mathbf{y}_k|\mathbf{X}) = E(\mathbf{J}\mathbf{Z}_k^*|\mathbf{X})\boldsymbol{\theta}_k^*$, where $\mathbf{Z}_k^* = [\mathbf{G}\mathbf{y}_k, \mathbf{G}\mathbf{y}_l, \mathbf{X}, \mathbf{G}\mathbf{X}]$ and $\boldsymbol{\theta}_k^* = (\lambda_{kk}^*, \lambda_{lk}^*, \boldsymbol{\beta}_k^{*'}, \boldsymbol{\gamma}_k^{*'})'$, for $k = 1, 2$ and $l = 3 - k$. To better understand Assumption 1 (i), consider two special cases where this assumption is violated. In the first case, suppose $\boldsymbol{\beta}_k^* = \boldsymbol{\gamma}_k^* = \mathbf{0}$, for $k = 1, 2$, in the data generating process. This case corresponds to the situation where none of the observed exogenous characteristics has an effect on \mathbf{y}_1 and \mathbf{y}_2 . In this case, $E(\mathbf{J}\mathbf{Z}_k^*|\mathbf{X})$ does not have full column rank as

$E(\mathbf{JGy}_1|\mathbf{X}) = E(\mathbf{JGy}_2|\mathbf{X}) = \mathbf{0}$. In the second case, suppose the restrictions, $\lambda_{11}^* = \lambda_{22}^*$ and $\lambda_{12}^* = \lambda_{21}^* = 0$, hold in the data generating process but the researcher estimates model (4.1) without imposing these restrictions. This case corresponds to the situation where the true model is a seemingly unrelated regression (SUR) network model with identical within-activity peer effects. In this case, $E(\mathbf{JGy}_k|\mathbf{X}) = \mathbf{JG}(\mathbf{I} - \lambda_{kk}^* \mathbf{G})^{-1}[\mathbf{X}, \mathbf{GX}](\boldsymbol{\beta}_k^{*'}, \boldsymbol{\gamma}_k^{*'})'$, for $k = 1, 2$, and hence $E(\mathbf{JZ}_k^*|\mathbf{X})$ does not have full column rank due to the perfect collinearity of $E(\mathbf{JGy}_1|\mathbf{X})$ and $E(\mathbf{JGy}_2|\mathbf{X})$.

Remark 2. As pointed out by Bramoullé et al. (2009), the powers of the adjacency matrix \mathbf{G} is closely related to the diameter of the network. In graph theory, the (i, j) -th element of \mathbf{G}^s is non-zero if there exists a path from node i to node j of length s , and the diameter of a network is the shortest distance between the two most distant nodes in the network. Hence, to check the linear independence of $\mathbf{I}_n, \mathbf{G}, \mathbf{G}^2, \mathbf{G}^3, \mathbf{G}^4$, one could simply check if there exists a pair of nodes i and j ($i \neq j$) in the network such that the shortest path from i to j is of length 4, i.e., if the diameter of the network is no less than 4.

4.2 Identification of structural parameters

If the pseudo reduced form parameters in (4.1) are identified, then the structural parameters in model (3.4) can be identified via (4.2) as in a classical linear simultaneous-equation model (see, e.g., Schmidt, 1976). To be more specific, Let $\boldsymbol{\theta}_1 = (1, -\phi_{21}, -\lambda_{11}, -\lambda_{21}, -\boldsymbol{\beta}'_1, -\boldsymbol{\gamma}'_1)'$, $\boldsymbol{\theta}_2 = (-\phi_{12}, 1, -\lambda_{12}, -\lambda_{22}, -\boldsymbol{\beta}'_2, -\boldsymbol{\gamma}'_2)'$, and $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]$. Suppose that, for $k = 1, 2$, there are q_k restrictions on $\boldsymbol{\theta}_k$ of the form $\mathbf{R}_k \boldsymbol{\theta}_k = \mathbf{0}$ where \mathbf{R}_k is a $q_k \times (4 + 2p)$ matrix of known constants. Then, the sufficient and necessary *rank* condition for $\boldsymbol{\theta}_k$ to be identified by the restrictions $\mathbf{R}_k \boldsymbol{\theta}_k = \mathbf{0}$ is that $\text{rank}(\mathbf{R}_k \boldsymbol{\Theta}) = 1$, and the necessary *order* condition is $q_k \geq 1$.

Assumption 2. *Suppose there are q_k restrictions on $\boldsymbol{\theta}_k$ of the form $\mathbf{R}_k \boldsymbol{\theta}_k = \mathbf{0}$, such that $\text{rank}(\mathbf{R}_k \boldsymbol{\Theta}) = 1$, for $k = 1, 2$.*

Proposition 2. *Under Assumptions 1 and 2, the simultaneous-equation network model (3.4) is identified.*

To better understand Proposition 2, especially the role played by Assumption 2 in identification, consider the following examples.

Example 1. Consider the model

$$\mathbf{y}_k = \phi_{lk}\mathbf{y}_l + \lambda_{kk}\mathbf{G}\mathbf{y}_k + \mathbf{X}\boldsymbol{\beta}_k + \mathbf{G}\mathbf{X}\boldsymbol{\gamma}_k + \mathbf{L}\boldsymbol{\alpha}_k + \boldsymbol{\epsilon}_k$$

for $k = 1, 2$ and $l = 3 - k$, with its within-transformed counterpart

$$\mathbf{J}\mathbf{y}_k = \phi_{lk}\mathbf{J}\mathbf{y}_l + \lambda_{kk}\mathbf{J}\mathbf{G}\mathbf{y}_k + \mathbf{J}\mathbf{X}\boldsymbol{\beta}_k + \mathbf{J}\mathbf{G}\mathbf{X}\boldsymbol{\gamma}_k + \mathbf{J}\boldsymbol{\epsilon}_k. \quad (4.3)$$

This model includes the self-simultaneity effect and within-activity peer effect but does not include the cross-activity peer effect. It has a pseudo reduced form defined in (4.1), where

$$\begin{aligned} \lambda_{kk}^* &= (1 - \phi_{12}\phi_{21})^{-1}\lambda_{kk}, & \lambda_{lk}^* &= (1 - \phi_{12}\phi_{21})^{-1}\phi_{lk}\lambda_{ll}, \\ \boldsymbol{\beta}_k^* &= (1 - \phi_{12}\phi_{21})^{-1}(\boldsymbol{\beta}_k + \phi_{lk}\boldsymbol{\beta}_l), & \boldsymbol{\gamma}_k^* &= (1 - \phi_{12}\phi_{21})^{-1}(\boldsymbol{\gamma}_k + \phi_{lk}\boldsymbol{\gamma}_l). \end{aligned} \quad (4.4)$$

Suppose Assumption 1 is satisfied and the pseudo reduced form parameters can be identified. Then, the parameters in model (4.3) can be identified via (4.4) if Assumption 2 holds. The exclusion restriction $\lambda_{21} = 0$ can be written as $\mathbf{R}_1\boldsymbol{\theta}_1 = 0$ where $\mathbf{R}_1 = [0, 0, 0, -1, \mathbf{0}_{1 \times p}, \mathbf{0}_{1 \times p}]$. Then $\mathbf{R}_1\boldsymbol{\Theta} = [0, \lambda_{22}]$, which has rank 1 if $\lambda_{22} \neq 0$. Similarly, the exclusion restriction $\lambda_{12} = 0$ can be written as $\mathbf{R}_2\boldsymbol{\theta}_2 = 0$ where $\mathbf{R}_2 = [0, 0, -1, 0, \mathbf{0}_{1 \times p}, \mathbf{0}_{1 \times p}]$. Then $\mathbf{R}_2\boldsymbol{\Theta} = [\lambda_{11}, 0]$, which has rank 1 if $\lambda_{11} \neq 0$. Indeed, if $\lambda_{11} = \lambda_{22} = 0$, then (4.3) becomes a classical linear simultaneous-equation model, which cannot be identified without imposing additional exclusion restrictions. \square

Example 2. Suppose $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$, where \mathbf{X}_1 and \mathbf{X}_2 are, respectively, $n \times p_1$ and $n \times p_2$ matrices of exogenous variables. Correspondingly, partition the parameter vectors as follows $\beta_1 = (\beta'_{11}, \beta'_{21})'$, $\beta_2 = (\beta'_{12}, \beta'_{22})'$, $\gamma_1 = (\gamma'_{11}, \gamma'_{21})'$ and $\gamma_2 = (\gamma'_{12}, \gamma'_{22})'$. Consider the model

$$\mathbf{y}_k = \phi_{lk}\mathbf{y}_l + \lambda_{kk}\mathbf{G}\mathbf{y}_k + \lambda_{lk}\mathbf{G}\mathbf{y}_l + \mathbf{X}_k\beta_{kk} + \mathbf{G}\mathbf{X}_k\gamma_{kk} + \mathbf{L}\alpha_k + \epsilon_k$$

for $k = 1, 2$ and $l = 3 - k$, with its within-transformed counterpart

$$\mathbf{J}\mathbf{y}_k = \phi_{lk}\mathbf{J}\mathbf{y}_l + \lambda_{kk}\mathbf{J}\mathbf{G}\mathbf{y}_k + \lambda_{lk}\mathbf{J}\mathbf{G}\mathbf{y}_l + \mathbf{J}\mathbf{X}_k\beta_{kk} + \mathbf{J}\mathbf{G}\mathbf{X}_k\gamma_{kk} + \mathbf{J}\epsilon_k. \quad (4.5)$$

This model has a pseudo reduced form defined in (4.1), where

$$\begin{aligned} \lambda_{kk}^* &= (1 - \phi_{12}\phi_{21})^{-1}(\lambda_{kk} + \phi_{lk}\lambda_{kl}), \\ \lambda_{lk}^* &= (1 - \phi_{12}\phi_{21})^{-1}(\lambda_{lk} + \phi_{lk}\lambda_{ll}), \\ [\beta_1^*, \beta_2^*] &= (1 - \phi_{12}\phi_{21})^{-1} \begin{bmatrix} \beta_{11} & \phi_{12}\beta_{11} \\ \phi_{21}\beta_{22} & \beta_{22} \end{bmatrix}, \\ [\gamma_1^*, \gamma_2^*] &= (1 - \phi_{12}\phi_{21})^{-1} \begin{bmatrix} \gamma_{11} & \phi_{12}\gamma_{11} \\ \phi_{21}\gamma_{22} & \gamma_{22} \end{bmatrix}. \end{aligned} \quad (4.6)$$

Suppose Assumption 1 is satisfied and the pseudo reduced form parameters can be identified. Then, the parameters in model (4.5) can be identified via (4.6) if Assumption 2 holds.

The exclusion restrictions $\beta_{21} = \gamma_{21} = \mathbf{0}$ can be written as $\mathbf{R}_1\boldsymbol{\theta}_1 = \mathbf{0}$ where

$$\mathbf{R}_1 = \begin{bmatrix} \mathbf{0}_{p_2 \times 4} & \mathbf{0}_{p_2 \times p_1} & -\mathbf{I}_{p_2} & \mathbf{0}_{p_2 \times p_1} & \mathbf{0}_{p_2 \times p_2} \\ \mathbf{0}_{p_2 \times 4} & \mathbf{0}_{p_2 \times p_1} & \mathbf{0}_{p_2 \times p_2} & \mathbf{0}_{p_2 \times p_1} & -\mathbf{I}_{p_2} \end{bmatrix}.$$

Then

$$\mathbf{R}_1 \Theta = \begin{bmatrix} \mathbf{0}_{p_2 \times 1} & \beta_{22} \\ \mathbf{0}_{p_2 \times 1} & \gamma_{22} \end{bmatrix},$$

which has rank 1 if $(\beta'_{22}, \gamma'_{22})' \neq \mathbf{0}$. Similarly, the exclusion restrictions $\beta_{12} = \gamma_{12} = \mathbf{0}$ can be written as $\mathbf{R}_2 \theta_2 = \mathbf{0}$ where

$$\mathbf{R}_2 = \begin{bmatrix} \mathbf{0}_{p_1 \times 4} & -\mathbf{I}_{p_1} & \mathbf{0}_{p_1 \times p_2} & \mathbf{0}_{p_1 \times p_1} & \mathbf{0}_{p_1 \times p_2} \\ \mathbf{0}_{p_1 \times 4} & \mathbf{0}_{p_1 \times p_1} & \mathbf{0}_{p_1 \times p_2} & -\mathbf{I}_{p_1} & \mathbf{0}_{p_1 \times p_2} \end{bmatrix}.$$

Then

$$\mathbf{R}_2 \Theta = \begin{bmatrix} \beta_{11} & \mathbf{0}_{p_1 \times 1} \\ \gamma_{11} & \mathbf{0}_{p_1 \times 1} \end{bmatrix},$$

which has rank 1 if $(\beta'_{11}, \gamma'_{11})' \neq \mathbf{0}$. □

5 Empirical Application

5.1 Data

To illustrate the empirical salience of the proposed model, we study the (peer) effects of screen-related activities (e.g., watching TV, playing video games, etc.) on the academic performance of a student. Indeed, there is a growing concern that screen-related activities are taking up the time of adolescents and that these activities have strong negative effects on education. In the United States, eight- to eighteen-year-olds spend more time with media than in any other activity besides (maybe) sleeping – an average of more than $7\frac{1}{2}$ hours a day, seven days a week (Cordes and Miller, 2000). The TV shows they watch, video games they play, and websites they visit have an enormous influence on their lives.

Moreover, there is strong evidence that screen-related activities have a negative impact

on education. For example, in a research synthesis of 23 studies of the relation between leisure television time and achievement, Williams et al. (1982) found an overall negative relation between achievement and TV time. The relation between achievement and TV watching seems to persist across research designs and background characteristics are controlled for. In particular, it has been shown that TV watching negatively impacts reading comprehension skills and reduces recreational reading (Koolstra et al., 1997). Moreover, a number of studies have documented a significant negative relationship between the amount of time spent with screen-based media (television, movies and video games) and school performance (see, e.g., Cordes and Miller, 2000; Chan and Rabinowitz, 2006; Gentile, 2009). For example, a survey on a large, nationally representative sample of American children and adolescents found that nearly half (47%) of heavy media users get poor grades compared to 23% of light media users (Rideout et al., 2010). A longitudinal study of elementary school children showed that total screen time significantly predicts poorer grades later in the school year, even while controlling for other relevant characteristics (Anderson et al., 2007).

These studies, however, did not take into account peer effects in these activities. To understand the impact of peers on education and screen-related activities, we use 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. The survey was conducted in 1994-1995 and was designed to capture information on friends, family, school and neighborhood influences on students behaviors, including academic performance, social decisions, extracurriculars, dangerous behaviors and more. Every student attending schools on the sampling day was provided with a questionnaire that covered topics on demographics, behavioral characteristics, education, family background and critically for our purposes, friendships. The in-school survey was followed by four waves of in-home

interviews with more detailed information. In this empirical study, we use the first wave of the in-home interview data.

We consider the estimation of model (3.3) where \mathbf{y}_1 and \mathbf{y}_2 measure, respectively, academic performance and screen-related activities. To be more specific, \mathbf{y}_1 is the average grade (converted to a four point scale, with A = 4, B = 3, etc.) in English (or language arts), mathematics, history (or social studies) and science at the most recent grading period. \mathbf{y}_2 is the logarithm of the total number of hours spent on watching TV/videos and playing video/computer games in a week. We use the logarithm to alleviate the problem of measurement errors when a student reports spending a large amount of time on screen-related activities. After taking the logarithm, \mathbf{y}_2 has similar mean and standard deviation as \mathbf{y}_1 . A list of the variables used in the empirical study, together with their summary statistics, is given in Table 1.

[Insert Table 1 here]

The adjacency matrix $\mathbf{G} = [g_{ij}]$ is constructed based on the friend-nomination information provided by the Add Health data. In the in-school survey questionnaire, students were asked to identify their 10 best friends (up to 5 female friends and 5 male friends) from a school roster. About 6.5% of the students in the sample nominated 5 female friends and about 3.9% of the students in the sample nominated 5 male friends. Thus the bound on the number of friend-nominations is not binding. We define $g_{ij} = 1/n_i$ if student i nominates student j as a friend and $g_{ij} = 0$ otherwise, where n_i is the number of nominated friends of student i . A network is defined as the smallest set of students such that all students in the same network are directly or indirectly connected through friend nominations while no students from different networks are connected. After removing isolated students (i.e. students who nominated no friends and were not nominated by any students) and students

with missing observations on \mathbf{y}_1 and \mathbf{y}_2 , the sample consists of 7,669 students distributed over 124 schools. A school usually consists of several networks. In the sample used by this empirical study, there are 1043 networks, with sizes ranging from 2 to 484.² Among all the networks in the sample, there are 315 networks with diameters no less than 4. Hence, the identification condition given in Assumption 1 (ii) in terms of the network topology is clearly satisfied for the sample considered.

5.2 Parameter estimates

We consider the estimation of model (3.4) under different exclusion restrictions. First, we impose the exclusion restrictions that $\lambda_{12} = \lambda_{21} = \phi_{12} = \phi_{21} = 0$. Under these exclusion restrictions, model (3.4) reduces to a single-activity social interaction model. The 2SLS estimation results with the IV matrix $\mathbf{Q}_1 = [\mathbf{JX}, \mathbf{JGX}, \mathbf{JG}^2\mathbf{X}]$ are reported in the left panel (under Model 1) of Table 2a. The estimates of the within-activity peer effect show that the academic performance of the peers has a statistically significantly positive effect on a student’s academic performance. This result is in line with studies in the literature showing positive peer effects in education (see, e.g., Calvó-Armengol et al., 2009; De Giorgi et al., 2010; Bifulco et al., 2011). Also, the time spent by the peers on screen-related activities has a positive effect on a student’s own time spent on these activities. However, the estimated peer effect in screen-related activities is not statistically significant.

It is worth noting that the validity of the IV matrix \mathbf{Q}_1 relies on the exogeneity of the network adjacency matrix \mathbf{G} . If the over-identifying restrictions (OIR) test (Lin and Lee, 2010) cannot reject the null hypothesis that the IV matrix \mathbf{Q}_1 is valid, then it provides evidence that \mathbf{G} is uncorrelated with the error term after controlling for the exogenous regressors \mathbf{X} and network fixed effects. As reported at the bottom of Table 2a, the p-value

²The estimation results reported in the following subsection are qualitatively unchanged when we drop networks of extremely small or large sizes from the sample.

of the OIR test is larger than conventional significance levels, which provides evidence for the exogeneity of \mathbf{G} .

[Insert Table 2a and Table 2b here]

Next, we impose the exclusion restrictions that $\lambda_{11} = \lambda_{22} = \lambda_{12} = \lambda_{21} = 0$. Under these exclusion restrictions, model (3.4) becomes a classical simultaneous-equation model without endogenous peer effects. It is well known that the identification of this model requires instruments (or exclusion restrictions). Let \mathbf{x}_1 be a vector of dummy variables set equal to 1 if at least one of the non-resident biological parents of the student is a college graduate, and 0 if the non-resident biological parents do not have a college degree or the student lives with both biological parents.³ The intelligence of a student is likely to be correlated with her biological parent’s education. However, the non-resident parent would have little influence on the amount of time the student spends on screen-related activities. Hence, we use \mathbf{x}_1 as an instrument for the academic performance \mathbf{y}_1 . On the other hand, let \mathbf{x}_2 be a vector of dummy variables coded as 1 if the resident parents let the student decide how much TV to watch, and 0 otherwise. We use \mathbf{x}_2 as an instrument for \mathbf{y}_2 , with the underlying exclusion restriction that whether the student is allowed to make her own decision on how much TV to watch only affects her academic performance indirectly through how much time she spends on watching TV. As the model includes contextual effects, we use $\mathbf{G}\mathbf{x}_1$ and $\mathbf{G}\mathbf{x}_2$ as additional instruments for \mathbf{y}_1 and \mathbf{y}_2 respectively. Thus, the model is over-identified. The 2SLS estimation results with the IV matrix $\mathbf{Q}_2 = [\mathbf{J}\mathbf{X}, \mathbf{J}\mathbf{G}\mathbf{X}]$, where \mathbf{X} is a matrix of all exogenous variables (listed in Table 1) including \mathbf{x}_1 and \mathbf{x}_2 , are reported in the right panel (under Model 2) of Table 2a. The p-value of the OIR test is larger than conventional

³The effects of living with both biological parents and the education level of resident parents are controlled for by other regressors in \mathbf{X} .

significance levels, which provides some evidence on the validity of the instruments. The Cragg-Donald F statistics (Stock and Yogo, 2005) suggest the instruments are informative. We find the time a student spends on screen-related activities is negatively affected by her own GPA, while a student’s academic performance is positively affected by the time spent on screen-related activities. However, neither effect is statistically significant.

Finally, we estimate model (3.4) with self-simultaneity effects and cross-activity peer effects. As discussed in Section 4, identification of this model requires exclusion restrictions. We consider three sets of exclusion restrictions. Model 3 imposes the exclusion restrictions that $\phi_{12} = \phi_{21} = 0$, i.e. no self-simultaneity effects. Model 4 imposes the exclusion restrictions that $\lambda_{12} = \lambda_{21} = 0$, i.e. no cross-activity peer effects. Model 5 imposes the same set of exclusion restrictions on the exogenous regressors as in Model 2. It is worth noting that Model 3 has the same specification of the pseudo reduced form (4.1), Model 4 conforms to the model in Example 1, and Model 5 conforms to the model in Example 2. Table 2b reports the 2SLS estimation results of these three models with the IV matrix $\mathbf{Q}_3 = [\mathbf{JX}, \mathbf{JGX}, \mathbf{JG}^2\mathbf{X}]$, where \mathbf{X} includes a subset of exogenous variables in \mathbf{X} . To be more specific, \mathbf{X} includes “Age”, “Female”, “Living condition”, and “Live with both bio parents”. As the Cragg-Donald F statistics reported in Table 2b suggest the instruments are weak, we only use a subset of the exogenous characteristics in \mathbf{X} to construct instruments to alleviate the potential weak instrument bias. The estimates of these three models are qualitatively consistent with each other. The estimates of λ_{kk} , λ_{lk} and ϕ_{lk} ($k = 1, 2$ and $l = 3 - k$) satisfy the condition given in (2.4), suggesting the reduced form equations of the system are well defined. From the estimates of Model 5, we find that the academic performance of a student is not only positively affected the academic performance of the peers, but also negatively affected by the time the peers spend on screen-related activities. Both types of peer effects are statistically significant. We also find that the academic

performance of a student is negatively affected by the time she spend on screen-related activities. However, only the negative effect of GPA on one’s own screen-related activities is statistically significant. These results confirm the studies cited at the beginning of this section and, more importantly, show the importance of peer effects in these activities.

5.3 Marginal effects

It follows from the reduced form of model (3.3) that

$$\frac{\partial \mathbf{y}_k}{\partial \mathbf{x}'_h} = \mathbf{S}^{-1}[(\phi_{lk}\beta_{l,h} + \beta_{k,h})\mathbf{I}_n + (\lambda_{lk}\beta_{l,h} - \lambda_{ll}\beta_{k,h} + \phi_{lk}\gamma_{l,h} + \gamma_{k,h})\mathbf{G} + (\lambda_{lk}\gamma_{l,h} - \lambda_{ll}\gamma_{k,h})\mathbf{G}^2]$$

for $k = 1, 2$ and $l = 3 - k$, where $\mathbf{x}_h = (x_{1h}, \dots, x_{nh})'$ is the h -th column of \mathbf{X} . For $k = 1, 2$, $\partial \mathbf{y}_k / \partial \mathbf{x}'_h$ is an $n \times n$ matrix of marginal effects, with its (i, j) -th element given by $\partial y_{ik} / \partial x_{jh}$. The off-diagonal element of $\partial \mathbf{y}_k / \partial \mathbf{x}'_h$, in general, is not zero, suggesting that a change in the h -th explanatory variable for an individual can potentially affect the dependent variable of all the other individuals in the network. Following LeSage and Pace (2009), we define the *average direct impact* of \mathbf{x}_h on \mathbf{y}_k as $n^{-1} \sum_{i=1}^n \partial y_{ik} / \partial x_{ih}$ and the *average indirect impact* of \mathbf{x}_h on \mathbf{y}_k as $n^{-1} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \partial y_{ik} / \partial x_{jh}$, for $h = 1, \dots, p$ and $k = 1, 2$. Table 3 reports the average direct and indirect impacts of the exogenous variables with standard errors calculated by the Delta method.

[Insert Table 3 here]

Due to the presence of simultaneity/peer effects, the average direct impact of a covariate is, in general, different from its coefficient estimate reported in Table 2b. Some of the average direct impacts even have opposite signs from the corresponding coefficient estimates. According to the reported marginal effects, a younger white female student, who

is in a higher grade, in excellent health, lives with both biological parents in a well-kept home, and has well-educated resident or non-resident biological parents, is more likely to have better academic performance. A younger non-white male student, who is in poor health and is allowed to make own decision on TV watching time, tends to spend more time on screen-related activities.

6 Conclusion

In this paper, we investigate the impact of peers on individual outcomes where individuals embedded in a network are involved in multiple activities. We develop a general simultaneous-equation 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 self-simultaneity effect and the cross-activity peer effect. We provide identification conditions for network models with the above effects. We then study the impact of peer effects on education and screen-related activities and show that a student's academic performance is not only affected by the academic performance of the peers but also affected by screen-related activities of the peers.

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 in constructing 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.

Some possible extensions of the current work are in order. First, different individu-

als may participate in different activities. Therefore, it would be interesting to study the sample selection issue (Heckman, 1976) in the context of social networks and multivariate choices. Second, people may form different social networks for different activities they participate. Hence, another thread of future research could be to consider activity-specific networks and to study the formation and evolution of such networks and associated identification problems. Third, sampling issues prevail in network data. It is very rare one can observe the whole network of the full population. For example, the Add Health data used in the empirical application does not provide information on students' friends outside school. For the single-activity network model, there is a growing literature on the sampling issue in network data (see, e.g., Sojourner, 2013; Liu, 2013; Liu et al., 2016; Chandrasekhar and Lewis, 2016). It would be interesting to extend these works to the simultaneous-equation network model.

References

- Anderson, C. A., Gentile, D. A. and Buckley, K. E. (2007). *Violent Video Game Effects on Children and Adolescents: Theory, Research, and Public Policy*, Oxford University Press.
- Ballester, C., Calvó-Armengol, A. and Zenou, Y. (2006). Who's who in networks. Wanted: the key player, *Econometrica* **74**: 1403–1417.
- Belhaj, M. and Deroian, F. (2014). Competing activities in social networks, *B.E. Journal of Economic Analysis & Policy* **14**: 1431–1466.
- Bifulco, R., Fletcher, J. M. and Ross, S. L. (2011). The effect of classmate characteristics on post-secondary outcomes: Evidence from the uppercaseAdd uppercaseHealth, *American Economic Journal: Economic Policy* **3**: 25–53.

- Blume, L. E., Brock, W. A., Durlauf, S. N. and Ioannides, Y. M. (2011). Identification of social interactions, in J. Benhabib, A. Bisin and M. O. Jackson (eds), *Handbook of Social Economics*, Vol. 1B, North-Holland, pp. 855–966.
- Bramoullé, Y., Djebbari, H. and Fortin, B. (2009). Identification of peer effects through social networks, *Journal of Econometrics* **150**: 41–55.
- Bramoullé, Y. and Kranton, R. (2007). Public goods in networks, *Journal of Economic Theory* **135**: 478–494.
- Bramoullé, Y., Kranton, R. and D’Amours, M. (2014). Strategic interaction and networks, *American Economic Review* **104**: 898–930.
- Calvó-Armengol, A., Patacchini, E. and Zenou, Y. (2009). Peer effects and social networks in education, *The Review of Economic Studies* **76**: 1239–1267.
- Chan, P. A. and Rabinowitz, T. (2006). A cross-sectional analysis of video games and attention deficit hyperactivity disorder symptoms in adolescents, *Annals of General Psychiatry* **5**: 1–10.
- Chandrasekhar, A. and Lewis, R. (2016). Econometrics of sampled networks. Working paper, Stanford University.
- Chen, Y.-J., Zenou, Y. and Zhou, J. (2015). Multiple activities for socially connected criminals. CEPR Discussion Paper No. 10709.
- Cohen-Cole, E. (2006). Multiple groups identification in the linear-in-means model, *Economic Letters* **92**: 157–162.
- Cordes, C. and Miller, E. (2000). *Fool’s Gold: A Critical Look at Computers in Childhood*, Alliance for Childhood.

- De Giorgi, G., Pellizzari, M. and Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups, *American Economic Journal: Applied Economics* **2**: 241–275.
- Gentile, D. (2009). Pathological video-game use among youth ages 8 to 18: A national study, *Psychological Science* **20**: 594–602.
- Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator of such models, *Annals of Economic and Social Measurement* **5**: 475–492.
- Ioannides, Y. M. (2012). *From Neighborhoods to Nations: The Economics of Social Interactions*, Princeton University Press.
- Ioannides, Y. M. and Loury, D. L. (2004). Job information networks, neighborhood effects, and inequality, *Journal of Economic Literature* **42**: 1056–1093.
- Jackson, M. O. and Zenou, Y. (2015). Games on networks, in P. Young and S. Zamir (eds), *Handbook of Game Theory*, Vol. 4, Elsevier, pp. 91–157.
- Kelejian, H. H. and Prucha, I. R. (2004). Estimation of simultaneous systems of spatially interrelated cross sectional equations, *Journal of Econometrics* **118**: 27–50.
- Koolstra, C., van der Voort, T. and van der Kamp, L. (1997). Television’s impact on reading comprehension and decoding skills: A 3-year panel study, *Reading Research Quarterly* **32**: 128–152.
- Lee, L. F., Liu, X. and Lin, X. (2010). Specification and estimation of social interaction models with network structures, *The Econometrics Journal* **13**: 145–176.
- LeSage, J. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*, CRC Press.

- Lin, X. and Lee, L. F. (2010). GMM estimation of spatial autoregressive models with unknown heteroskedasticity, *Journal of Econometrics* **157**: 34–52.
- Liu, X. (2013). Estimation of a local-aggregate network model with sampled networks, *Economics Letters* **118**: 243–246.
- Liu, X., Patacchini, E. and Rainone, E. (2016). Peer effects in bed time decisions among adolescents: a social network model with sampled data. *The Econometrics Journal*, forthcoming.
- Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem, *The Review of Economic Studies* **60**: 531–542.
- Neyman, J. and Scott, E. L. (1948). Consistent estimates based on partially consistent observations, *Econometrica* **16**: 1–32.
- Patacchini, E. and Zenou, Y. (2012). Juvenile delinquency and conformism, *Journal of Law, Economics, and Organization* **28**: 1–31.
- Rideout, V., Foehr, U. and Roberts, D. (2010). Generation M2: Media in the lives of 8 to 18 year olds. The Henry J. Kaiser Family Foundation, <http://files.eric.ed.gov/fulltext/ED527859.pdf>.
- Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far?, in E. A. Hanushek, S. Machin and L. Woessmann (eds), *Handbook of Economics of Education*, Vol. 3, Elsevier, pp. 249–277.
- Schmidt, P. (1976). *Econometrics*, Marcel Dekker.
- Sojourner, A. (2013). Identification of peer effects with missing peer data: Evidence from project star, *The Economic Journal* **123**: 574–605.

- Stock, J. and Yogo, M. (2005). Testing for weak instruments in linear iv regression, *in* D. W. K. Andrews and J. Stock (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press, pp. 80–108.
- Williams, P. A., Haertel, E. H., Haertel, G. D. and Walberg, H. J. (1982). The impact of leisure-time television on school learning: A research synthesis, *American Educational Research Journal* **19**: 19–50.
- Yang, K. and Lee, L. F. (2017). Identification and QML estimation of multivariate and simultaneous equations spatial autoregressive models, *Journal of Econometrics* **196**: 196–214.

Table 1: Data Summary

	Definition	Mean	SD
Dependent variables			
GPA	The average grade in English, math, history, and science at the most recent grading period	2.87	0.73
TV	The logarithm of weekly time spent on watching TV/videos and playing video/computer games	2.84	0.82
Control variables			
Age	Age	15.29	1.66
Female	1 if female	0.53	0.50
(White)	1 if White American	0.53	0.50
African American	1 if African American	0.22	0.41
Other races	1 if race is not White or African American	0.25	0.43
(Freshman)	1 if in Grade 7 or 8	0.34	0.47
Junior	1 if in Grade 9 or 10	0.38	0.48
Senior	1 if in Grade 11 or 12	0.28	0.45
Health	1 if health is excellent	0.30	0.46
Living condition	1 if the building in which the respondent lives is well kept	0.59	0.49
Live with both bio parents	1 if live with both biological parents	0.55	0.50
(Res parent: less than HS)	1 if the resident parent's education is less than high school	0.13	0.34
Res parent: HS grad	1 if the resident parent's education is high school or higher but no college degree	0.54	0.50
Res parent: college grad	1 if the resident parent's education is college or higher	0.29	0.45
Res parent: educ missing	1 if the resident parent's education information is missing	0.04	0.20
Res parent: professional	1 if the resident parent's job is a doctor, lawyer, scientist, teacher, librarian, nurse, manager, executive, director, technical/computer specialist, or radiologist	0.30	0.46
Res parent: office worker	1 if the resident parent's job is office worker, bookkeepers, clerk, secretary, sales worker, insurance agent, or store clerk	0.22	0.41
Res parent: other job	1 if the resident parent's job is not listed above	0.34	0.47
(Res parent: no job)	1 if the resident parent does not have a job	0.13	0.34
Res parent: job missing	1 if the resident parent's job information is missing	0.01	0.10
Bio parent: college grad	1 if the non-resident bio parent is a college graduate	0.08	0.26
Own TV time decision	1 if the resident parents let the respondent decide how much TV to watch	0.81	0.39

The variable in the parentheses is the reference category.

If both parents are in the household, the education and job of the mother is considered.

Table 2a: Parameter Estimates

	Model 1		Model 2	
	GPA	TV	GPA	TV
Peer GPA	0.3999*** (0.1304)			
Peer TV		0.2924 (0.2902)		
Own GPA				-0.4652 (0.2908)
Own TV			0.1941 (0.1985)	
Age	-0.0570*** (0.0112)	-0.0582*** (0.0184)	-0.0496*** (0.0173)	-0.1008*** (0.0232)
Female	0.2086*** (0.0162)	-0.2650*** (0.0198)	0.2566*** (0.0552)	-0.1734*** (0.0626)
African American	-0.0985*** (0.0338)	0.2916*** (0.0422)	-0.1652*** (0.0660)	0.2303*** (0.0533)
Other races	-0.0191 (0.0284)	0.1071*** (0.0362)	-0.0461 (0.0371)	0.1067*** (0.0355)
Junior	0.0461 (0.0454)	-0.0965* (0.0549)	0.0701 (0.0485)	-0.0607 (0.0580)
Senior	0.2403*** (0.0555)	-0.0765 (0.0723)	0.2619*** (0.0565)	0.0665 (0.1029)
Health	0.1138*** (0.0175)	-0.0659*** (0.0207)	0.1421*** (0.0210)	-0.0026 (0.0437)
Living condition	0.0727*** (0.0176)	0.0065 (0.0209)	0.0814*** (0.0173)	0.0496 (0.0326)
Live with both bio parents	0.1519*** (0.0192)	-0.0101 (0.0230)	0.1722*** (0.0184)	0.0731 (0.0470)
Res parent: HS grad	0.1098*** (0.0267)	0.0104 (0.0336)	0.1239*** (0.0267)	0.0810 (0.0500)
Res parent: college grad	0.2418*** (0.0325)	-0.0375 (0.0373)	0.2806*** (0.0319)	0.0923 (0.0920)
Res parent: educ missing	-0.0425 (0.0477)	-0.0480 (0.0657)	-0.0212 (0.0496)	-0.0578 (0.0642)
Res parent: professional	0.0575** (0.0277)	0.0464 (0.0343)	0.0556* (0.0298)	0.0774* (0.0400)
Res parent: office worker	0.0514* (0.0281)	0.0918*** (0.0351)	0.0433 (0.0340)	0.1145*** (0.0393)
Res parent: other job	0.0238 (0.0258)	0.0627* (0.0320)	0.0153 (0.0291)	0.0745** (0.0337)
Res parent: job missing	0.2490*** (0.0869)	0.0226 (0.1326)	0.2558*** (0.0918)	0.1389 (0.1528)
Bio parent: college grad	0.1409*** (0.0331)	-0.0564 (0.0394)	0.1375*** (0.0345)	
Own TV time decision	0.0116 (0.0209)	0.1015*** (0.0255)		0.1138*** (0.0268)
Contextual effects	Yes	Yes	Yes	Yes
OIR test p-value	0.300	0.664	0.947	0.393
Cragg-Donald F statistic	18.016	5.367	8.422	8.702

Heteroskedastic-robust standard errors in parentheses.

Statistical significance: ***p<0.01; **p<0.05; *p<0.1.

To save space, estimates of contextual effects are not reported.

Table 2b: Parameter Estimates

	Model 3		Model 4		Model 5	
	GPA	TV	GPA	TV	GPA	TV
Peer GPA	0.4877*** (0.1549)	-0.1320 (0.3835)	0.5876*** (0.2246)		0.5319*** (0.1789)	0.2487 (0.2675)
Peer TV	-0.2283 (0.2116)	0.4545 (0.5609)		0.4155 (0.3471)	-0.2774* (0.1613)	0.1844 (0.4590)
Own GPA				-0.3097 (0.4108)		-0.4488* (0.2579)
Own TV			-0.5608 (0.4678)		-0.0688 (0.1916)	
Age	-0.0659*** (0.0141)	-0.0530** (0.0245)	-0.0940*** (0.0336)	-0.0723*** (0.0270)	-0.0720*** (0.0182)	-0.0876*** (0.0233)
Female	0.2068*** (0.0166)	-0.2645*** (0.0206)	0.0600 (0.1255)	-0.2003** (0.0882)	0.1884*** (0.0533)	-0.1720*** (0.0581)
African American	-0.1035*** (0.0349)	0.2931*** (0.0439)	0.0651 (0.1423)	0.2615*** (0.0593)	-0.0845 (0.0665)	0.2458*** (0.0505)
Other races	-0.0104 (0.0302)	0.1002*** (0.0425)	0.0485 (0.0662)	0.0957*** (0.0406)	-0.0009 (0.0388)	0.1032*** (0.0411)
Junior	0.0532 (0.0467)	-0.1003* (0.0581)	-0.0052 (0.0691)	-0.0846 (0.0599)	0.0491 (0.0501)	-0.0730 (0.0591)
Senior	0.2579*** (0.0588)	-0.0869 (0.0811)	0.2064*** (0.0724)	-0.0093 (0.1165)	0.2588*** (0.0588)	0.0391 (0.0941)
Health	0.1126*** (0.0180)	-0.0623*** (0.0240)	0.0710* (0.0418)	-0.0270 (0.0562)	0.1068*** (0.0240)	-0.0165 (0.0404)
Living condition	0.0733*** (0.0180)	0.0079 (0.0220)	0.0735*** (0.0214)	0.0307 (0.0389)	0.0733*** (0.0182)	0.0390 (0.0303)
Live with both bio parents	0.1521*** (0.0196)	-0.0069 (0.0255)	0.1403*** (0.0258)	0.0405 (0.0717)	0.1503*** (0.0204)	0.0545 (0.0442)
Res parent: HS grad	0.1143*** (0.0279)	0.0104 (0.0350)	0.1132*** (0.0328)	0.0455 (0.0582)	0.1158*** (0.0280)	0.0606 (0.0447)
Res parent: college grad	0.2358*** (0.0336)	-0.0277 (0.0481)	0.2066*** (0.0491)	0.0467 (0.1181)	0.2294*** (0.0368)	0.0673 (0.0864)
Res parent: educ missing	-0.0415 (0.0493)	-0.0465 (0.0686)	-0.0728 (0.0654)	-0.0593 (0.0700)	-0.0464 (0.0512)	-0.0681 (0.0678)
Res parent: professional	0.0559** (0.0285)	0.0487 (0.0361)	0.0802** (0.0386)	0.0663 (0.0446)	0.0588* (0.0301)	0.0715* (0.0402)
Res parent: office worker	0.0447 (0.0298)	0.0980*** (0.0406)	0.0954* (0.0493)	0.1129*** (0.0463)	0.0491 (0.0333)	0.1117*** (0.0436)
Res parent: other job	0.0216 (0.0267)	0.0648* (0.0335)	0.0564 (0.0409)	0.0718** (0.0356)	0.0257 (0.0292)	0.0727** (0.0348)
Res parent: job missing	0.2419*** (0.0898)	0.0294 (0.1369)	0.2535** (0.1139)	0.1054 (0.1736)	0.2290*** (0.0912)	0.1319 (0.1536)
Bio parent: college grad	0.1502*** (0.0348)	-0.0652 (0.0484)	0.1200*** (0.0440)	-0.0202 (0.0632)	0.1497*** (0.0356)	
Own TV time decision	0.0123 (0.0215)	0.1025*** (0.0266)	0.0663 (0.0521)	0.1064*** (0.0276)		0.1062*** (0.0272)
Contextual effects	Yes	Yes	Yes	Yes	Yes	Yes
OIR test p-value	0.374	0.518	0.643	0.650	0.501	0.839
Cragg-Donald F statistic	4.871	1.383	0.796	2.028	2.235	1.374

Heteroskedastic-robust standard errors in parentheses. Statistical significance: ***p<0.01; **p<0.05; *p<0.1.

To save space, estimates of contextual effects are not reported.

Table 3: Marginal Effects of Model 5

	GPA		TV	
	Direct Effects	Indirect Effects	Direct Effects	Indirect Effects
Age	-0.0686*** (0.0153)	-0.0439 (0.0595)	-0.0619*** (0.0200)	-0.0755 (0.0729)
Female	0.2074*** (0.0337)	-0.0053 (0.1818)	-0.2663*** (0.0267)	-0.0249 (0.2214)
African American	-0.1003** (0.0510)	0.0395 (0.1982)	0.2923*** (0.0537)	-0.0089 (0.2609)
Other races	0.0066 (0.0408)	0.1308 (0.1171)	0.1050** (0.0513)	-0.0332 (0.1317)
Junior	0.0540 (0.0578)	0.0846 (0.1632)	-0.0851 (0.0644)	0.2007 (0.1699)
Senior	0.2679*** (0.0769)	0.1810 (0.2666)	-0.0612 (0.0854)	0.2806 (0.3440)
Health	0.1206*** (0.0267)	0.0576 (0.1307)	-0.0706*** (0.0270)	-0.0656 (0.1374)
Living condition	0.0734*** (0.0219)	0.0204 (0.0605)	0.0082 (0.0249)	0.0284 (0.0550)
Live with both bio parents	0.1581*** (0.0245)	0.0723 (0.0947)	-0.0128 (0.0263)	0.0078 (0.0901)
Res parent: HS grad	0.1327*** (0.0384)	0.1553 (0.1394)	0.0068 (0.0446)	-0.0376 (0.1336)
Res parent: college grad	0.2608*** (0.0477)	0.2374 (0.2358)	-0.0426 (0.0462)	-0.0976 (0.2335)
Res parent: educ missing	-0.0420 (0.0596)	0.0218 (0.1357)	-0.0473 (0.0749)	0.0226 (0.1049)
Res parent: professional	0.0503 (0.0355)	-0.0467 (0.0906)	0.0472 (0.0418)	0.0099 (0.0782)
Res parent: office worker	0.0350 (0.0392)	-0.0640 (0.1053)	0.0941* (0.0492)	0.0266 (0.1003)
Res parent: other job	0.0173 (0.0330)	-0.0475 (0.0810)	0.0621 (0.0389)	-0.0144 (0.0786)
Res parent: job missing	0.2480** (0.1085)	0.1700 (0.2679)	0.0255 (0.1496)	-0.0766 (0.2475)
Bio parent: college grad	0.1512*** (0.0432)	-0.0183 (0.1042)	-0.0676 (0.0438)	0.0274 (0.0598)
Own TV time decision	-0.0113 (0.0279)	-0.0402 (0.0671)	0.1092*** (0.0315)	-0.0046 (0.1023)

Standard errors in parentheses. Statistical significance: ***p<0.01; **p<0.05; *p<0.1.

Online Appendices to
“Multivariate Choices and Identification of Social Interactions”

by Ethan Cohen-Cole, Xiaodong Liu, and Yves Zenou

A Model of Social Conformity

The usefulness of the proposed econometric model is not limited to the specific structural model considered in the main text. Here, we present another theoretical model to motivate the econometric model.

Patacchini and Zenou (2012) consider a social conformity model where the social norm is given by the average behavior of peers in a certain activity. We generalize their model by defining the social norm based on the weighted average behavior of two activities. Suppose a set of n individuals interact in a social network. Given the adjacency matrix $\mathbf{G} = [g_{ij}]$, individual i chooses effort levels y_{i1}, y_{i2} simultaneously to maximize her utility function

$$U_i(\mathbf{y}_1, \mathbf{y}_2) = \sum_{k=1}^2 \varpi_{ik} y_{ik} - \frac{1}{2} (\varphi_{11} y_{i1}^2 + 2\varphi_{12} y_{i1} y_{i2} + \varphi_{22} y_{i2}^2) - \frac{1}{2} \sum_{k=1}^2 \rho_k (y_{ik} - \sum_{l=1}^2 \varrho_{lk} \sum_{j=1}^n g_{ij} y_{jl})^2.$$

The first term of the utility captures the payoff from the efforts with the productivity of individual i in activity k given by ϖ_{ik} . The second term is the cost from the efforts with the substitution effect between efforts in different activities captured by φ_{12} . The last term reflects the influence of an individual’s friends on her own behavior. It is such that each individual wants to minimize the social distance between her own behavior y_{ik} to the social norm of that activity. The social norm for activity k is given by the weighted average behavior of her friends in the two activities $\sum_{l=1}^2 \varrho_{lk} \sum_{j=1}^n g_{ij} y_{jl}$ with the weights ϱ_{lk} such that $\varrho_{1k} + \varrho_{2k} = 1$. The coefficient ρ_k captures the taste for conformity.

Maximizing the utility function yields the best response function

$$y_{ik} = \phi_{lk}y_{ik} + \lambda_{kk} \sum_{j=1}^n g_{ij}y_{jk} + \lambda_{lk} \sum_{j=1}^n g_{ij}y_{jl} + \pi_{ik}, \quad \text{for } k = 1, 2 \text{ and } l = 3 - k,$$

where $\phi_{lk} = -\varphi_{12}/(\varphi_{kk} + \rho_k)$, $\lambda_{kk} = \rho_k \varrho_{kk}/(\varphi_{kk} + \rho_k)$, $\lambda_{lk} = \rho_k \varrho_{lk}/(\varphi_{kk} + \rho_k)$, and $\pi_{ik} = \varpi_{ik}/(\varphi_{kk} + \rho_k)$. Let $\pi_{ik} = \mathbf{x}'_i \boldsymbol{\beta}_k + \sum_{j=1}^n g_{ij} \mathbf{x}'_j \boldsymbol{\gamma}_k + \alpha_k + \epsilon_{ik}$ (the network subscript r is suppressed for simplicity). Then, the best response function implies the econometric model considered in this paper.

B Proofs

Proof of Proposition 1. The reduced form of the model is

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{S}^{-1}[\mathbf{X}(\phi_{21}\boldsymbol{\beta}_2 + \boldsymbol{\beta}_1) + \mathbf{G}\mathbf{X}(\lambda_{21}\boldsymbol{\beta}_2 - \lambda_{22}\boldsymbol{\beta}_1 + \phi_{21}\boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_1) + \mathbf{G}^2\mathbf{X}(\lambda_{21}\boldsymbol{\gamma}_2 - \lambda_{22}\boldsymbol{\gamma}_1) \\ &\quad + (\phi_{21} + \lambda_{21})\mathbf{L}\boldsymbol{\alpha}_2 + (1 - \lambda_{22})\mathbf{L}\boldsymbol{\alpha}_1 + (\mathbf{I}_n - \lambda_{22}\mathbf{G})\boldsymbol{\epsilon}_1 + (\phi_{21}\mathbf{I}_n + \lambda_{21}\mathbf{G})\boldsymbol{\epsilon}_2] \\ \mathbf{y}_2 &= \mathbf{S}^{-1}[\mathbf{X}(\phi_{12}\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2) + \mathbf{G}\mathbf{X}(\lambda_{12}\boldsymbol{\beta}_1 - \lambda_{11}\boldsymbol{\beta}_2 + \phi_{12}\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2) + \mathbf{G}^2\mathbf{X}(\lambda_{12}\boldsymbol{\gamma}_1 - \lambda_{11}\boldsymbol{\gamma}_2) \\ &\quad + (\phi_{12} + \lambda_{12})\mathbf{L}\boldsymbol{\alpha}_1 + (1 - \lambda_{11})\mathbf{L}\boldsymbol{\alpha}_2 + (\mathbf{I}_n - \lambda_{11}\mathbf{G})\boldsymbol{\epsilon}_2 + (\phi_{12}\mathbf{I}_n + \lambda_{12}\mathbf{G})\boldsymbol{\epsilon}_1], \end{aligned} \quad (\text{B.1})$$

where

$$\mathbf{S} = (1 - \phi_{12}\phi_{21})\mathbf{I}_n - (\lambda_{11} + \lambda_{22} + \phi_{21}\lambda_{12} + \phi_{12}\lambda_{21})\mathbf{G} + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\mathbf{G}^2. \quad (\text{B.2})$$

$\text{E}(\mathbf{J}\mathbf{Z}_1|\mathbf{X}) = [\text{E}(\mathbf{J}\mathbf{y}_2|\mathbf{X}), \text{E}(\mathbf{J}\mathbf{G}\mathbf{y}_1|\mathbf{X}), \text{E}(\mathbf{J}\mathbf{G}\mathbf{y}_2|\mathbf{X}), \mathbf{J}\mathbf{X}, \mathbf{J}\mathbf{G}\mathbf{X}]$ has full column rank if and only if

$$\text{E}(\mathbf{J}\mathbf{y}_2|\mathbf{X})d_1 + \text{E}(\mathbf{J}\mathbf{G}\mathbf{y}_1|\mathbf{X})d_2 + \text{E}(\mathbf{J}\mathbf{G}\mathbf{y}_2|\mathbf{X})d_3 + \mathbf{J}\mathbf{X}d_4 + \mathbf{J}\mathbf{G}\mathbf{X}d_5 = \mathbf{0} \quad (\text{B.3})$$

implies that $d_1 = d_2 = d_3 = 0$ and $\mathbf{d}_4 = \mathbf{d}_5 = \mathbf{0}$. As $\mathbf{JGJ} = \mathbf{JG}$, $\mathbf{JSJ} = \mathbf{JS}$ and $\mathbf{SG} = \mathbf{GS}$, if we premultiply (B.3) by \mathbf{JS} , then it follows from the reduced form equations (B.1) that

$$\mathbf{JX}\eta_1 + \mathbf{JGX}\eta_2 + \mathbf{JG}^2\mathbf{X}\eta_3 + \mathbf{JG}^3\mathbf{X}\eta_4 = \mathbf{0}$$

where

$$\eta_1 = (\phi_{12}\beta_1 + \beta_2)d_1 + (1 - \phi_{12}\phi_{21})\mathbf{d}_4$$

$$\begin{aligned} \eta_2 = & (\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \phi_{12}\gamma_1 + \gamma_2)d_1 + (\beta_1 + \phi_{21}\beta_2)d_2 + (\phi_{12}\beta_1 + \beta_2)d_3 \\ & - (\lambda_{11} + \lambda_{22} + \lambda_{12}\phi_{21} + \lambda_{21}\phi_{12})\mathbf{d}_4 + (1 - \phi_{12}\phi_{21})\mathbf{d}_5 \end{aligned}$$

$$\begin{aligned} \eta_3 = & (\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2)d_1 + (\lambda_{21}\beta_2 - \lambda_{22}\beta_1 + \phi_{21}\gamma_2 + \gamma_1)d_2 + (\lambda_{12}\beta_1 - \lambda_{11}\beta_2 + \phi_{12}\gamma_1 + \gamma_2)d_3 \\ & + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\mathbf{d}_4 - (\lambda_{11} + \lambda_{22} + \lambda_{12}\phi_{21} + \lambda_{21}\phi_{12})\mathbf{d}_5 \end{aligned}$$

$$\eta_4 = (\lambda_{21}\gamma_2 - \lambda_{22}\gamma_1)d_2 + (\lambda_{12}\gamma_1 - \lambda_{11}\gamma_2)d_3 + (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\mathbf{d}_5.$$

If $d_2 = (\lambda_{12} + \lambda_{11}\phi_{12})d_1/(\phi_{12}\phi_{21} - 1)$, $d_3 = (\lambda_{22} + \lambda_{21}\phi_{12})d_1/(\phi_{12}\phi_{21} - 1)$, $\mathbf{d}_4 = (\phi_{12}\beta_1 + \beta_2)d_1/(\phi_{12}\phi_{21} - 1)$ and $\mathbf{d}_5 = (\phi_{12}\gamma_1 + \gamma_2)d_1/(\phi_{12}\phi_{21} - 1)$, then (B.3) holds. Therefore, $E(\mathbf{JZ}_1|\mathbf{X})$ does not have full column rank. Similarly, $E(\mathbf{JZ}_2|\mathbf{X})$ does not have full column rank. ■

Proof of Proposition 2. The identification of the structural parameters takes two steps. In the first step, we show that the pseudo reduced form parameters can be identified under Assumption 1. In the second step, we show that the structural parameters can be identified from the pseudo reduced form parameters under Assumption 2.

Step 1. The proof follows a similar argument as in Bramoullé, Djebbari and Fortin (2009). We first show that, under Assumption 1 (ii), $\mu_0\mathbf{I}_n + \mu_1\mathbf{G} + \mu_2\mathbf{G}^2 + \mu_3\mathbf{G}^3$ has identical rows implies $\mu_0 = \mu_1 = \mu_2 = \mu_3 = 0$. If $\mu_0\mathbf{I}_n + \mu_1\mathbf{G} + \mu_2\mathbf{G}^2 + \mu_3\mathbf{G}^3$ has identical

rows, then

$$\mu_0 \boldsymbol{\iota}_n + \mu_1 \mathbf{G} \boldsymbol{\iota}_n + \mu_2 \mathbf{G}^2 \boldsymbol{\iota}_n + \mu_3 \mathbf{G}^3 \boldsymbol{\iota}_n = c_0 \boldsymbol{\iota}_n, \quad (\text{B.4})$$

for some constant c_0 . As $\mathbf{G} \boldsymbol{\iota}_n = \boldsymbol{\iota}_n$, multiplying both sides of (B.4) by \mathbf{G} gives

$$\mu_0 \mathbf{G} \boldsymbol{\iota}_n + \mu_1 \mathbf{G}^2 \boldsymbol{\iota}_n + \mu_2 \mathbf{G}^3 \boldsymbol{\iota}_n + \mu_3 \mathbf{G}^4 \boldsymbol{\iota}_n = c_0 \boldsymbol{\iota}_n. \quad (\text{B.5})$$

Subtracting (B.4) from (B.5) gives $\mu_0 \boldsymbol{\iota}_n + (\mu_1 - \mu_0) \mathbf{G} \boldsymbol{\iota}_n + (\mu_2 - \mu_1) \mathbf{G}^2 \boldsymbol{\iota}_n + (\mu_3 - \mu_2) \mathbf{G}^3 \boldsymbol{\iota}_n - \mu_3 \mathbf{G}^4 \boldsymbol{\iota}_n = \mathbf{0}$, which implies $\mu_0 = \mu_1 = \mu_2 = \mu_3 = 0$ under Assumption 1 (ii).

The moment conditions $\mathbb{E}(\mathbf{J} \boldsymbol{\epsilon}_1 | \mathbf{X}) = \mathbb{E}(\mathbf{J} \boldsymbol{\epsilon}_2 | \mathbf{X}) = \mathbf{0}$ imply that

$$\begin{aligned} \mathbb{E}(\mathbf{J} \mathbf{y}_1 | \mathbf{X}) &= \lambda_{11}^* \mathbb{E}(\mathbf{J} \mathbf{G} \mathbf{y}_1 | \mathbf{X}) + \lambda_{21}^* \mathbb{E}(\mathbf{J} \mathbf{G} \mathbf{y}_2 | \mathbf{X}) + \mathbf{J} \mathbf{X} \boldsymbol{\beta}_1^* + \mathbf{J} \mathbf{G} \mathbf{X} \boldsymbol{\gamma}_1^* \\ \mathbb{E}(\mathbf{J} \mathbf{y}_2 | \mathbf{X}) &= \lambda_{22}^* \mathbb{E}(\mathbf{J} \mathbf{G} \mathbf{y}_2 | \mathbf{X}) + \lambda_{12}^* \mathbb{E}(\mathbf{J} \mathbf{G} \mathbf{y}_1 | \mathbf{X}) + \mathbf{J} \mathbf{X} \boldsymbol{\beta}_2^* + \mathbf{J} \mathbf{G} \mathbf{X} \boldsymbol{\gamma}_2^*. \end{aligned}$$

Let $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^{*'}, \boldsymbol{\theta}_2^{*'})'$ with $\boldsymbol{\theta}_k^* = (\lambda_{kk}^*, \lambda_{lk}^*, \boldsymbol{\beta}_k^{*'}, \boldsymbol{\gamma}_k^{*'})'$, for $k = 1, 2$ and $l = 3 - k$. If $[\mathbb{E}(\mathbf{J} \mathbf{G} \mathbf{y}_1 | \mathbf{X}), \mathbb{E}(\mathbf{J} \mathbf{G} \mathbf{y}_2 | \mathbf{X}), \mathbf{J} \mathbf{X}, \mathbf{J} \mathbf{G} \mathbf{X}]$ has full column rank, then $\boldsymbol{\theta}^*$ and $\tilde{\boldsymbol{\theta}}^*$ leading to the same $\mathbb{E}(\mathbf{J} \mathbf{y}_1 | \mathbf{X})$ and $\mathbb{E}(\mathbf{J} \mathbf{y}_2 | \mathbf{X})$ implies $\boldsymbol{\theta}^* = \tilde{\boldsymbol{\theta}}^*$, i.e. $\boldsymbol{\theta}^*$ is identified. $[\mathbb{E}(\mathbf{J} \mathbf{G} \mathbf{y}_1 | \mathbf{X}), \mathbb{E}(\mathbf{J} \mathbf{G} \mathbf{y}_2 | \mathbf{X}), \mathbf{J} \mathbf{X}, \mathbf{J} \mathbf{G} \mathbf{X}]$ has full column rank if

$$\mathbb{E}(\mathbf{J}_r \mathbf{G}_r \mathbf{y}_{1,r} | \mathbf{X}_r) d_1 + \mathbb{E}(\mathbf{J}_r \mathbf{G}_r \mathbf{y}_{2,r} | \mathbf{X}_r) d_2 + \mathbf{J}_r \mathbf{X}_r \mathbf{d}_3 + \mathbf{J}_r \mathbf{G}_r \mathbf{X}_r \mathbf{d}_4 = \mathbf{0} \quad (\text{B.6})$$

implies that $d_1 = d_2 = 0$ and $\mathbf{d}_3 = \mathbf{d}_4 = \mathbf{0}$, for some network r . The pseudo reduced form equations imply

$$\begin{aligned} \mathbb{E}(\mathbf{J}_r \mathbf{y}_{1,r} | \mathbf{X}_r) &= \mathbf{J}_r \mathbf{S}_r^{*-1} [\mathbf{X}_r \boldsymbol{\beta}_1^* + \mathbf{G}_r \mathbf{X}_r (\lambda_{21}^* \boldsymbol{\beta}_2^* - \lambda_{22}^* \boldsymbol{\beta}_1^* + \boldsymbol{\gamma}_1^*) + \mathbf{G}_r^2 \mathbf{X}_r (\lambda_{21}^* \boldsymbol{\gamma}_2^* - \lambda_{22}^* \boldsymbol{\gamma}_1^*)] \\ \mathbb{E}(\mathbf{J}_r \mathbf{y}_{2,r} | \mathbf{X}_r) &= \mathbf{J}_r \mathbf{S}_r^{*-1} [\mathbf{X}_r \boldsymbol{\beta}_2^* + \mathbf{G}_r \mathbf{X}_r (\lambda_{12}^* \boldsymbol{\beta}_1^* - \lambda_{11}^* \boldsymbol{\beta}_2^* + \boldsymbol{\gamma}_2^*) + \mathbf{G}_r^2 \mathbf{X}_r (\lambda_{12}^* \boldsymbol{\gamma}_1^* - \lambda_{11}^* \boldsymbol{\gamma}_2^*)] \end{aligned}$$

where $\mathbf{S}_r^* = \mathbf{I}_{n_r} - (\lambda_{11}^* + \lambda_{22}^*)\mathbf{G}_r + (\lambda_{11}^*\lambda_{22}^* - \lambda_{12}^*\lambda_{21}^*)\mathbf{G}_r^2$. As $\mathbf{J}_r\mathbf{G}_r\mathbf{J}_r = \mathbf{J}_r\mathbf{G}_r$, $\mathbf{J}_r\mathbf{S}_r^*\mathbf{J}_r = \mathbf{J}_r\mathbf{S}_r^*$ and $\mathbf{S}_r^*\mathbf{G}_r = \mathbf{G}_r\mathbf{S}_r^*$, premultiplying (B.6) by $\mathbf{J}_r\mathbf{S}_r^*$ gives

$$\sum_{h=1}^p (\eta_{0,h}\mathbf{I}_{n_r} + \eta_{1,h}\mathbf{G}_r + \eta_{2,h}\mathbf{G}_r^2 + \eta_{3,h}\mathbf{G}_r^3)\mathbf{x}_{r,h} = c_1\mathbf{t}_{n_r} \quad (\text{B.7})$$

where $\mathbf{x}_{r,h}$ is the h -th column of \mathbf{X}_r ,

$$\begin{aligned} \boldsymbol{\eta}_0 &= (\eta_{0,1}, \dots, \eta_{0,p})' = \mathbf{d}_3 \\ \boldsymbol{\eta}_1 &= (\eta_{1,1}, \dots, \eta_{1,p})' = \boldsymbol{\beta}_1^*d_1 + \boldsymbol{\beta}_2^*d_2 - (\lambda_{11}^* + \lambda_{22}^*)\mathbf{d}_3 + \mathbf{d}_4 \\ \boldsymbol{\eta}_2 &= (\eta_{2,1}, \dots, \eta_{2,p})' = (\lambda_{21}^*\boldsymbol{\beta}_2^* - \lambda_{22}^*\boldsymbol{\beta}_1^* + \boldsymbol{\gamma}_1^*)d_1 + (\lambda_{12}^*\boldsymbol{\beta}_1^* - \lambda_{11}^*\boldsymbol{\beta}_2^* + \boldsymbol{\gamma}_2^*)d_2 \\ &\quad + (\lambda_{11}^*\lambda_{22}^* - \lambda_{12}^*\lambda_{21}^*)\mathbf{d}_3 - (\lambda_{11}^* + \lambda_{22}^*)\mathbf{d}_4 \\ \boldsymbol{\eta}_3 &= (\eta_{3,1}, \dots, \eta_{3,p})' = (\lambda_{21}^*\boldsymbol{\gamma}_2^* - \lambda_{22}^*\boldsymbol{\gamma}_1^*)d_1 + (\lambda_{12}^*\boldsymbol{\gamma}_1^* - \lambda_{11}^*\boldsymbol{\gamma}_2^*)d_2 + (\lambda_{11}^*\lambda_{22}^* - \lambda_{12}^*\lambda_{21}^*)\mathbf{d}_4 \end{aligned}$$

and $c_1 = n_r^{-1}\mathbf{t}_{n_r}'(\mathbf{X}_r\boldsymbol{\eta}_0 + \mathbf{G}_r\mathbf{X}_r\boldsymbol{\eta}_1 + \mathbf{G}_r^2\mathbf{X}_r\boldsymbol{\eta}_2 + \mathbf{G}_r^3\mathbf{X}_r\boldsymbol{\eta}_3)$. As (B.7) holds for all possible realizations of $\mathbf{x}_{r,h}$, $\eta_{0,h}\mathbf{I}_{n_r} + \eta_{1,h}\mathbf{G}_r + \eta_{2,h}\mathbf{G}_r^2 + \eta_{3,h}\mathbf{G}_r^3$ has identical rows. Therefore, $\boldsymbol{\eta}_0 = \boldsymbol{\eta}_1 = \boldsymbol{\eta}_2 = \boldsymbol{\eta}_3 = \mathbf{0}$, which implies that $d_1 = d_2 = 0$ and $\mathbf{d}_3 = \mathbf{d}_4 = \mathbf{0}$ under Assumption 1 (i). Hence, $[\mathbf{E}(\mathbf{J}\mathbf{G}\mathbf{y}_1|\mathbf{X}), \mathbf{E}(\mathbf{J}\mathbf{G}\mathbf{y}_2|\mathbf{X}), \mathbf{J}\mathbf{X}, \mathbf{J}\mathbf{G}\mathbf{X}]$ has full column rank and thus $\boldsymbol{\theta}^*$ is identified.

Step 2. Under Assumption 2, the identification of the structural parameters from the pseudo reduced form parameters follows the same argument as in a classical simultaneous-equation model (see, e.g., Schmidt, 1976), and thus the proof is omitted here. ■

References

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

Patacchini, E. and Zenou, Y. (2012). Juvenile delinquency and conformism, *Journal of Law, Economics, and Organization* **28**: 1–31.

Schmidt, P. (1976). *Econometrics*, Marcel Dekker.