Who is the Key Player?

A Network Analysis of Juvenile Delinquency*

Lung-Fei Lee† Xiaodong Liu‡ Eleonora Patacchini§ Yves Zenou¶

February 21, 2018

Abstract

We generalize the theoretical model of Ballester et al. (2006) and develop a novel GMM estimation approach that is robust to potential endogenous networks. Using unique information on friendship networks among U.S. teenagers, we determine the key player in delinquent activities. We show that, compared to a policy that removes the most active delinquent from the network, a key player policy engenders a much higher delinquency reduction.

Key words: Network centrality measures, linear social interaction models, crime policies.

JEL Classification: A14, C31, D85, K42, Z13

*A previous version of this paper circulated with the title: "Criminal Networks: Who is the Key Player?" (see Liu et al. 2012). 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.

†The Ohio State University, USA. E-mail: lee.1777@osu.edu.

‡University of Colorado Boulder, USA. E-mail: xiaodong.liu@colorado.edu.

§Cornell University, EIEF and CEPR. E-mail: ep454@cornell.edu.

¶Corresponding author. Monash University, IFN and CEPR. E-mail: yves.zenou@monash.edu.

1 Introduction

There are 2.3 million people behind bars in the United States, and that number continues to grow. It is the highest level of incarceration per capita in the world. Moreover, since the crime explosion of the 1960s, the prison population in the United States has multiplied fivefold, to one prisoner for every hundred adults — a rate unprecedented in American history and unmatched anywhere in the world.[1] However, in spite of a continuously falling crime rate, the prisoner head count continues to rise, and poor people as well as minorities still bear the brunt of both crime and punishment.

One possible way to reduce crime is to detect, apprehend, convict, and punish criminals. This is what has been done in the United States and all of these actions cost money, currently about $200 billion per year nationwide. For example, in California, even if this "brute force" policy has partly worked (the rate of every major crime category is now less than half of what it was 20 years ago), the cost of this policy has been tremendous. For example, the cost of the justice system is higher than the cost of education.[2]

In his book published in 2009, Mark Kleiman argues that simply locking up more people for lengthier terms is no longer a workable crime-control strategy. But, says Kleiman, there has been a revolution in controlling crime by means other than brute-force incarceration: substituting swiftness and certainty of punishment for randomized severity, concentrating enforcement resources rather than dispersing them, communicating specific threats of punishment to specific offenders, and enforcing probation and parole conditions to make community corrections a genuine alternative to incarceration. As Kleiman shows, "zero tolerance" is nonsense: there are always more offenses than there is punishment capacity.

Is there an alternative to brute force? In this paper, we argue that concentrating efforts by targeting "key players", i.e. players who once removed generate the highest possible reduction in aggregate activity level in a network. The "key players" policy can be more effective in reducing delinquency and crime because of the snow-ball effects or "social multipliers" at work (see, in particular, Glaeser et al., 1996; Calvó-Armengol and Zenou, 2004; Verdier and Zenou, 2004; Kleiman, 2009). Furthermore, the impact of social networks may be particularly important for adolescents because this developmental period overlaps with the initiation and continuation of many risky, un-

[1] See Cook and Ludwig (2010) and the references therein.

[2] For example, the "Three Strikes" law passed in California in 1994 mandates extremely long prison terms (between 29 years and life) for anyone previously convicted in two serious or violent felonies (including residential burglary) when convicted of a third felony, even for something as minor as petty theft.

healthy, and delinquent behaviors and is a period of maximal response to peer pressure (Warr, 2002; Thornberry et al., 2003).

It is indeed well-established that delinquency and crime are, to some extent, a group phenomenon, and that the sources of delinquency and crime are located in the intimate social networks of individuals (see e.g. Haynie, 2001; Sarnecki, 2001; Patacchini and Zenou, 2012; Warr, 2002). Delinquents often have friends who have themselves committed several offences, and social ties among delinquents are seen as a means whereby individuals exert an influence over one another to commit crimes. In fact, not only friends, but also the *structure* of social networks, matters in explaining an individual's own delinquent behavior. This suggests that the underlying structural properties of friendship networks must be taken into account to better understand the impact of peer influence on delinquent behavior and to design adequate and novel delinquency-reducing policies.

In the fast growing literature of network models (see Jackson, 2008; Jackson and Zenou, 2015; Jackson et al., 2017, for overviews), two types of linear social interaction models are the most popular, namely, the *local-aggregate model* and the *local-average model*. In the local-aggregate model (see, in particular, Ballester et al., 2006, 2010; Bramoullé and Kranton, 2007; Galeotti et al., 2009; Calvó-Armengol et al., 2009; Bramoullé et al., 2014), endogenous peer effects are captured by the sum of friends' efforts in some activity so that the more active friends an individual has, the higher is her marginal utility of exerting effort. In the local-average model (e.g. Glaeser and Scheinkman, 2003; Patacchini and Zenou, 2012; Boucher et al., 2014), peers' choices are viewed as a social norm and individuals pay a cost for deviating from it. In this model, each individual wants to conform as much as possible to the social norm of her reference group, which is defined as the average effort of her friends.

The aim of this paper is to generalize the model of Ballester et al. (2006, 2010) to incorporate local-aggregate (i.e. strategic complementarity) as well as local-average (i.e. taste for conformity) effects and develop a novel methodology to bring the generalized model to the data. Indeed, to determine the key player, we need to estimate the intensity of interactions between connected individuals in a social network. The conventional two stage least square (2SLS) estimation method uses instrumental variables (IVs) based on the network structure and individual characteristics to account for the endogenous interaction effects (see, e.g. Bramoullé et al., 2009; Calvó-Armengol et al., 2009; Lin, 2010). However, the validity of the IV strategy relies on the assumption that the network structure is *exogenous*. If the formation of the network depends on some unobservable covariates that are correlated with the error term of the social interaction equation, then the adjacency matrix is endogenous and the aforementioned IVs would be invalid. To overcome this problem, we use the predicted adjacency matrix based on exogenous dyadic characteristics to construct the IVs for the social interaction equation. Furthermore, to alleviate the potential weak IV problem when the exogenous dyadic characteristics are not very informative in explaining the formation of the network, we introduce some additional quadratic moment conditions based on the correlation structure of the error term and propose a generalized method of moments (GMM) estimator that utilizes both linear and quadratic moment conditions. We investigate the performance of the proposed estimator in Monte Carlo simulations.

Using information on friendship networks among U.S. teenagers from the National Longitudinal Study of Adolescent Health (Add Health), we apply the proposed estimation strategy to determine who is the teenager whose removal is associated with the largest reduction in the aggregate delinquency level of her network. We find that the key player is *not* necessarily the most active delinquent in the network. We also find that it is *not* straightforward to determine which delinquent should be removed from a network based on other standard network centrality measures (which are not microfounded). In at least 60% of the networks with more than 50 members, the key player is not the student with the highest betweenness centrality, closeness centrality, or eigenvector centrality. Finally, we look at the portray of the "key" adolescent in juvenile delinquency. Perhaps surprisingly, we do not observe that the most disruptive kids are those coming from low-income families (as measured by having at least one parent on welfare) and low-educated parents. We do see that, compared to other students, the key players in crime are more likely to come from single-parent families, but on average their parents have college and above education level, work as professionals, and do not receive public assistance. We also find that key players have more friends than other students, and their friends are more likely to come from single-parent families and tend to be non-whites.

We show that targeting the most active delinquent is less effective than the key-player policy. Our key-player policy can be applied to reduce crime in the real world when criminal network data are available. In fact, similar policies aiming at reducing crime have already been implemented in the U.S. but without the analytical tools of the key-player policy (see e.g. Kennedy, 1998, 2008). It can also be used to determine the key player in other contexts. We discuss the case of financial networks, R&D networks, social networks in developing countries and political networks.

The rest of the paper unfolds as follows. In the next section, we discuss the related literature and explain our contribution. In Section 3, we introduce the network game, characterize the equilibrium, and define the key player. Our estimation strategy is described in Section 4. Section 5 presents the empirical results. In Section 6, we discuss the implications of the key-player policy. Finally, Section 7 concludes.

2 Related literature

Our paper lies at the intersection of different literatures.

Econometrics of networks The literature on identification and estimation of social network models has recently significantly progressed (see Durlauf and Ioannides, 2010; Blume et al., 2011; Graham, 2015; de Paula, 2017; Dolton, 2017; Advani and Malde, 2018; Graham and de Paula, 2018; for recent surveys). In spite of the new advances in network econometrics, the estimation of linear social interaction models with network data is still the most used practice in applied work. Bramoullé et al. (2009) provide identification conditions for this model based on the intransitivities in the network structure and propose an 2SLS estimation strategy exploiting exogenous characteristics of indirect connections. This estimation strategy gains its popularity due to its simplicity. Yet, it has some embedded problems. First, the validity of the IVs relies on the assumption that the network structure captured by the adjacency matrix is exogenous. If the formation of the network depends on some unobservable covariates that are correlated with the error term of the social interaction equation, then the adjacency matrix is endogenous and this IV-based estimator would be inconsistent. Second, when the network is dense, it becomes difficult to find an indirect connection for each individual that is not a direct connection or herself. As a result, the IVs based on exogenous characteristics of indirect connections could be weak.

In this paper, we propose a new estimation strategy to overcome these two issues. First, to avoid the invalidity of IVs due to endogenous networks, we use the predicted adjacency matrix based on exogenous dyadic characteristics to construct IVs for the social interaction equation. Second, to alleviate the potential weak IV problem, we introduce some additional quadratic moment conditions based on the correlation structure of the error term and propose a GMM estimator utilizing both linear and quadratic moment conditions. We investigate the performance of the proposed estimator in Monte Carlo simulations and find the proposed GMM estimator performs well when the adjacency matrix is endogenous and the IVs are weak.

The key-player problem The problem of identifying key players has a long tradition in the sociological literature, which have proposed different measures of network centralities to define the key players (see, in particular, Wasserman and Faust, 1994). Borgatti (2003, 2006) was among the first researchers to analytically study the issue of key players by explicitly measuring the contribution of a set of actors to the cohesion of a network. Borgatti measures the amount of *reduction* in cohesiveness of the network that would occur if some nodes were not present. In the economics literature, Ballester et al. (2006) were the first to define the key-player problem in terms of *behavior* of agents so that total activity is measured as the sum of efforts of all agents at a Nash equilibrium (see Zenou, 2016, for an overview). Ballester et al. (2010) show that this model is well suited to analyzing key player policies aiming at reducing crime.

The (few) recent papers aiming at identifying key players in networks are based on estimation strategies that either do not directly tackle the issue of network endogeneity (e.g. Lindquist and Zenou, 2014; Denbee et al., 2017) or achieve identification exploiting context specific information (e.g. König et al., 2014; König et al., 2017). In particular, König et al. (2014) combine datasets of R&D alliances and firms' annual balance sheets and income statements to study technology spillovers and identify key players in R&D networks. Their identification strategy exploits the panel structure of the data to use time-lagged R&D collaborations and time-lagged R&D tax credits to construct valid IVs. König et al. (2017) use data from the Second Congo War, a conflict over a complex network of alliances and rivalries across many ethnic groups, to study to which extent the removal of each ethnic group involved in the conflict would reduce the conflict intensity. Their identification strategy exploits the exogenous variation in the average weather conditions facing, respectively, the set of allies and of enemies of each ethnic group.

This paper develops a methodology to bring the model in Ballester et al. (2006, 2010) to the data taking into account potential network endogeneity, and it does so in a way that can be of immediate applicability for the practitioner. Monte Carlo simulations show that the proposed estimation strategy work well without additional (context specific) IVs for the network formation process. Therefore, our methodology can be easily applied to other contexts. Similarly to the other papers cited above, our key player policy abstracts from the possibility that new links can be formed (or ceased) in response to the policy itself. As such, our key player analysis should be qualified as a short-run policy, that is it helps detecting the key player for a given network topology.[3]

[3] The long-run key player analysis that takes into account the dynamic evolution of the network would be sensitive to

Empirical studies of peer effects in crime There is a growing body of empirical literature suggesting that peer effects are very strong in criminal decisions. Ludwig et al. (2001) and Kling et al. (2005) study the relocation of families from high- to low-poverty neighborhoods using data from the Moving to Opportunity (MTO) experiment. They find that this policy reduces juvenile arrests for violent offences by 30 to 50 percent, relative to a control group. This also suggests very strong social interactions in crime behaviors. Patacchini and Zenou (2012) find that peer effects in crime are strong, especially for petty crimes. Bayer et al. (2009) consider the influence that juvenile offenders serving time in the same correctional facility have on each other's subsequent criminal behavior. They also find strong evidence of learning effects in criminal activities since exposure to peers with a history of committing a particular crime increases the probability that an individual who has already committed the same type of crime recidivates that crime. More recently, Damm and Dustmann (2014) and Corneo (2017) investigate the influence of peers on crime. The former exploit a Danish natural experiment that randomly allocates parents of young children to neighborhoods with different shares of youth criminals while the latter uses data collected among the homeless. Both find strong peer effects in crime.

In this paper, we provide a micro-foundation for a structural model that distinguishes local-aggregate (i.e. strategic complementarity) and local-average (i.e. taste for conformity) peer effects. We propose a novel estimation strategy for this network model taking into account potentially endogenous networks. We bring this model to the data and find evidence of local-aggregate peer effects in juvenile delinquency.

3 Network Game, Equilibrium, and Key Player

3.1 Network and utility

Consider a finite set of agents $\mathcal{N} = \{1, \ldots, n\}$ with each agent corresponding to a node in a network $\mathcal{G}$. We keep track of social connections in network $\mathcal{G}$ through its adjacency matrix $\mathbf{G} = [g_{ij}]$, where $g_{ij} = 1$ if nodes i and j ($i \neq j$) are connected and $g_{ij} = 0$ otherwise.[4] We set $g_{ii} = 0$. Furthermore, let $\mathbf{G}^* = [g^*_{ij}]$, with $g^*_{ij} = g_{ij} / \sum_{j=1}^n g_{ij}$,[5] denote the row-normalized adjacency matrix.

the underlying network formation model assumed. Despite the recent development of the literature (see, in particular, Sheng, 2016; Graham, 2017; and Mele, 2017), there is not a consensus way to model the dynamic network formation process. We introduce a simple homophily network formation model in Section 4.2 only to motivate instruments. This model is not sophisticated enough for the purpose of conducting a long-run key player analysis.

[4] For the ease of the presentation, we focus on undirected unweighted networks so that $\mathbf{G}$ is a symmetric square matrix. All our theoretical results also hold for directed and/or weighted networks.

[5] For simplicity, we assume none of the agents in the network are isolated so that $\sum_{j=1}^n g_{ij} \neq 0$ for all i.

Agents in network $\mathcal{G}$ decide how much effort to exert in delinquent activities. We denote by y_i the effort level of agent i and by $\mathbf{y} = (y_1, ..., y_n)'$ the population effort profile in network $\mathcal{G}$. Each agent i selects an effort $y_i \geq 0$, and obtains a payoff $U_i(\mathbf{y}, \mathcal{G})$ that depends on the effort profile $\mathbf{y}$ and on the underlying network $\mathcal{G}$ in the following way:

$$U_i(\mathbf{y}, \mathcal{G}) = \underbrace{(\varpi_i + \lambda_1 \textstyle\sum_{j \in \mathcal{N}} g_{ij} y_j) y_i}_{\text{payoff}} - \underbrace{[p\, f\, y_i + \tfrac{1}{2} y_i^2 + \tfrac{1}{2}\lambda_2 (y_i - \textstyle\sum_{j \in \mathcal{N}} g_{ij}^* y_j)^2]}_{\text{cost}} \tag{1}$$

with $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. This utility has a standard cost-payoff structure as in Becker (1968). The payoff increases in own effort y_i with the marginal payoff given by $\varpi_i + \lambda_1 \sum_{j=1}^n g_{ij} y_j$. The term ϖ_i represents the *exogenous heterogeneity* of agent i's "productivity" in delinquent activities, and is given by

$$\varpi_i = \mathbf{x}_i' \boldsymbol{\gamma}_1 + \textstyle\sum_{j \in \mathcal{N}} g_{ij}^* \mathbf{x}_j' \boldsymbol{\gamma}_2 + \xi + v_i.$$

where $\mathbf{x}_i$ is a $K \times 1$ vector of *observable* exogenous characteristics of agent i (e.g. age, sex, race, parental education, etc.), $\sum_{j=1}^n g_{ij}^* \mathbf{x}_j'$ is the *average* exogenous characteristics of agent i's connections, with its coefficient vector $\boldsymbol{\gamma}_2$ representing *contextual effects* (Manski, 1993), ξ denotes the *unobservable* (to the econometrician) exogenous network characteristics (e.g., the prosperity level of the neighborhood of network $\mathcal{G}$), and v_i denotes the *unobservable* (to the econometrician) characteristics of agent i.

Compared to the standard crime model (Becker, 1968), the marginal payoff in the utility function (1) has a new component $\lambda_1 \sum_{j=1}^n g_{ij} y_j$, which reflects the impact of the total effort of an agent's connections on her own "productivity". Indeed, an agent may benefit directly from the effort of her connections if they are co-offenders in some delinquent activity. An agent may also benefit indirectly through the form of know-how sharing about delinquent behavior with her friends.[6] We assume that the more delinquent connections an agent has and the more these connections are involved in delinquent activities, the higher is the marginal payoff of the agent's own delinquent effort. Thus, we call λ_1 the *social-multiplier* coefficient.

The cost part of the utility function (1) has three components. The cost of being caught is captured by the probability of being caught $0 < p < 1$ times the fine $f\, y_i$, which increases with

[6] Sutherland (1947) and Akers (1998) argue that criminal behavior is learned from others in the same way that *all* human behavior is learned. Indeed, young people may be influenced by their peers in all categories of behavior — music, speech, dress, sports, and *delinquency*.

the effort level y_i, as the severity of the punishment increases with one's involvement in delinquent activities. Also, individuals have a *direct* cost of exerting effort given by $\frac{1}{2}y_i^2$. Finally, different from Ballester et al. (2006, 2010), the cost in the utility function (1) has an additional term $\frac{1}{2}\lambda_2(y_i - \sum_{j=1}^n g_{ij}^* y_j)^2$, which represents the moral cost due to deviation from the *social norm* of the reference group (i.e., the average effort of agent i's connections). We call λ_2 the *social-conformity* coefficient.[7]

3.2 Equilibrium characterization

In equilibrium, each agent maximizes her utility and the best-response function is given by:

$$y_i = \phi_1 \sum_{j\in\mathcal{N}} g_{ij}y_j + \phi_2 \sum_{j\in\mathcal{N}} g_{ij}^* y_j + \pi_i, \tag{2}$$

with

$$\pi_i = \mathbf{x}_i'\boldsymbol{\beta}_1 + \sum_{j\in\mathcal{N}} g_{ij}^*\mathbf{x}_j'\boldsymbol{\beta}_2 + \eta + u_i, \tag{3}$$

where $\phi_1 = \lambda_1/(1+\lambda_2)$, $\phi_2 = \lambda_2/(1+\lambda_2)$, $\boldsymbol{\beta}_1 = \boldsymbol{\gamma}_1/(1+\lambda_2)$, $\boldsymbol{\beta}_2 = \boldsymbol{\gamma}_2/(1+\lambda_2)$, $\eta = (\xi - p\,f)/(1+\lambda_2)$ and $u_i = v_i/(1+\lambda_2)$. As $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$, we have $\phi_1 \geq 0$ and $0 \leq \phi_2 < 1$. The coefficient ϕ_1 captures the *local-aggregate* endogenous peer effect. As $\phi_1 \geq 0$, this coefficient reflects *strategic complementarity* in efforts. The coefficient ϕ_2 captures the *local-average* endogenous peer effect, which reflects the *taste for conformity.* Note that, $\phi_1/\phi_2 = \lambda_1/\lambda_2$. That is, the relative magnitude of ϕ_1 and ϕ_2 is the same as that of the social-multiplier coefficient λ_1 and the social-conformity coefficient λ_2.

Let $\rho(\mathbf{A})$ denote the spectral radius of a square matrix $\mathbf{A}$, and $\mathbf{I}_n$ denote the $n \times n$ identity matrix. The following proposition provides a complete characterization of the equilibrium effort profile. Its proof can be found in Appendix A.

Proposition 1 *If $|\phi_2| < 1$ and $|\phi_1| < 1/\rho(\mathbf{G}(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1})$, then the network game with the utility function (1) has a unique Nash equilibrium in pure strategies with the equilibrium effort vector $\mathbf{y}^* = (y_1^*, ..., y_n^*)'$ given by:*

$$\mathbf{y}^* = \mathbf{y}^*(\mathcal{G}) = (\mathbf{I}_n - \phi_1\mathbf{G} - \phi_2\mathbf{G}^*)^{-1}\boldsymbol{\pi}, \tag{4}$$

[7] A similar utility function has been used by Liu et al. (2014). The focus of that paper is to differentiate *social-multiplier* and *social-conformity* effects, while the focus of the current paper is to conduct the key player analysis in the presence of both effects. We also propose a new estimation method that takes into account the potential endogenous adjacency matrix.

where $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_n)'$ *and* π_i *is defined in Equation (3). Furthermore, if* $\phi_1 \geq 0$*,* $\phi_2 \geq 0$ *and* $\boldsymbol{\pi} \geq \mathbf{0}$*, then* $\mathbf{y}^* \geq \mathbf{0}$*.*

Two special cases of this network game are of particular interest. The first case is when $\lambda_2 = 0$ in the utility (1). In this case, $\phi_2 = 0$ and the best-response function becomes

$$y_i = \phi_1 \textstyle\sum_{j \in \mathcal{N}} g_{ij} y_j + \pi_i,$$

with equilibrium effort levels given by

$$\mathbf{y}^* = (\mathbf{I}_n - \phi_1 \mathbf{G})^{-1} \boldsymbol{\pi}.$$

As the equilibrium effort of an agent depends on the aggregate effort of her connections, we call this case the *local-aggregate* network game. The other case is when $\lambda_1 = 0$ in the utility (1). In this case, $\phi_1 = 0$ and the best-response function becomes

$$y_i = \phi_2 \textstyle\sum_{j \in \mathcal{N}} g_{ij}^* y_j + \pi_i,$$

with equilibrium effort levels given by

$$\mathbf{y}^* = (\mathbf{I}_n - \phi_2 \mathbf{G}^*)^{-1} \boldsymbol{\pi}.$$

As the equilibrium effort of an agent depends on the average effort of her connections, we call this case the *local-average* network game.

3.3 Key player

The *key player* in a network is the agent whose removal from the network leads to the largest reduction in the aggregate effort level (Ballester et al., 2006 and 2010). More formally, let $\mathcal{G}_{-i}$ denote the network obtained by removing the ith node from network $\mathcal{G}$, and $||\cdot||_1$ denote the l_1 matrix norm. Then the key player is defined as

$$i^* = \arg\max_{i \in \mathcal{N}} (||\mathbf{y}^*(\mathcal{G})||_1 - ||\mathbf{y}^*(\mathcal{G}_{-i})||_1) = \arg\min_{i \in \mathcal{N}} ||\mathbf{y}^*(\mathcal{G}_{-i})||_1, \quad (5)$$

where $\mathbf{y}^*(\mathcal{G})$ and $\mathbf{y}^*(\mathcal{G}_{-i})$ are equilibrium effort vectors for the underlying networks $\mathcal{G}$ and $\mathcal{G}_{-i}$, respectively, defined in Equation (4).[8]

For the *local-aggregate* network game (i.e. $\lambda_2 = \phi_2 = 0$), Ballester et al. (2006) and Ballester and Zenou (2014) show that the key player is the agent with the highest *intercentrality* measure in the case without contextual effects (i.e. $\boldsymbol{\gamma}_2 = \boldsymbol{\beta}_2 = \mathbf{0}$) and the agent with the highest *generalized intercentrality* measure in the case with contextual effects (i.e. $\boldsymbol{\gamma}_2 = \boldsymbol{\beta}_2 \neq \mathbf{0}$). However, in presence of the *local-average* endogenous peer effect, it is difficult to identify the key player based on a single network topology measure. The difficulty comes from the fact that the row-normalized adjacency matrix of network $\mathcal{G}_{-i}$ is, in general, not a submatrix of the row-normalized adjacency matrix of network $\mathcal{G}$. Yet, provided that the unknown parameters in the best-response function (2) can be consistently estimated, we can still determine the key player based on its definition in Equation (5).

In the rest of the paper, we present an estimation strategy for Equation (2) with potentially endogenous adjacency matrices, and apply this method to identify key players in juvenile deliquency networks.

4 GMM Estimation with Endogenous Networks

4.1 Econometric model

Let $\bar{r}$ be the total number of networks, n_r be the number of agents in the rth network, and $n = \sum_{r=1}^{\bar{r}} n_r$ be the total number of agents in the sample. The econometric model corresponding to the best-response function (2) can be written as

$$y_{i,r} = \phi_1 \textstyle\sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + \phi_2 \sum_{j=1}^{n_r} g^*_{ij,r} y_{j,r} + \mathbf{x}'_{i,r}\boldsymbol{\beta}_1 + \sum_{j=1}^{n_r} g^*_{ij,r}\mathbf{x}'_{j,r}\boldsymbol{\beta}_2 + \eta_r + u_{i,r}, \tag{6}$$

where $u_{i,r}$'s are i.i.d. with zero mean and variance σ^2. Let $\mathbf{G}_r = [g_{ij,r}]$ and $\mathbf{G}^*_r = [g^*_{ij,r}]$ denote, respectively, the adjacency matrix and the row-normalized adjacency matrix for the rth network. Let $\mathbf{y}_r = (y_{1,r}, \cdots, y_{n_r,r})'$ denote an $n_r \times 1$ vector of observations on the dependent variable and $\mathbf{X}_r = (\mathbf{x}_{1,r}, \cdots, \mathbf{x}_{n_r,r})'$ denote an $n_r \times K$ matrix of observations on K exogenous variables. In

[8] The definition of the key player relies on the assumption that the rest of the network remains unchanged after a player is removed from the network. This assumption can be justified if network links are formed in a pair-wise independence manner. Alternatively, considering that it takes time for the network to evolve in response to the removal of a player, our definition should be qualified as a short-run definition of the key player.

matrix form, Equation (6) can be written as

$$\mathbf{y}_r = \phi_1 \mathbf{G}_r \mathbf{y}_r + \phi_2 \mathbf{G}_r^* \mathbf{y}_r + \mathbf{X}_r \boldsymbol{\beta}_1 + \mathbf{G}_r^* \mathbf{X}_r \boldsymbol{\beta}_2 + \eta_r \boldsymbol{\iota}_{n_r} + \mathbf{u}_r,$$

where $\boldsymbol{\iota}_{n_r}$ is an $n_r \times 1$ vector of ones and $\mathbf{u}_r = (u_{1,r}, \cdots, u_{n_r,r})'$.

For a sample with $\bar{r}$ networks, stack up the data by defining $\mathbf{y} = (\mathbf{y}_1', \cdots, \mathbf{y}_{\bar{r}}')'$, $\mathbf{X} = (\mathbf{X}_1', \cdots, \mathbf{X}_{\bar{r}}')'$, $\mathbf{u} = (\mathbf{u}_1', \cdots, \mathbf{u}_{\bar{r}}')'$, $\mathbf{G} = \mathrm{diag}\{\mathbf{G}_r\}_{r=1}^{\bar{r}}$, $\mathbf{G}^* = \mathrm{diag}\{\mathbf{G}_r^*\}_{r=1}^{\bar{r}}$, $\mathbf{L} = \mathrm{diag}\{\boldsymbol{\iota}_{n_r}\}_{r=1}^{\bar{r}}$ and $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_{\bar{r}})'$, where $\mathrm{diag}\{\mathbf{A}_k\}$ denotes a "generalized" block diagonal matrix in which the diagonal blocks are $m_k \times n_k$ matrices $\mathbf{A}_k$'s. For the entire sample, the model is

$$\mathbf{y} = \phi_1 \mathbf{G}\mathbf{y} + \phi_2 \mathbf{G}^* \mathbf{y} + \mathbf{X}\boldsymbol{\beta}_1 + \mathbf{G}^* \mathbf{X}\boldsymbol{\beta}_2 + \mathbf{L}\boldsymbol{\eta} + \mathbf{u}. \tag{7}$$

In Equation (7), ϕ_1 represents the *local-aggregate endogenous peer effect*, where an agent's effort may depend on the aggregate effort level of her friends; ϕ_2 represents the *local-average endogenous peer effect*, where an agent's effort may depend on the average effort level of her friends; and $\boldsymbol{\beta}_2$ represents *contextual effects*, where an agent's effort may depend on the exogenous characteristics of her friends. The vector of network fixed effects given by $\boldsymbol{\eta}$ captures *the correlated effect* where agents in the same network may behave similarly as they have similar unobserved individual characteristics or they face a similar (institutional) environment. The network fixed effect serves as a remedy for the selection bias that originates from the possible sorting of individuals with similar unobserved characteristics into a network. Since the seminal work of Manski (1993), identification and estimation of social network models in presence of endogenous, contextual and correlated effects has attracted a lot of attention in the literature (see Blume et al., 2011 for an excellent review).

The network fixed effect also captures the deterrence effect on delinquency, i.e., the term $p\,f$ in the utility function (1). Indeed, because networks are within schools in our data, the use of network fixed effects also accounts for differences in the strictness of anti-delinquency regulations across schools. Thus, instead of directly estimating deterrence effects (i.e. coefficients of observable measures of deterrence, such as local police expenditures or the arrest rate in the local area), we focus our attention on the estimation of peer effects in delinquency, accounting for the deterrence effect.

In the econometric model, we allow for network fixed effects to depend on $\mathbf{G}, \mathbf{G}^*$ and $\mathbf{X}$ by treating $\boldsymbol{\eta}$ as a vector of unknown parameters (as in a fixed effect panel data model). When the number of groups $\bar{r}$ is large, we may have the incidental parameter problem. To avoid the incidental parameter problem, we eliminate network fixed effect parameters $\boldsymbol{\eta}$ by pre-multiplying Equation (7) by a projection matrix $\mathbf{J} = \text{diag}\{\mathbf{J}_r\}_{r=1}^{\bar{r}}$, where $\mathbf{J}_r = \mathbf{I}_{n_r} - n_r^{-1}\boldsymbol{\iota}_{n_r}\boldsymbol{\iota}_{n_r}'$. This transformation is analogous to the *within* transformation for the fixed effect panel data model. As $\mathbf{JL} = \mathbf{0}$, the transformed model is

$$\mathbf{Jy} = \phi_1\mathbf{JGy} + \phi_2\mathbf{JG}^*\mathbf{y} + \mathbf{JX}\boldsymbol{\beta}_1 + \mathbf{JG}^*\mathbf{X}\boldsymbol{\beta}_2 + \mathbf{Ju}. \tag{8}$$

4.2 Linear moment conditions and 2SLS

Linear social interaction models can be estimated by the 2SLS method based on linear moment conditions (see, e.g., Bramoullé et al., 2009 and Lin, 2010). When the linear moment condition $\text{E}(\mathbf{u}|\mathbf{G}, \mathbf{G}^*, \mathbf{X}) = \mathbf{0}$ is satisfied, Equation (8) can be estimated based on the empirical linear moment function $\mathbf{Q}'\mathbf{u}(\boldsymbol{\theta})$, where $\mathbf{Q}$ is an IV matrix and

$$\mathbf{u}(\boldsymbol{\theta}) = \mathbf{Jy} - \phi_1\mathbf{JGy} - \phi_2\mathbf{JG}^*\mathbf{y} - \mathbf{JX}\boldsymbol{\beta}_1 - \mathbf{JG}^*\mathbf{X}\boldsymbol{\beta}_2 = \mathbf{J}(\mathbf{y} - \mathbf{Z}\boldsymbol{\theta}), \tag{9}$$

with $\mathbf{Z} = [\mathbf{Gy}, \mathbf{G}^*\mathbf{y}, \mathbf{X}, \mathbf{G}^*\mathbf{X}]$ and $\boldsymbol{\theta} = (\phi_1, \phi_2, \boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$. A possible candidate for the IV matrix is $\mathbf{Q} = \mathbf{J}[\mathbf{X}, \mathbf{G}^*\mathbf{X}, \mathbf{GX}, (\mathbf{G}^*)^2\mathbf{X}, \mathbf{G}\boldsymbol{\iota}_n]$, where $(\mathbf{G}^*)^2\mathbf{X}$ represents the characteristics of friends' friends and $\mathbf{G}\boldsymbol{\iota}_n$ reflects the number of friends of each agent.[9] The model parameters can be identified via the 2SLS method, provided that a certain level of intransitivity exists in the network such that $\text{rank}(\mathbf{Q}) \geq \dim(\boldsymbol{\theta})$ (see, e.g., Bramoullé et al., 2009 and Liu et al., 2014).

However, the validity of the IV method relies on the assumption that the adjacency matrices $\mathbf{G}$ and $\mathbf{G}^*$ are exogenous. If the formation of friendship depends on some unobservable covariates that are correlated with the error term of Equation (8), then $\mathbf{G}$ and $\mathbf{G}^*$ are endogenous and the 2SLS estimator based on the aforementioned IV matrix would be inconsistent.

Suppose the formation of friendship within a network can be (partially) explained by homophily, i.e. similar agents are more likely to be friends than dissimilar ones (McPherson et al., 2001; Currarini

[9] Liu and Lee (2010) and Liu et al. (2014) suggest to use the leading order terms of the Bonacich centrality (Bonacich, 1987), such as $\mathbf{GL}, \mathbf{G}^2\mathbf{L}, \mathbf{G}^3\mathbf{L}, \cdots$, as additional IVs for the local-aggregate endogenous peer effect. Observe that $\mathbf{L}$ has $\bar{r}$ columns, where $\bar{r}$ is the number of networks. As the number of networks increases with the sample size, including $\mathbf{GL}, \mathbf{G}^2\mathbf{L}, \mathbf{G}^3\mathbf{L}, \cdots$ in the IV matrix would induce the many-instrument bias. To avoid this problem, we only include the vector $\mathbf{G}\boldsymbol{\iota}_n$ in the IV matrix at the cost of losing some asymptotic efficiency.

et al., 2009, 2010). Then, the 2SLS method can be remedied by replacing the actual adjacency matrices in the IV matrix by the predicted ones based on some exogenous dyadic characteristics.[10] Suppose the (i,j)th element of the adjacency matrix $\mathbf{G}_r$ is given by:

$$g_{ij,r} = \mathbf{1}(\mathbf{w}'_{ij,r}\boldsymbol{\gamma} + \epsilon_{ij,r} > 0), \tag{10}$$

where $\mathbf{1}(\cdot)$ is an indicator function, $\mathbf{w}_{ij,r}$ is a vector of homophily measures reflecting the similarity in exogenous characteristics between agents i and j, and $\epsilon_{ij,r}$ is an error term. Let $\widehat{\boldsymbol{\gamma}}$ be some estimator of $\boldsymbol{\gamma}$. To fix ideas, suppose $\widehat{\boldsymbol{\gamma}}$ is an estimator obtained from a probit regression of $g_{ij,r}$ on $\mathbf{w}_{ij,r}$. Then, the predicted adjacency matrix is given by $\widehat{\mathbf{G}}_r = [\widehat{g}_{ij,r}]$ with $\widehat{g}_{ij,r} = \Phi(\mathbf{w}'_{ij,r}\widehat{\boldsymbol{\gamma}})$, where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. Let $\widehat{\mathbf{G}}^*_r$ denote the row-normalized predicted adjacency matrix. Furthermore, let $\widehat{\mathbf{G}} = \mathrm{diag}\{\widehat{\mathbf{G}}_r\}_{r=1}^{\bar{r}}$ and $\widehat{\mathbf{G}}^* = \mathrm{diag}\{\widehat{\mathbf{G}}^*_r\}_{r=1}^{\bar{r}}$. Equation (8) can be estimated using the empirical linear moment function

$$\mathbf{g}_1(\boldsymbol{\theta}) = \widehat{\mathbf{Q}}'\mathbf{u}(\boldsymbol{\theta}), \tag{11}$$

where $\widehat{\mathbf{Q}} = \mathbf{J}[\mathbf{X}, \widehat{\mathbf{G}}^*\mathbf{X}, \widehat{\mathbf{G}}\mathbf{X}, (\widehat{\mathbf{G}}^*)^2\mathbf{X}, \widehat{\mathbf{G}}\boldsymbol{\iota}_n]$ and $\mathbf{u}(\boldsymbol{\theta})$ is defined in Equation (9). The 2SLS estimator based on the empirical linear moment function (11) is given by

$$\widehat{\boldsymbol{\theta}}_{2sls} = \arg\min \mathbf{g}_1(\boldsymbol{\theta})'(\widehat{\mathbf{Q}}'\widehat{\mathbf{Q}})^{-1}\mathbf{g}_1(\boldsymbol{\theta}) = (\mathbf{Z}'\widehat{\mathbf{P}}\mathbf{Z})^{-1}\mathbf{Z}'\widehat{\mathbf{P}}\mathbf{y}, \tag{12}$$

where $\widehat{\mathbf{P}} = \widehat{\mathbf{Q}}(\widehat{\mathbf{Q}}'\widehat{\mathbf{Q}})^{-1}\widehat{\mathbf{Q}}'$.

4.3 Quadratic moment conditions and GMM

The IVs based on the predicted adjacency matrices are likely to be weak if the dyadic characteristics are not very informative in explaining the formation of friendship within a network. To alleviate the potential weak IV problem, we introduce some quadratic moment conditions for the estimation of Equation (8) based on the second moment of the error term $\mathbf{u}$.

For a square matrix $\mathbf{A}$, let $\mathbf{A}_{(t)} = \mathbf{JAJ} - (n-\bar{r})^{-1}\mathrm{tr}(\mathbf{JA})\mathbf{J}$. Then, $\mathrm{E}[(\mathbf{Ju})'\widehat{\mathbf{G}}_{(t)}\mathbf{Ju}|\mathbf{X}] = \sigma^2\mathrm{tr}(\widehat{\mathbf{G}}_{(t)}) = 0$ and $\mathrm{E}[(\mathbf{Ju})'\widehat{\mathbf{G}}^*_{(t)}\mathbf{Ju}|\mathbf{X}] = \sigma^2\mathrm{tr}(\widehat{\mathbf{G}}^*_{(t)}) = 0$, which suggests the empirical quadratic

[10] The use of predicted adjacency matrices to construct IVs has also been seen in spatial panel data models (Kelejian and Piras, 2014).

moment function

$$\mathbf{g}_2(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{u}(\boldsymbol{\theta})'\widehat{\mathbf{G}}_{(t)}\mathbf{u}(\boldsymbol{\theta}) \\ \mathbf{u}(\boldsymbol{\theta})'\widehat{\mathbf{G}}^*_{(t)}\mathbf{u}(\boldsymbol{\theta}) \end{bmatrix}, \tag{13}$$

with $\mathbf{u}(\boldsymbol{\theta})$ defined in Equation (9).

Combining the linear and quadratic moment functions defined in Equations (11) and (13), the GMM estimator is given by

$$\widehat{\boldsymbol{\theta}}_{gmm} = \arg\min \mathbf{g}(\boldsymbol{\theta})'\widehat{\boldsymbol{\Omega}}^{-1}\mathbf{g}(\boldsymbol{\theta}), \tag{14}$$

where $\mathbf{g}(\boldsymbol{\theta}) = [\mathbf{g}_1(\boldsymbol{\theta})', \mathbf{g}_2(\boldsymbol{\theta})']'$ and $n^{-1}\widehat{\boldsymbol{\Omega}}$ is a consistent estimator of the variance-covariance matrix of the moment functions $n^{-1}\mathrm{Var}[\mathbf{g}(\boldsymbol{\theta})|\mathbf{X}]$.

The consistency of the proposed 2SLS and GMM estimators does not rely on the consistency of the estimator $\widehat{\boldsymbol{\gamma}}$. Suppose $\widehat{g}_{ij,r} = \Phi(\mathbf{w}'_{ij,r}\widehat{\gamma})$ converges in probability to a well defined limit $\overline{g}_{ij,r}$, for all i, j and r. Then, the row and column sums of the diagonal matrix $\overline{\mathbf{G}} = \mathrm{diag}\{\overline{\mathbf{G}}_r\}_{r=1}^{\bar{r}}$ with $\overline{\mathbf{G}}_r = [\overline{g}_{ij,r}]$ are bounded uniformly in absolute value, if we follow the underlying asymptotic scheme of many works in the social network literature (including Bramoullé et al., 2009) by assuming that network sizes n_r are fixed and bounded, and the number of networks $\bar{r}$ goes to infinity. This would allow us to apply the results in the spatial econometrics literature (e.g., Kelejian and Prucha, 1998 and Lee, 2007) to establish the root-n consistency and asymptotic normality of the proposed 2SLS and GMM estimators.[11] The asymptotic distributions of the proposed estimators are given in Appendix B.

4.4 Monte Carlo simulations

To investigate the finite sample performance of the proposed estimators, we conduct some simulation experiments. In the experiments, the data generating process (DGP) is given by

$$y_{i,r} = \phi_1 \textstyle\sum_{j=1}^{n_r} g_{ij,r}y_{j,r} + \phi_2 \sum_{j=1}^{n_r} g^*_{ij,r}y_{j,r} + \beta_1 x_{i,r} + \beta_2 \sum_{j=1}^{n_r} g^*_{ij,r}x_{i,r} + \eta_r + u_{i,r},$$

and

$$g_{ij,r} = \mathbf{1}(\gamma w_{ij,r} + \epsilon_{ij,r} > 0),$$

[11] If some networks sizes become large as $\bar{r}$ tends to infinity, then the theory on endogenous network links in Qu and Lee (2015) via spatial near epoch dependence can be extended to cover such a situation. The assumption of bounded network sizes is appropriate for our empirical study with AddHealth data.

where $w_{ij,r} = 2 - (x_{i,r} - x_{j,r})^2$ measures the similarity between agents i and j in terms of exogenous characteristics captured by $x_{i,r}$, and $\epsilon_{ij,r} = v_{i,r} + v_{j,r}$, for $r = 1, \cdots, \bar{r}$. We generate $x_{i,r}$ and η_r from independent standard normal distributions. We generate $u_{i,r}$ and $v_{i,r}$ jointly from a bivariate normal distribution

$$\begin{bmatrix} u_{i,r} \\ v_{i,r} \end{bmatrix} \sim \mathrm{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \sigma_{12} \\ \sigma_{12} & 1 \end{bmatrix}\right).$$

We consider $\bar{r} \in \{50, 100, 200\}$ equal-sized groups with the group size $n_r = 10$. In the DGP, we set $\phi_1 = 0.05$, $\phi_2 = 0.2$, $\beta_1 = \beta_2 = 2$, and experiment with different values for γ and σ_{12}.

We consider the following estimators: (a) "2SLS-1", the conventional 2SLS estimator with the IV matrix $\mathbf{Q} = \mathbf{J}[\mathbf{X}, \mathbf{G}^*\mathbf{X}, \mathbf{G}\mathbf{X}, (\mathbf{G}^*)^2\mathbf{X}, \mathbf{G}\boldsymbol{\iota}_n]$; (b) "2SLS-2", the 2SLS estimator defined in Equation (12) with the IV matrix $\widehat{\mathbf{Q}} = \mathbf{J}[\mathbf{X}, \widehat{\mathbf{G}}^*\mathbf{X}, \widehat{\mathbf{G}}\mathbf{X}, (\widehat{\mathbf{G}}^*)^2\mathbf{X}, \widehat{\mathbf{G}}\boldsymbol{\iota}_n]$; and (c) "GMM", the GMM estimator defined in Equation (14). The estimation results are reported in Tables 1 and 2. We use robust measures of central tendency and dispersion, namely, the median, the median of the absolute deviations (Med. AD), the difference between the 0.1 and 0.9 quantile (Dec. Rge) in the empirical distribution of the estimates from 1000 simulation replications. The main observations from the experiment are summarized as follows.

- The conventional 2SLS estimator, namely "2SLS-1", is biased when the adjacency matrix is endogenous. The "2SLS-1" estimates of ϕ_1 and β_2 are upwards biased and the "2SLS-1" estimate of ϕ_2 is downwards biased. As σ_{12} increases and γ decreases, the level of endogeneity captured by the correlation between $g_{ij,r}$ and $u_{i,r}$ increases and, as a result, the magnitude of the bias increases. The bias does not reduce as the sample size increases.

- When $\gamma = 1$, the McFadden's pseudo-R^2 of the probit regression of $g_{ij,r}$ on $w_{ij,r}$ is about 0.418, which suggests $w_{ij,r}$ is informative in explaining friendship formation. In this case, the "2SLS-2" and "GMM" estimates based on the predicted adjacency matrix are unbiased when the number of groups is reasonably large ($\bar{r} \in \{100, 200\}$).

- When $\gamma = 0.2$, the McFadden's pseudo-R^2 of the probit regression of $g_{ij,r}$ on $w_{ij,r}$ is about 0.052, which means $w_{ij,r}$ is less informative in explaining friendship formation. In this case, most "2SLS-2" and "GMM" estimates are still unbiased, except that the estimates of ϕ_2 and β_2 are slightly biased when the number of groups is small ($\bar{r} = 50$). The bias reduces as the sample size increases.

- The "GMM" estimator substantially reduces the dispersion (in terms of Med. AD and Dec. Rge) of the "2SLS-2" estimator. The reduction in dispersion is even more pronounced when $w_{ij,r}$ is less informative in explaining friendship formation (i.e. $\gamma = 0.2$).

[*Insert Tables* 1 *and* 2 *here*]

5 Empirical Analysis

5.1 Data

Our analysis is made possible by the use of the unique information on friendship relationships among teenagers contained in the National Longitudinal Survey of Adolescent Health (Add Health). The survey collects data on the social environment of students in grades 7-12 from a nationally representative sample of roughly 130 private and public schools in the Unites States in the academic year 1994-95. Every student attending the sampled schools on the interview day completed a questionnaire (*in-school survey*) asking to identify the best friends from a school roster (up to five males and five females), and containing questions on basic socio-demographic characteristics.[12] A subset of adolescents selected from the rosters of the sampled schools, was then asked to complete a longer questionnaire containing more sensitive individual and household information (*in-home survey*). This questionnaire was administered to about 20,000 individuals. It contains an extensive set of questions on juvenile delinquency, that are used to construct our dependent variable. Specifically, the Add Health data contains information on 15 delinquency items.[13] The survey asks students how often they participate in each of these activities during the past year.[14] Each response is coded using an ordinal scale ranging from 0 (never participate), 1 (participate 1 or 2 times), 2 (participate 3 or 4 times), up to 3 (participate 5 or more times). To derive quantitative information on crime activity using qualitative answers to a battery of related questions, we calculate an index of delinquency

[12] The limit in the number of nominations is not binding. About 1% of the students nominate ten best friends, about 6% of the students nominate five best female friends, and about 4% of the students nominate five best male friends.

[13] Namely, painting graffiti or signs on someone else's property or in a public place; deliberately damaging property that didn't belong to you; lying to your parents or guardians; taking something from a store without paying for it; getting into a serious physical fight; hurting someone badly enough to need bandages or care from a doctor or nurse; runing away from home; driving a car without its owner's permission; stealing something worth more than $50; going into a house or building to steal something; use or threaten to use a weapon to get something from someone; selling marijuana or other drugs; stealing something worth less than $50; taking part in a fight where a group of your friends was against another group; acting loud, rowdy, or unruly in a public place.

[14] Respondents listened to pre-recorded questions through earphones and then entered their answers directly on laptop computers. This administration of the survey for sensitive topics minimizes the potential for interview and parental influence, while maintaining data security.

involvement for each respondent via a standard factor analysis, where the factor loadings of the different variables are used to calculate the total score. The delinquency index ranges between 0 and 10.27, with a mean of 0.84 and a standard deviation of 1.08.[15] We construct friendship networks assuming friendship to be reciprocal so that the adjacency matrix $\mathbf{G} = [g_{ij}]$ is symmetric, with $g_{ij} = 1$ if either student i or j nominated the other as a friend.

After removing isolated students and pairs (i.e. networks with only two students) as well as students with missing information, the sample contains 9,396 students distributed over 579 networks.[16] The large reduction in sample size with respect to the original sample is common when working with Add Health networks since roughly 20 percent of the students do not nominate any friends and another 20 percent cannot be correctly linked.[17] There are 577 networks with network size range between 3 and 305 students, and 2 networks of roughly 1,000 students. We drop those two networks, since peer effects may be very different in large networks (see, e.g., Calvó Armengol et al. 2009).[18] Our final sample thus consists of 7,735 students distributed over 577 networks, with network sizes ranging from 3 to 305. The median, mean and the standard deviation of network sizes are, respectively, 4, 13.41 and 26.68. On average, the students in our sample nominate roughly 3 friends, with a standard deviation of 2 friends. The list of variables used in the empirical analysis, with their definition and summary statistics, can be found in Table C.1 in Appendix C.

5.2 Estimation results

We consider the following estimators for the general network model given by Equation (8), namely, (a) "2SLS-1", the conventional 2SLS estimator with the IV matrix $\mathbf{Q} = \mathbf{J}[\mathbf{X}, \mathbf{G}^*\mathbf{X}, \mathbf{G}\mathbf{X}, (\mathbf{G}^*)^2\mathbf{X}, \mathbf{G}\boldsymbol{\iota}_n]$; (b) "2SLS-2", the 2SLS estimator defined in Equation (12) with the IV matrix $\widehat{\mathbf{Q}} = \mathbf{J}[\mathbf{X}, \widehat{\mathbf{G}}^*\mathbf{X}, \widehat{\mathbf{G}}\mathbf{X}, (\widehat{\mathbf{G}}^*)^2\mathbf{X}, \widehat{\mathbf{G}}\boldsymbol{\iota}_n]$; and (c) "GMM", the GMM estimator defined in Equation (14).

The estimation results of Equation (8) are reported in Table 3. For the "2SLS-1" estimator, the p-value of the overidentifying restrictions (OIR) test is 0.061, suggesting the "conventional" IVs might be invalid. On the other hand, for the "2SLS-2" and "GMM" estimators, the p-values of the OIR test are very much higher than the conventional significance level, lending us confidence in

[15] The delinquency index is zero if the student reported never participating in any of the listed delinquent activities. In the sample considered, only about 7% of the students have the delinquency index being zero.

[16] A network is defined as the largest set of students who are directly or indirectly connected through friend nominations. By definition, students from two different networks cannot be friends.

[17] The representativeness of the sample is, however, preserved.

[18] The representativeness of the sample is not affected by the trimming procedure (see Table C.1 in Appendix C). Moreover, the results of our analysis, which are obtained when using the entire netwok size range remain qualitatively unchanged. We report them in the last column of Table 5 in Section 5.2 below.

the estimation results. For the local-aggregate peer effect coefficient ϕ_1, the "2SLS-2" and "GMM" yield similar estimates, while the more efficient "GMM" estimator produces a much smaller standard error. Thus, the "GMM" estimate of ϕ_1 is statistically significant, which means that we find evidence of local-aggregate peer effects in juvenile delinquency. The "2SLS-2" and "GMM" estimates of the local-average peer effect coefficient ϕ_2 are statistically insignificant. Furthermore, the "GMM" estimates of ϕ_1 and ϕ_2 satisfy the sufficient condition given in Proposition 1 for the existence of a unique Nash equilibrium of the underlying network game.

[*Insert Table* 3 *here*]

Table 4 reports a probit regression of Equation (10), the auxiliary regression to obtain the predicted adjacency matrices $\widehat{\mathbf{G}}$ and $\widehat{\mathbf{G}}^*$ for the construction of the IV matrix $\widehat{\mathbf{Q}}$. In this regression, we find strong evidence of *homophily.* Students are more likely to be friends, if they are of the same age, gender and race, if they are in the same grade, and if they have similar family background. The McFadden's pseudo-R^2 of the probit regression is 0.08, which is comparable to a case considered in the Monte Carlo simulations. In the Monte Carlo experiment, we find that, when the dyadic characteristics are less informative to predict friendship formation, the "GMM" estimator is more reliable than the "2SLS-2" estimator.

[*Insert Table* 4 *here*]

Table 5 collects the estimation results with alternative sets of regressors. Column 4 of this table corresponds to the last column of Table 3. The "GMM" estimates of ϕ_1 and ϕ_2 do not vary much across columns, showing the robustness against potential model misspecification. We also report the estimation results using the sample with all networks (i.e. including very large networks). The estimates are qualitatively unchanged, showing the result does not depend crucially on the network size threshold used to mitigate possible differences in peer effects across networks of very different sizes.

[*Insert Table* 5 *here*]

Table 6 reports the estimation results for the local-aggregate model (under the restriction $\phi_2 = 0$) and the local-average model (under the restriction $\phi_1 = 0$). As the "GMM" estimate of ϕ_2 is insignificant in the regression with both local-aggregate and local-average peer effects, imposing the restriction $\phi_2 = 0$ in the local-aggregate model has little impact on the estimation result. On the other hand, imposing the restriction $\phi_1 = 0$ in the local-average model biases the estimate of ϕ_2 upwards a little, but the estimated ϕ_2 remains statistically insignificant.

[*Insert Table* 6 *here*]

5.3 Who is the key player?

With the "GMM" estimates of the general network model reported in Table 3, the key player can be determined for each network according to Equation (5).[19] We contrast the key player with two benchmarks: the most criminal teenager in the network and the teenager with the highest values of other standard measures of network centrality (which are not supported by a theoretical analysis). The results are illustrated in Figure 1. First, we find that the key player is not necessarily the most *active* student in delinquent activities. Only in 124 networks out of the 577 networks in the sample, the key player is the most active student in delinquent activities. Second, the key player is not necessarily the student with the highest betweenness centrality, closeness centrality, or eigenvector centrality, especially when the network size is large. In at least 60% of the networks in the sample with more than 50 members, the key player is not the student with the highest betweenness centrality, closeness centrality, or eigenvector centrality.[20]

[*Insert Figure* 1 *here*]

Once we identify the key player for each network, we can draw her "profile" by comparing the characteristics of the key players with those of the other students in the network. Table 7 compares the characteristics of the key player with those of the other students. Surprisingly, we do not observe that the most disruptive kids are those coming from low-income families (as measured by having at least one parent on welfare) and low-educated parents. We do see that, compared to other students,

[19] As the estimate of the local-average effect is statistically insignificant in the empirical study, the key players identified based on the GMM estimates of the general network model coincides with those identified based on the GMM estimates of the local-aggregate network model reported in the first column of Table 6. In 567 networks out of the 577 networks in the sample, the key players identified based on those two models are the same.

[20] Closeness centrality measures the length of the average shortest path passing between a node and each other node. Betweenness is equal to the number of shortest paths from all nodes to each other that passes through that node. Eigenvalue centrality of node i is the ith component of the eigenvector associated to the highest eigenvalue of $\mathbf{G}$. See Jackson (2008) for an introduction and detailed description of these measures.

the key players in crime are more likely to come from single-parent families, but on average their parents have college and above education level, work as professionals, and do not receive public assistance. We also find that key players have more friends than other students, and their friends are more likely to come from single-parent families and tend to be non-white.

[*Insert Table 7 here*]

We then compare the aggregate-delinquency reductions by removing, respectively, the key player and the most active delinquent. In the 577 networks of our sample, on average, the aggregate-delinquency reductions by removing the key player and the most active delinquent are, respectively, 4.51% and 2.44%. A paired two-sample t test suggests that the aggregate-delinquency reduction by removing the key player is statistically significantly larger than that by removing the most active delinquent. The median network size in our sample is 4. For a network with 4 students, the percentage reductions in aggregate delinquency by removing the key player and the most active delinquent are, respectively, 97.23% and 83.14%. Hence, targeting key players is more effective than targeting the most active delinquent in reducing total delinquency.

6 Policy Implications

In this section, we discuss how the key-player policy can be used for fighting crime and for other purposes. First, we list some available data on criminal networks and discuss the extent to which our analysis can be used in practice to address policies against crime. Second, we show that the methodology developed in this paper can be used to design and implement policies in other contexts such as financial networks, R&D networks, networks in developing economies and political networks.

6.1 Data on criminal networks

In order to apply the key-player policy one needs to collect detailed data on the social networks linking individuals and on their criminal activity. We make below a (selective) list of available network data.

Juvenile delinquency in schools There are data similar to the Add Health data for other countries. For example, the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) "School Study" focuses on social networks and the role of peers in delinquency among secondary school students. Survey questionnaires were administered in two waves. Students were provided with a numbered list of all students in the same grade in their school and were asked to fill in the numbers of those fellow students they spend time with at school ("with which of these students do you associate regularly?"), with a maximum of ten possible nominations. Students' delinquent behavior is measured using self-reports on a variety of offenses. The final sample consisted of 1,156 students in 10 schools that participated in both waves (see Weerman, 2011).

Adult crime (police data) The police has a lot of information on criminal networks. For example, each time two (or more) persons are suspected of a crime, the police in Sweden registers this information. A link in a network could then be created by linking individuals who are suspected of a crime together. Lindquist and Zenou (2014) map the network of all criminals in Sweden for several years. This type of information can be obtained from the police in many other countries. Analyzing these networks can help law enforcement agencies in designing more effective strategies for crime prevention and reduction (Tayebi et al., 2011). Also in the United States there are similar data. For example, Coplink (Hauck et al., 2002) was one of the first large scale research projects in crime data mining, and an excellent work in criminal network analysis. It is remarkable in its practicality, being integrated with and used in the workflow of the Tucson Police Department. Coplink has information about the perpetrators' habits and close associations in crime to capture the *connections* between people, places, events, and vehicles, based on past crimes. Xu and Chen (2005) built on this when they created CrimeNet Explorer, a framework for criminal network knowledge discovery incorporating hierarchical clustering, social network analysis methods, and multidimensional scaling.

Gang networks McGloin (2004, 2005) uses data from the Newark portion of the North Jersey Gang Task Force, a regional problem analysis project that sought to define the local gang landscape in Northern New Jersey. These data came from the experiential knowledge of representatives of various criminal justice agencies, including the Newark Police Department, Essex County Sheriff's Office, Essex County Department of Parole, and Juvenile Justice Commission of New Jersey. In particular, groups of law enforcement officials from this variety of criminal justice agencies engaged in collective semi-structured interviews — 32 over the course of one year — that solicited information on the gang landscape. They provided information on known gang members, as well as the quantity and type of their respective associates. The classification of the profession of gang members in the questionnaire relied on New Jersey code. This code defines a gang as three or more people who are associated, that is people who have a common group name or a common identifying sign, such as tattoos, or other indices of association, and who have committed criminal offenses while engaged in gang-related activity.[21]

6.2 Implementation and relevance of the key-player policy in crime

Different policies can be implemented to reduce crime based on key-player analysis.

First, the police can offer to the key player(s) incentives to leave the gang or the criminal network. For example, the police can offer them a job or a conditional transfer (by asking them to move to another city) or monitor them more. These types of policies have been implemented in Canada, where some gang members of criminal networks were persuaded to abandon gang life in return for needed employment training, educational training, and skills training (Tremblay et al., 1996).

Second, key-player analysis can be useful in implementing the "focused deterrence" strategy (also referred to as "pulling levers" policing), which is a recent innovation in policing based on the growing evidence of the effectiveness of police deterrence strategies (Kennedy, 1998, 2008). This strategy was pioneered in Boston (*Boston Gun Project*) with the intent of understanding the purported nexus of rising youth violence and use of firearms. As part of its analysis, representatives of various criminal justice agencies defined and characterized problematic local gangs (Braga et al., 2001; Kennedy et al., 1996, 1997, 2001). This process included elaborating on the relationships among the street gangs, which Kennedy et al. (1996, 1997, 2001) translated into sociograms illustrating connections within the gang landscape. This seemingly simple information was invaluable for the problem analysis and construction of *Operation Ceasefire*. This latter policy combines a strong law enforcement response with a "pulling levers" deterrence effort aimed at chronic gang offenders. The key to the success is to use a "lever pulling" approach, which is a crime deterrence strategy that attempts to prevent violent behavior by using a *targeted individual or group's vulnerability* to law enforcement as a means of gaining their compliance. Operation Ceasefire was first launched in Boston and youth homicide fell by two-thirds after the Ceasefire strategy was put in place in 1996 (Kennedy, 1998). It was then implemented in Los Angeles in 2000: police beefed up patrols in the area, attempting to locate gang members who had *outstanding arrest warrants* or *had violated probation or parole regulations*. Gang members who had *violated public housing rules*, *failed to pay child support*, or were similarly vulnerable were also subjected to stringent enforcement (Tita et al., 2003).

Finally, key-player analysis can also be helpful for other related issues. For example, there is a

[21] See also Mastrobuoni and Patacchini (2012) who use a data set provided by the Federal Bureau of Narcotics on criminal profiles of 800 US Mafia members active in the 1950s and 1960s and on their connections within the Cosa Nostra network.

lot of debate in the U.S. on how to allocate under-age offenders into juvenile detention centers. To reduce peer influence and criminal learning behind bars, our model suggests key players should be locked away from other young delinquents.

6.3 Key players in other contexts

An important advantage of our methodology is that it can be applied to other contexts where network data are available. Let us provide some examples.

Financial networks There is an abundance of information available on financial networks. Networks have proved to be a useful analytical tool for studying financial contagion and systemic risk from both theoretical and empirical perspectives.[22] Billio et al. (2012) propose several measures of interconnectedness among the monthly returns of hedge funds, banks, brokers, and insurance companies, and show their predictive power for market dislocation. Using unique data set of transactions from two financial futures contracts traded on the Chicago Mercantile Exchange (CME), Cohen-Cole et. al. (2014) show that network pattern of trades captures the relations between behavior in the market and returns. Cohen-Cole et al. (2015) test a model of network interactions using transaction level data on interbank lending from an electronic interbank market, the e-MID SPA (or e-MID), which was the reference marketplace for liquidity trading in the Euro area from January 2002 to December 2009. In the context of financial networks, key-player analysis can be used to identify the bank that will need to be bailed out in a financial crisis in order to reduce systemic risk or minimize total loss.[23] This is an extremely important issue because the recent financial crisis has shown that a bank is not necessary "too big to fail" but "too interconnected to fail".

R&D networks There is also a lot of information on R&D networks. For example, García-Canal et al. (2008) use alliance data stemming from the Thomson Securities Data Company (SDC) Platinum data base. Three types of alliances are reported in the SDC database: (1) alliances that imply the transmission of an existing technology from one partner to another or to the alliance; (2) alliances that imply the cross-transfer of existing technologies between two or more partners or between these and the alliance, and (3) alliances that include the undertaking of R&D activities. König et al. (2014) merge the SDC alliance database with the Cooperative Agreements and Technology Indicators (CATI) database. The CATI database only records agreements for which a combined innovative activity or an exchange of technology is at least part of the agreement. In the context

[22] See Babus and Allen (2009) and Cohen-Cole et. al. (2013) for recent reviews of this literature.
[23] See Denbee et al. (2017) for a first attempt to identify "key" banks in financial networks.

of R&D networks, key-player analysis can be used, for example, to design optimal R&D subsidy policies (see König et al., 2014).

Networks in developing economies There are many network data for developing countries.[24] Banerjee et al. (2013) identify key players in the context of diffusion using eigenvector centrality. Using network data from a survey on 75 rural villages in India, the authors look at the diffusion of a microfinance program in these villages and show that, if the bank in charge of this program had targeted individuals in the village with the highest eigenvector centrality, the *diffusion* of the microfinance program (i.e. take-up rates) would have been much higher.[25] More generally, in developing countries, one could apply the key player policy to the issue of adoption of a new technology since there is strong evidence of social learning (Conley and Udry, 2010) and take-up rates in microfinance programs.[26] Another application of the key player methodology is the administration of vaccines. Targeting individuals who would reduce the spread of infectious diseases the most would be highly cost effective, freeing up resources to make more vaccination projects viable, and reducing the overall infection rate.

Political networks The key player policy could also be applied in the political arena. Recent advances in data collection have fostered a renaissance of interest in the study of the importance of social connections in politics among economists (Lazer, 2011, Battaglini and Patacchini 2018b). Cohen and Malloy (2014) provide evidence that personal connections have a significant impact on the voting behavior of U.S. politicians. Battaglini and Patacchini (2018a) show that interest groups allocate more campaign contributions to legislators with higher Katz-Bonacich network centralities. Using a field experiment to study social influences on political participation during the 2012 U.S. presidential election campaign, Perez-Truglia and Cruces (2017) document that making an individual's contributions more visible to her neighbors increases the contributions of supporters of the local majority party and decreases those of supporters of the minority party.[27] Competition to persuade public opinion does not only occur in marketing between rival firms, but also in lobbying by interests

[24] See Munshi (2014), Chuang and Schechter (2015) and Breza (2016) for recent overviews of the literature on networks and development.

[25] Other approaches have been implemented to identify key players in the context of diffusion. For example, in *viral marketing*, firms try to use word-of-mouth effects to market a product with a limited advertising budget, relying on the fact that early adopters may convince friends and colleagues to use the product, creating a large wave of adoptions. For overviews, see Kleinberg (2007) and Mayzlin (2016).

[26] See also the recent paper by Fafchamps and Labonne (2017) who, using data on 3.6 million households from the Philippines, show that households with high *betweenness centrality* receive more public services from their local government.

[27] See Victor et al. (2018) for a review of studies on the importance of networks in political science.

groups on opposite sides of a legislation and it is the essence of political campaigns. Our key-player policy would suggest that resources should be spent on *key politicians or key voters* who have an influential position in the social network. Our methodology may provide a cost effective instrument for electoral campaigns. For example, consider the Obama-Romney 2012 U.S. presidential election, which recorded the most extensive campaign spending in U.S. history.[28] Implementation of our key-player methodology would shift attention away from "swing states" and rather target "swing voters", who would have the greatest possible impact on voter decision within their social networks. Such an approach could easily have tempered the USD 2.3 billion total campaign cost.[29] The cost of gathering the necessary data could not be expected to be more than a fraction of this.

7 Concluding remarks

This paper presents a methodology for determining the key player. The key-player methodology has great scope for practical implementation, since it takes advantage of multiplier effects in naturally occurring networks. Given the availability of network data, it can be implemented in many contexts. Even where data is not yet readily available, implementing the key-player policy could be cost-effective overall, if the cost of data gathering is recouped by the cost saving in the actual policy intervention that stems from targeting key-players rather than all individuals.

References

[1] Advani, A. and B. Malde (2018), "Credibly identifying social effects: Accounting for network formation and measurement error," *Journal of Economic Surveys*, forthcoming.

[2] Akers, R.L. (1998), *Social Learning and Social Structure: A General Theory of Crime and Deviance*, Boston: Northeastern University Press.

[3] Babus, A., and F. Allen (2009): "*Networks in finance*," In: P. Kleindorfer and J. Wind (Eds.), *Network-based Strategies and Competencies*, New Jersey: Pearson Education.

[4] Badev, A. (2017), "Discrete games in endogenous networks: Theory and policy," Unpublished manucript, University of Pennsylvania.

[28] Indeed, in the 2016 U.S. presidential election, both Hillary Clinton and Donald Trump raised less than President Obama and Mitt Romney did in 2012. See: https://www.washingtonpost.com/graphics/politics/2016-election/campaign-finance/.

[29] See http://www.opensecrets.org/pres12/index.php#out.

[5] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2006), "Who's who in networks. Wanted: the key player," *Econometrica* 74, 1403-1417.

[6] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2010), "Delinquent networks," *Journal of the European Economic Association* 8, 34-61.

[7] Ballester, C. and Y. Zenou (2014), "Key player policies when contextual effects matter," *Journal of Mathematical Sociology* 38, 233-248.

[8] Banerjee, A., Chandrasekhar, A.G., Duflo, E. and M.O. Jackson (2013), "The diffusion of microfinance," *Science* 26, Vol. 341, no. 6144.

[9] Battaglini, M. and E. Patacchini (2018a), "Influencing connected legislators," *Journal of Political Economy*, forthcoming.

[10] Battaglini, M. and E. Patacchini (2018b), "Social networks and policy making," Unpublished manuscript, Cornell University.

[11] Bayer, P., Hjalmarsson, R. and D. Pozen (2009), "Building criminal capital behind bars: Peer effects in juvenile corrections," *Quarterly Journal of Economics* 124, 105-147.

[12] Becker, G.S. (1968), "Crime and punishment: An economic approach," *Journal of Political Economy* 76, 169-217.

[13] Billio, M., M. Getmanskyb, A. W. Lo, and L. Pelizzona (2012), "Econometric measures of connectedness and systemic risk in the finance and insurance sectors," *Journal of Financial Economics* 104, 535-559.

[14] Blume, L.E., Brock, W.A., Durlauf, S.N. and Y.M. Ioannides (2011), "Identification of social interactions," In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics*, Amsterdam: Elsevier Science, pp. 853-964.

[15] Borgatti, S.P. (2003), "The key player problem," In: R. Breiger, K. Carley and P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, New York: National Academy of Sciences Press, pp. 241-252.

[16] Borgatti, S.P. (2006), "Identifying sets of key players in a network," *Computational, Mathematical and Organizational Theory* 12, 21-34.

[17] Boucher, V., Bramoullé, Y., Djebbari, H. and B. Fortin (2014), "Do peers affect student achievement? Evidence from Canada using group size variation," *Journal of Applied Econometrics* 29, 91-109.

[18] Braga, A.A., Kennedy, D.M., Waring, E.J. and A.M. Piehl (2001), "Problem oriented policing, deterrence, and youth violence: An evaluation of Boston's Operation Ceasefire," *Journal of Research in Crime and Delinquency* 38, 195-225.

[19] Bramoullé, Y., Djebbari, H. and B. Fortin (2009), "Identification of peer effects through social networks," *Journal of Econometrics* 150, 41-55.

[20] Bramoullé, Y. and R.E. Kranton (2007), "Public goods in networks," *Journal of Economic Theory* 135, 478-494.

[21] Bramoullé, Y., Kranton, R.E. and M. d'Amours (2014), "Strategic interaction and networks," *American Economic Review* 104, 898-930.

[22] Breza, E. (2016), "Field experiments, social networks, and development," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press.

[23] Calvó-Armengol, A., Patacchini, E. and Y. Zenou (2009), "Peer effects and social networks in education," *Review of Economic Studies* 76, 1239-1267.

[24] Calvó-Armengol, A. and Y. Zenou (2004), "Social networks and crime decisions: The role of social structure in facilitating delinquent behavior," *International Economic Review* 45, 935-954.

[25] Chuang, Y. and L. Schechter (2015), "Social networks in developing countries," *Annual Review of Resource Economics* 7, 451-472.

[26] Cohen-Cole E., Kirilenko A. and E. Patacchini (2013), "Strategic interactions on financial networks for the analysis of systemic risk," In: J.-P. Fouque and J.A. Langsam (Eds.), *Handbook on Systemic Risk*, Cambridge: Cambridge University Press.

[27] Cohen-Cole E., Kirilenko A. and E. Patacchini (2014), "Trading networks and liquidity provision," *Journal of Financial Economics* 113, 235-251.

[28] Cohen-Cole, E., Patacchini, E. and Y. Zenou (2015), "Static and dynamic networks in nterbank markets," *Network Science* 3, 98–123.

[29] Cohen, L. and C. Malloy (2014), "Friends in high places," *American Economic Journal: Economic Policy* 6, 63-91.

[30] Conley, T.J. and C.R. Udry (2010), "Learning about a new technology: Pineapple in Ghana," *American Economic Review* 100, 35-69.

[31] Cook, P.J. and J. Ludwig (2010), "Economical crime control," NBER Working Paper No. 16513.

[32] Corno, L. (2017), "Homelessness and crime: Do your friends matter?" *Economic Journal* 127, 959-995.

[33] Currarini, S., Jackson, M.O., and P. Pin (2009), "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica* 77, 1003-1045.

[34] Currarini, S., Jackson, M.O., and P. Pin (2010), "Identifying the roles of choice and chance in network formation: Racial biases in high school friendships," *Proceedings of the National Academic of Sciences of the USA* 107, 4857-4861.

[35] Damm, A.P. and C. Dustmann (2014), "Does growing up in a high crime neighborhood affect youth criminal behavior?" *American Economic Review* 104, 1806-1832.

[36] De Paula, A. (2017), "Econometrics of network models," In: B. Honore, A. Pakes, M. Piazzesi and L. Samuelson (Eds.), *Advances in Economics and Econometrics: Theory and Applications: Eleventh World Congress*, Cambridge: Cambridge University Press, pp. 268-323.

[37] Denbee, E., Julliard, C., Li, Y. and K. Yuan (2017), "Network risk and key players: A structural analysis of interbank liquidity," Unpublished manuscript, London School of Economics.

[38] Dolton, P. (2017), "Identifying social network effects," *Economic Record* 93, 1-15.

[39] Durlauf, S.N. and Y.M. Ioannides (2010), "Social interactions," *Annual Review of Economics* 2, 451-478.

[40] Eck, J.E., Chainey, S., Cameron, J.G., Leitner, M. and R.E.Wilson (2005), *Mapping Crime: Understanding Hot Spots*, Washington, DC: Office of Justice Programs, U.S. Department of Justice.

[41] Fafchamps, M. and J. Labonne (2017), "Family networks and distributive politics," Unpublished manuscript, Stanford University.

[42] Galeotti A., Goyal S., Jackson M.O., Vega-Redondo F. and L. Yariv (2009), "Network games," *Review of Economic Studies* 77, 218-244.

[43] García-Canal, E., Valdés-Llaneza, A. and P. Sánchez-Lorda (2008), "Technological flows and choice of joint ventures in technology alliances," *Research Policy* 37, 97-114.

[44] Glaeser, E.L., Sacerdote, B. and J. Scheinkman (1996), "Crime and social interactions," *Quarterly Journal of Economics* 111, 508-548.

[45] Glaeser, E.L. and J. Scheinkman (2003), "Non-market interactions", In: M. Dewatripont, L. P. Hansen and S. Turnovsky (Eds.), *Advances in Economics and Econometrics: Theory and Applications*, Cambridge: Cambridge University Press.

[46] Graham, B.S. (2015), "Methods of identification in social networks," *Annual Review of Economics* 7, 465-485.

[47] Graham, B.S. (2017), "An econometric model of network formation with degree heterogeneity," *Econometrica* 85, 1033-1063.

[48] Graham, B.S. and A. de Paula (2018), *The Econometric Analysis of Network Data*, New York: Academic Press, forthcoming.

[49] Hall, B.H., Jaffe, A.B. and M. Trajtenberg (2001), "The NBER patent citations data file: Lessons, insights and methodological tools," NBER Working Paper No. 8498.

[50] Hansen, C., Hausman, J. and W. Newey (2008), "Estimation with many instrumental variables," *Journal of Business and Economic Statistics* 26, 398-422.

[51] Hauck, R.V., Atabakhsh, H., Ongvasith, P., Gupta, H. and H. Chen (2002), "Using Coplink to analyze criminal-justice data," *IEEE Computer* 35, 3: 3037.

[52] Haynie, D.L. (2001), "Delinquent peers revisited: Does network structure matter?" *American Journal of Sociology* 106, 1013-1057.

[53] Jackson, M.O. (2008), *Social and Economic Networks* , Princeton University Press.

[54] Jackson, M. O., B. W. Rogers, and Y. Zenou (2017), "The impact of social networks on economic behavior," *Journal of Economic Literature* 55, 49-95.

[55] Jackson, M.O. and Y. Zenou (2015), "Games on networks," In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier, pp. 91-157.

[56] Kelejian, H.H and G. Piras (2014), "Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes," *Regional Science and Urban Economics* 46, 140-149.

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

[58] Kennedy, D.M. (1998), "Pulling levers: Getting deterrence right," *National Institute of Justice Journal* 236, 2-8.

[59] Kennedy, D.M. (2008), *Deterrence and Crime Prevention: Reconsidering the Prospect of Sanction*, New York: Routledge.

[60] Kennedy, D.M., Braga, A.A. and A.M. Piehl (1996), *Youth Gun Violence in Boston: Gun Markets, Serious Youth Offenders, and a Use Reduction Strategy*, Boston, MA: John F. Kennedy School of Government, Harvard University Press.

[61] Kennedy, D.M., Braga, A.A. and A.M. Piehl (1997), "The (un)known universe: Mapping gangs and gang violence in Boston," In: D. Weisburd and T. McEwen (Eds.), *Crime Mapping and Crime Prevention*, Monsey, N.Y: Criminal Justice Press.

[62] Kennedy, D.M., Braga, A.A., Piehl, A.M. and E.J. Waring (2001), *Reducing Gun Violence: The Boston Gun Project's Operation Ceasefire*, Washington, D.C: U.S. National Institute of Justice.

[63] Kleiman, M.A. (2009), *When Brute Force Fails. How to Have Less Crime and Less Punishment*, Princeton: Princeton University Press.

[64] Kleinberg, J. (2007), "Cascading behavior in networks: Algorithmic and economic issues," In: N. Nisan, T. Roughgarden, É. Tardos and V. Vazirani (Eds.), *Algorithmic Game Theory*, Cambridge: Cambridge University Press, pp. 613-632.

[65] Kling, J.R., Ludwig, J. and L.F. Katz (2005), "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment," *Quarterly Journal of Economics* 120, 87-130.

[66] König, M., Liu, X. and Y. Zenou (2014), "R&D networks: Theory, empirics and policy implications", CEPR Discussion Paper No. 9872.

[67] König, M., Rohner, D., Thoenig, M. and F. Zilibotti (2017), "Networks in conflict: Theory and evidence from the great war of Africa," *Econometrica* 85, 1093-1132.

[68] Lazer, D. (2011), "Networks in political science: back to the future," *PS: Political Science and Politics* 44, 61-68.

[69] Lee, L.F. (2007), "GMM and 2SLS estimation of mixed regressive, spatial autoregressive models," *Journal of Econometrics* 137, 489-514.

[70] Lin, X. (2010), "Identifying peer effects in student academic achievement by a spatial autoregressive model with group unobservables," *Journal of Labor Economics* 28, 825-860.

[71] Lindquist, M. and Y. Zenou (2014), "Key players in co-offending networks," CEPR Discussion Paper No. 9889.

[72] Liu, X. and L.F. Lee (2010), "GMM estimation of social interaction models with centrality," *Journal of Econometrics* 159, 99-115.

[73] Liu, X., Patacchini, E., Zenou, Y. and L-F. Lee (2012), "Criminal networks: Who is the key player?" CEPR Discussion Paper No. 8772.

[74] Liu, X., Patacchini, E. and Y. Zenou (2014), "Endogenous peer effects: Local aggregate or local average?" *Journal of Economic Behavior and Organization* 103, 39-59.

[75] Ludwig, J., Duncan, G.J. and Hirschfield, P. (2001), "Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment," *Quarterly Journal of Economics* 116, 655-679.

[76] Manski, C.F. (1993), "Identification of endogenous effects: The reflection problem," *Review of Economic Studies* 60, 531-542.

[77] Mastrobuoni, G. and E. Patacchini (2012), "Organized crime networks: An application of network analysis techniques to the American mafia," *Review of Network Economics* 11(3), Article 10.

[78] Mayzlin, D. (2016), "Marketing and networks," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press.

[79] McPherson, M., Smith-Lovin, L. and J. M. Cook (2001)," Birds of a feather: Homophily in social networks," *Annual Review of Sociology* 27, 415-444.

[80] McGloin, J.M. (2004), *Associations among Criminal Gang Members as a Defining Factor of Organization and as a Predictor of Criminal Behavior: The Gang Landscape of Newark, New Jersey*, Ann Arbor, MI: University of Michigan Press.

[81] McGloin, J.M. (2005), "Policy and intervention considerations of a network analysis of street gangs," *Criminology and Public Policy* 4, 607-636.

[82] Mele, A. (2017), "A structural model of dense network formation," *Econometrica* 85, 825-850.

[83] Munshi K. (2014), "Community networks and the process of development," *Journal of Economic Perspectives* 28, 49-76.

[84] Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economic, and Organization* 28, 1-31.

[85] Perez-Truglia, R. and Cruces, G. (2017), "Partisan interactions: Evidence from a field experiment in the United States," *Journal of Political Economy* 125, 1208-1243.

[86] Qu, X. and L.F. Lee (2015), "Estimating a spatial autoregressive model with an endogenous spatial weight matrix," *Journal of Econometrics* 184, 209-232.

[87] Sarnecki, J. (2001), *Delinquent Networks: Youth Co-Offending in Stockholm*, Cambridge: Cambridge University Press.

[88] Sheng, S. (2016), "A structural econometric analysis of network formation games," Working Paper, UCLA.

[89] Stock, J.H, Wright J.H. and M. Yogo (2002), "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business and Economic Statistics* 2, 518-529.

[90] Stock, J.H and M. Yogo (2005), "Testing for weak instruments in linear IV regression," In: D.W.K. Andrews and J.H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge: Cambridge University Press, pp. 80-108.

[91] Sutherland, E.H. (1947), *Principles of Criminology*, fourth edition, Chicago: J.B. Lippincott.

[92] Tayebi, M.A., Bakker, L., Glässer, U. and V. Dabbaghian (2011), "Organized crime structures in co-offending networks," *IEEE Ninth International Conference on Autonomic and Secure Computing* (DASC), pp. 846-853.

[93] Tita, G.K. Riley, J., Ridgeway, G., Grammich, C., Abrahamse, A. and P.W. Greenwood (2003), *Reducing Gun Violence: Results from an Intervention in East Los Angeles*, Santa Monica, CA: RAND Corporation.

[94] Thornberry, T.P., Krohn, M.D., Lizotte, A.J., Smith, C.A. and K. Tobin (2003), *Gangs and Delinquency in Developmental Perspective*, Cambridge: Cambridge University Press.

[95] Tremblay, R.E., Masse, L., Pagani, L. and F. Vitaro (1996), "From childhood physical aggression to adolescent maladjustment: The Montreal Prevention Experiment," In R.D. Peters and R.J. McMahon (Eds.), *Preventing Childhood Disorders, Substance Abuse, and Delinquency*, Thousand Oaks, CA: Sage Publications.

[96] United States Treasury Department (2007), *MAFIA: The Government's Secret File on Organized Crime*, Bureau of Narcotics, New York: Harper Collins Publishers.

[97] Victor, J. N., Montgomery A. H. and M. Lubell (2018), *The Oxford Handbook of Political Networks*, Oxford University Press.

[98] Verdier, T. and Y. Zenou (2004), “Racial beliefs, location and the causes of crime,” *International Economic Review* 45, 727-756.

[99] Warr, M. (2002), *Companions in Crime: The Social Aspects of Criminal Conduct*, Cambridge: Cambridge University Press.

[100] Wasserman, S. and K. Faust (1994), *Social Network Analysis. Methods and Applications*, Cambridge: Cambridge University Press.

[101] Weerman, F.M. (2011), “Delinquent peers in context: A longitudinal network analysis of selection and influence effects,” *Criminology* 49, 253-286.

[102] Xu, J.J. and H. Chen (2005), “CrimeNet Explorer: A framework for criminal network knowledge discovery,” *ACM Transactions on Information Systems* 23, 201-226.

[103] Zenou, Y. (2016), “Key players,” In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press.

APPENDIX

A Proofs

Proof of Proposition 1. If $|\phi_2| < 1$, then $\mathbf{I}_n - \phi_2\mathbf{G}^*$ is invertible and its inverse is a Neumann series

$$(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1} = \textstyle\sum_{k=0}^{\infty} \phi_2^k \mathbf{G}^{*k}. \tag{15}$$

If, in addition, $|\phi_1| < 1/\rho(\mathbf{G}(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1})$, then $\mathbf{I}_n - \phi_1\mathbf{G}(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1}$ is invertible and its inverse is a Neumann series

$$[\mathbf{I}_n - \phi_1\mathbf{G}(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1}]^{-1} = \textstyle\sum_{k=0}^{\infty} \phi_1^k [\mathbf{G}(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1}]^k. \tag{16}$$

The invertibility of $\mathbf{I}_n - \phi_2\mathbf{G}^*$ and $\mathbf{I}_n - \phi_1\mathbf{G}(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1}$ implies that $\mathbf{I}_n - \phi_1\mathbf{G} - \phi_2\mathbf{G}^*$ is invertible. Furthermore, if $\phi_1 \geq 0$ and $\phi_2 \geq 0$, then it follows by Equations (15) and (16) that the elements of $(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1}$ and $[\mathbf{I}_n - \phi_1\mathbf{G}(\mathbf{I}_n - \phi_2\mathbf{G}^*)^{-1}]^{-1}$ are nonnegative, which, in turn, implies that the elements of $(\mathbf{I}_n - \phi_1\mathbf{G} - \phi_2\mathbf{G}^*)^{-1}$ are nonnegative. If, in addition, $\boldsymbol{\pi} \geq \mathbf{0}$, then $\mathbf{y}^* \geq \mathbf{0}$ according to Equation (4). ■

B Asymptotic Distributions of 2SLS and GMM

Assume $\widehat{g}_{ij,r}$ converges in probability to a well defined limit $\overline{g}_{ij,r}$, such that $|\overline{g}_{ij,r}| \leq 1$ for all i, j and r, and the row and column sums of the diagonal matrix $\overline{\mathbf{G}} = \text{diag}\{\overline{\mathbf{G}}_r\}_{r=1}^{\bar{r}}$ with $\overline{\mathbf{G}}_r = [\overline{g}_{ij,r}]$ are uniformly bounded in absolute value. In addition to the setting that sizes of all networks are finite and bounded while the number of networks tends to infinity, assume sample observations are independent across newtorks. Then, it follows by similar arguments as in Kelejian and Prucha (1998) that

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_{2sls} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \sigma^2(\text{plim}_{n\to\infty} n^{-1}\mathbf{Z}'\widehat{\mathbf{P}}\mathbf{Z})^{-1}).$$

Furthermore, let

$$\boldsymbol{\Omega} = \text{Var}[\mathbf{g}(\boldsymbol{\theta})|\mathbf{X}] = \begin{bmatrix} \sigma^2\widehat{\mathbf{Q}}'\widehat{\mathbf{Q}} & \mu_3\widehat{\mathbf{Q}}'\boldsymbol{\omega} \\ \mu_3\boldsymbol{\omega}'\widehat{\mathbf{Q}} & (\mu_4 - 3\sigma^4)\boldsymbol{\omega}'\boldsymbol{\omega} + \sigma^4\boldsymbol{\Delta} \end{bmatrix}$$

and

$$\mathbf{D} = -\mathrm{E}[\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}|\mathbf{X}] = \begin{bmatrix} \widehat{\mathbf{Q}}'\mathbf{G}\mathrm{E}(\mathbf{y}|\mathbf{X}) & \widehat{\mathbf{Q}}'\mathbf{G}^*\mathrm{E}(\mathbf{y}|\mathbf{X}) & \widehat{\mathbf{Q}}'\mathbf{X} & \widehat{\mathbf{Q}}'\mathbf{G}^*\mathbf{X} \\ \sigma^2\mathrm{tr}((\widehat{\mathbf{G}}_{(t)} + \widehat{\mathbf{G}}'_{(t)})\mathbf{G}\mathbf{S}^{-1}) & \sigma^2\mathrm{tr}((\widehat{\mathbf{G}}_{(t)} + \widehat{\mathbf{G}}'_{(t)})\mathbf{G}^*\mathbf{S}^{-1}) & \mathbf{0} & \mathbf{0} \\ \sigma^2\mathrm{tr}((\widehat{\mathbf{G}}^*_{(t)} + \widehat{\mathbf{G}}^{*\prime}_{(t)})\mathbf{G}\mathbf{S}^{-1}) & \sigma^2\mathrm{tr}((\widehat{\mathbf{G}}^*_{(t)} + \widehat{\mathbf{G}}^{*\prime}_{(t)})\mathbf{G}^*\mathbf{S}^{-1}) & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

where $\boldsymbol{\omega} = [\mathrm{vec}_D(\widehat{\mathbf{G}}_{(t)}), \mathrm{vec}_D(\widehat{\mathbf{G}}^*_{(t)})]$, $\boldsymbol{\Delta} = \frac{1}{2}[\mathrm{vec}(\widehat{\mathbf{G}}_{(t)}+\widehat{\mathbf{G}}'_{(t)}), \mathrm{vec}(\widehat{\mathbf{G}}^*_{(t)}+\widehat{\mathbf{G}}^{*\prime}_{(t)})]'[\mathrm{vec}(\widehat{\mathbf{G}}_{(t)}+\widehat{\mathbf{G}}'_{(t)}), \mathrm{vec}(\widehat{\mathbf{G}}^*_{(t)}+\widehat{\mathbf{G}}^{*\prime}_{(t)})]$, and $\mathbf{S} = \mathbf{I}_n - \phi_1\mathbf{G} - \phi_2\mathbf{G}^*$.[30] It follows by similar arguments as in Lee (2007) that

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_{gmm} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, (\mathrm{plim}_{n\to\infty} n^{-1}\mathbf{D}'\boldsymbol{\Omega}^{-1}\mathbf{D})^{-1}).$$

C Data Summary

Table C1 provides the data description and summary statistics.

[Insert Table C1 here]

[30] For a square matrix $\mathbf{A} = [a_{ij}]$, let $\mathrm{vec}_D(\mathbf{A}) = (a_{11}, \cdots, a_{nn})'$ denote the vector containing the diagonal elements of $\mathbf{A}$, and $\mathrm{vec}(\mathbf{A}) = (a_{11}, \cdots, a_{n1}, a_{12}, \cdots, a_{nn})'$ denote the vectorization of $\mathbf{A}$.

TABLE C.1: DATA DESCRIPTION

		All Networks ($n = 9{,}396$)		Networks of Sizes 3~305 ($n = 7{,}735$)	
Variable	*Definition*	*Mean*	*SD*	*Mean*	*SD*
Delinquency index	In the text	0.84	1.08	0.83	1.08
Age	Age	15.63	1.68	15.43	1.71
Grade	7th-12th grades are coded as 1,…,6	3.70	1.62	3.49	1.65
Female	1 if female	0.53	0.50	0.54	0.50
(White American)	1 if White American	0.52	0.50	0.55	0.50
African American	1 if African American	0.21	0.40	0.22	0.42
Other races	1 if race is not White or African American	0.28	0.45	0.23	0.42
Live with both parents	1 if live with both parents	0.55	0.50	0.55	0.50
Parent education: college	1 if at least one parent has at least college education	0.35	0.48	0.36	0.48
Parent job: professional	1 if at least one parent's job is a doctor, lawyer, scientist, teacher, librarian, nurse, manager, executive, director, technical/computer specialist, or radiologist	0.40	0.49	0.41	0.49
Parent on welfare	1 if at least one parent receives public assistance	0.09	0.29	0.09	0.29
Residential building quality	1 if the building in which the respondent lives is well kept	0.58	0.49	0.59	0.49
Single-family home	1 if the respondent lives in a detached single-family house or a single-family town house.	0.82	0.39	0.80	0.40

Notes: The variable in the parentheses is the reference category.

TABLE 1: MONTE CARLO SIMULATION RESULTS ($\gamma = 1, R^2 \approx 0.418$)

	$\phi_1 = 0.05$	$\phi_2 = 0.2$	$\beta_1 = 2$	$\beta_2 = 2$
	$n_r = 10, \bar{r} = 50$			
$\sigma_{12} = 0.1$				
2SLS-1	0.050(0.006)[0.022]	0.205(0.061)[0.232]	2.002(0.058)[0.212]	1.984(0.275)[1.122]
2SLS-2	0.050(0.012)[0.051]	0.227(0.186)[0.702]	2.018(0.118)[0.448]	1.892(1.126)[4.371]
GMM	0.050(0.011)[0.042]	0.208(0.090)[0.367]	2.008(0.090)[0.335]	1.981(0.537)[2.150]
$\sigma_{12} = 0.5$				
2SLS-1	0.052(0.009)[0.033]	0.103(0.105)[0.296]	2.001(0.060)[0.230]	2.418(0.483)[1.416]
2SLS-2	0.050(0.012)[0.050]	0.218(0.183)[0.721]	2.014(0.115)[0.452]	1.903(1.127)[4.279]
GMM	0.050(0.010)[0.041]	0.202(0.093)[0.365]	2.002(0.088)[0.340]	2.011(0.536)[2.172]
$\sigma_{12} = 0.9$				
2SLS-1	0.056(0.013)[0.046]	-0.084(0.284)[0.402]	1.989(0.071)[0.271]	3.235(1.235)[1.766]
2SLS-2	0.050(0.012)[0.050]	0.208(0.177)[0.738]	2.008(0.112)[0.433]	1.965(1.076)[4.454]
GMM	0.050(0.010)[0.042]	0.196(0.100)[0.383]	2.004(0.087)[0.321]	2.049(0.553)[2.271]
	$n_r = 10, \bar{r} = 100$			
$\sigma_{12} = 0.1$				
2SLS-1	0.050(0.004)[0.015]	0.200(0.040)[0.161]	1.997(0.038)[0.149]	2.011(0.194)[0.737]
2SLS-2	0.050(0.009)[0.035]	0.210(0.118)[0.510]	2.000(0.077)[0.311]	1.977(0.757)[3.018]
GMM	0.050(0.008)[0.028]	0.199(0.058)[0.229]	1.997(0.060)[0.231]	2.017(0.354)[1.315]
$\sigma_{12} = 0.5$				
2SLS-1	0.052(0.006)[0.023]	0.115(0.085)[0.205]	1.997(0.044)[0.159]	2.379(0.389)[0.990]
2SLS-2	0.050(0.009)[0.033]	0.215(0.124)[0.490]	2.001(0.075)[0.312]	1.967(0.743)[3.025]
GMM	0.050(0.008)[0.029]	0.199(0.062)[0.244]	1.999(0.059)[0.235]	2.022(0.377)[1.420]
$\sigma_{12} = 0.9$				
2SLS-1	0.055(0.009)[0.033]	-0.080(0.280)[0.270]	1.986(0.051)[0.193]	3.195(1.195)[1.275]
2SLS-2	0.050(0.009)[0.035]	0.205(0.124)[0.512]	1.999(0.078)[0.316]	1.992(0.766)[3.104]
GMM	0.050(0.008)[0.029]	0.201(0.066)[0.269]	1.998(0.061)[0.235]	2.029(0.413)[1.550]
	$n_r = 10, \bar{r} = 200$			
$\sigma_{12} = 0.1$				
2SLS-1	0.050(0.003)[0.011]	0.197(0.029)[0.112]	2.001(0.027)[0.100]	2.012(0.140)[0.552]
2SLS-2	0.050(0.006)[0.022]	0.206(0.084)[0.350]	2.006(0.054)[0.216]	1.953(0.530)[2.123]
GMM	0.050(0.005)[0.019]	0.197(0.041)[0.160]	2.002(0.041)[0.155]	2.026(0.239)[0.937]
$\sigma_{12} = 0.5$				
2SLS-1	0.052(0.005)[0.017]	0.110(0.090)[0.143]	1.994(0.029)[0.107]	2.379(0.381)[0.683]
2SLS-2	0.050(0.006)[0.022]	0.205(0.088)[0.349]	2.004(0.054)[0.214]	1.984(0.530)[2.108]
GMM	0.050(0.005)[0.019]	0.197(0.043)[0.166]	2.002(0.042)[0.159]	2.027(0.242)[0.969]
$\sigma_{12} = 0.9$				
2SLS-1	0.056(0.007)[0.024]	-0.080(0.280)[0.198]	1.983(0.034)[0.129]	3.181(1.181)[0.933]
2SLS-2	0.050(0.006)[0.022]	0.201(0.082)[0.337]	2.004(0.053)[0.212]	1.999(0.510)[2.065]
GMM	0.050(0.005)[0.019]	0.195(0.047)[0.181]	2.003(0.043)[0.159]	2.031(0.277)[1.068]

Notes: Robust measures of central tendency and dispersion in the empirical distribution of the estimates from 1000 simulation replications are reported: the median, the median of the absolute deviations (Med. AD), and the difference between the 0.1 and 0.9 quantile [Dec. Rge].

TABLE 2: MONTE CARLO SIMULATION RESULTS ($\gamma = 0.2, R^2 \approx 0.052$)

	$\phi_1 = 0.05$	$\phi_2 = 0.2$	$\beta_1 = 2$	$\beta_2 = 2$
	$n_r = 10, \bar{r} = 50$			
$\sigma_{12} = 0.1$				
2SLS-1	0.051(0.006)[0.025]	0.193(0.068)[0.260]	1.999(0.039)[0.148]	2.019(0.201)[0.761]
2SLS-2	0.050(0.034)[0.152]	0.211(0.431)[2.016]	2.000(0.088)[0.360]	1.974(1.374)[6.512]
GMM	0.048(0.028)[0.113]	0.241(0.260)[1.069]	1.998(0.073)[0.287]	1.864(0.862)[3.432]
$\sigma_{12} = 0.5$				
2SLS-1	0.063(0.015)[0.043]	-0.019(0.220)[0.396]	2.024(0.047)[0.178]	2.548(0.552)[1.096]
2SLS-2	0.051(0.033)[0.155]	0.234(0.447)[1.987]	1.996(0.085)[0.365]	1.813(1.401)[6.589]
GMM	0.051(0.027)[0.110]	0.209(0.262)[1.101]	1.991(0.076)[0.301]	1.855(0.850)[3.564]
$\sigma_{12} = 0.9$				
2SLS-1	0.086(0.036)[0.065]	-0.398(0.598)[0.639]	2.066(0.074)[0.217]	3.484(1.484)[1.573]
2SLS-2	0.053(0.034)[0.146]	0.176(0.444)[1.964]	2.014(0.092)[0.390]	2.085(1.403)[6.368]
GMM	0.055(0.026)[0.102]	0.133(0.259)[1.009]	2.005(0.073)[0.281]	2.147(0.786)[3.340]
	$n_r = 10, \bar{r} = 100$			
$\sigma_{12} = 0.1$				
2SLS-1	0.051(0.004)[0.016]	0.187(0.047)[0.172]	2.002(0.029)[0.103]	2.038(0.149)[0.539]
2SLS-2	0.049(0.025)[0.113]	0.228(0.383)[1.785]	1.992(0.067)[0.292]	1.930(1.273)[5.935]
GMM	0.050(0.020)[0.081]	0.233(0.194)[0.842]	1.992(0.051)[0.195]	1.858(0.649)[2.769]
$\sigma_{12} = 0.5$				
2SLS-1	0.063(0.013)[0.030]	-0.016(0.216)[0.300]	2.024(0.035)[0.114]	2.547(0.547)[0.796]
2SLS-2	0.050(0.026)[0.114]	0.231(0.399)[1.907]	1.995(0.066)[0.290]	1.873(1.323)[5.946]
GMM	0.050(0.020)[0.081]	0.234(0.200)[0.908]	1.992(0.050)[0.206]	1.881(0.652)[2.881]
$\sigma_{12} = 0.9$				
2SLS-1	0.085(0.035)[0.043]	-0.412(0.612)[0.483]	2.069(0.070)[0.161]	3.548(1.548)[1.160]
2SLS-2	0.049(0.024)[0.106]	0.160(0.373)[1.813]	2.003(0.063)[0.269]	2.189(1.193)[5.744]
GMM	0.050(0.019)[0.078]	0.186(0.200)[0.856]	2.000(0.049)[0.201]	1.993(0.691)[2.849]
	$n_r = 10, \bar{r} = 200$			
$\sigma_{12} = 0.1$				
2SLS-1	0.051(0.003)[0.012]	0.190(0.033)[0.122]	2.001(0.019)[0.068]	2.022(0.097)[0.372]
2SLS-2	0.050(0.018)[0.076]	0.214(0.330)[1.600]	1.996(0.048)[0.207]	1.972(1.085)[5.167]
GMM	0.050(0.014)[0.057]	0.227(0.145)[0.697]	1.994(0.035)[0.140]	1.895(0.495)[2.194]
$\sigma_{12} = 0.5$				
2SLS-1	0.063(0.013)[0.021]	-0.011(0.211)[0.204]	2.023(0.027)[0.079]	2.529(0.529)[0.541]
2SLS-2	0.049(0.017)[0.069]	0.190(0.315)[1.336]	1.999(0.044)[0.180]	2.037(0.999)[4.580]
GMM	0.050(0.014)[0.053]	0.210(0.147)[0.666]	1.994(0.034)[0.133]	1.923(0.512)[2.210]
$\sigma_{12} = 0.9$				
2SLS-1	0.086(0.036)[0.030]	-0.393(0.593)[0.327]	2.066(0.067)[0.109]	3.492(1.492)[0.782]
2SLS-2	0.051(0.017)[0.068]	0.179(0.322)[1.383]	2.003(0.040)[0.185]	2.097(1.079)[4.579]
GMM	0.051(0.013)[0.053]	0.195(0.151)[0.697]	1.997(0.034)[0.136]	2.005(0.501)[2.345]

Notes: Robust measures of central tendency and dispersion in the empirical distribution of the estimates from 1000 simulation replications are reported: the median, the median of the absolute deviations (Med. AD), and the difference between the 0.1 and 0.9 quantile [Dec. Rge].

TABLE 3: ESTIMATION OF THE NETWORK MODEL

Dep. Var.: Delinquency Index			
	2SLS-1	2SLS-2	GMM
Local-aggregate peer effect	0.0219***	0.0979	0.0937*
	(0.0092)	(0.0883)	(0.0516)
Local-average peer effect	0.3452*	0.1859	-0.0660
	(0.2076)	(0.3405)	(0.1746)
Age	-0.0344*	-0.0178	-0.0164
	(0.0204)	(0.0234)	(0.0219)
Female	-0.3245***	-0.3296***	-0.3259***
	(0.0265)	(0.0293)	(0.0272)
African American	-0.0356	0.0057	-0.0072
	(0.0576)	(0.0867)	(0.0821)
Other races	0.0406	0.0592	0.0609
	(0.0447)	(0.0510)	(0.0484)
Grade	0.0247	-0.0135	-0.0134
	(0.0260)	(0.0383)	(0.0364)
Live with both parents	-0.0568*	-0.0644	-0.0813***
	(0.0305)	(0.0398)	(0.0295)
Parent education: college	0.0113	0.0171	0.0083
	(0.0332)	(0.0385)	(0.0353)
Parent job: professional	0.0107	-0.0051	0.0012
	(0.0309)	(0.0350)	(0.0319)
Parent on welfare	0.1298***	0.1295**	0.1279***
	(0.0472)	(0.0566)	(0.0516)
Residential building quality	-0.1068***	-0.1366***	-0.1357***
	(0.0286)	(0.0327)	(0.0304)
Single-family home	-0.0613*	-0.0520	-0.0649*
	(0.0371)	(0.0425)	(0.0367)
Contextual effects	Yes	Yes	Yes
Network fixed effects	Yes	Yes	Yes
Cragg-Donald statistic	2.388	17.422	
OIR test *p-value*	*0.061*	*0.739*	*0.798*
N. Obs.	7,735	7,735	7,735

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

TABLE 4: DYADIC REGRESSION MODEL

Dep. Var.: Probability to be friends	
	PROBIT
Age	0.0294***
	(0.0023)
Grade	0.1317***
	(0.0031)
Female	0.0747***
	(0.0092)
Race	0.1684***
	(0.0113)
Live with both parents	0.0437***
	(0.0093)
Parent education: college	0.0323***
	(0.0099)
Parent job: professional	0.0313***
	(0.0097)
Parent on welfare	0.1014***
	(0.0129)
Residential building quality	0.0771***
	(0.0094)
Single-family home	0.1105***
	(0.0105)
McFadden's pseudo R^2	0.080
N. Obs.	252,944

Notes: The explanatory variables are differences of the listed characteristics between individuals. Standard errors in parentheses are reported. Statistical significance: ***p<0.01; **p<0.05; *p<0.1.

TABLE 5: ESTIMATION OF THE NETWORK MODEL
– DIFFERENT SAMPLES AND SETS OF CONTROLS –

Dep. Var.: Delinquency Index				
	Networks of Sizes 3~305			All Networks
	GMM	GMM	GMM	GMM
Local-aggregate peer effect	0.1198***	0.1136*	0.0937*	0.0793*
	(0.0497)	(0.0605)	(0.0516)	(0.0415)
Local-average peer effect	-0.0971	-0.1022	-0.0660	0.2669
	(0.1746)	(0.1929)	(0.1746)	(0.2067)
Age	-0.0026	-0.0037	-0.0164	-0.0252
	(0.0206)	(0.0219)	(0.0219)	(0.0203)
Female	-0.3144***	-0.3259***	-0.3259***	-0.3389***
	(0.0257)	(0.0272)	(0.0272)	(0.0271)
African American	0.0198	-0.0028	-0.0072	0.0075
	(0.0497)	(0.0823)	(0.0821)	(0.0808)
Other races	0.0721*	0.0648	0.0609	0.0643
	(0.0427)	(0.0472)	(0.0484)	(0.0470)
Grade	-0.0177	-0.0279	-0.0134	-0.0097
	(0.0230)	(0.0363)	(0.0364)	(0.0369)
Live with both parents			-0.0813***	-0.0415
			(0.0295)	(0.0273)
Parent education: college			0.0083	-0.0259
			(0.0353)	(0.0317)
Parent job: professional			0.0012	0.0300
			(0.0319)	(0.0294)
Parent on welfare			0.1279***	0.1069**
			(0.0516)	(0.0469)
Residential building quality			-0.1357***	-0.1356***
			(0.0304)	(0.0279)
Single-family home			-0.0649*	-0.0649*
			(0.0367)	(0.0350)
Contextual effects	No	Yes	Yes	Yes
Network fixed effects	Yes	Yes	Yes	Yes
N. Obs.	7,735	7,735	7,735	9,396

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

TABLE 6: ESTIMATION OF THE NETWORK MODEL
– ALTERNATIVE SPECIFICANTIONS OF PEER EFFECTS –

Dep. Var.: Delinquency Index		
	Local-aggregate Model	Local-average Model
	GMM	GMM
Local-aggregate peer effect	0.0723***	
	(0.0206)	
Local-average peer effect		0.0733
		(0.1038)
Age	-0.0192	-0.0236
	(0.0213)	(0.0212)
Female	-0.3256***	-0.3177***
	(0.0271)	(0.0269)
African American	-0.0127	0.0133
	(0.0818)	(0.0811)
Other races	0.0594	0.0761
	(0.0481)	(0.0477)
Grade	-0.0119	0.0075
	(0.0365)	(0.0362)
Live with both parents	-0.0803***	-0.0958***
	(0.0295)	(0.0297)
Parent education: college	0.0079	0.0070
	(0.0353)	(0.0354)
Parent job: professional	0.0037	0.0095
	(0.0318)	(0.0317)
Parent on welfare	0.1212***	0.1010**
	(0.0504)	(0.0498)
Residential building quality	-0.1333***	-0.1263***
	(0.0300)	(0.0297)
Single-family home	-0.0654*	-0.0729**
	(0.0367)	(0.0366)
Contextual effects	Yes	Yes
Network fixed effects	Yes	Yes
N. Obs.	7,735	7,735

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

TABLE 7: CHARACTERISTICS OF THE KEY PLAYER

	Key Players		Other Students		
	Mean	SD	Mean	SD	*p-value*
Own Characteristics					
Age	15.48	1.73	15.42	1.71	*0.24*
Grade	3.50	1.65	3.49	1.65	*0.43*
Female	0.52	0.50	0.54	0.50	*0.11*
White American	0.51	0.50	0.55	0.50	*0.03*
African American	0.24	0.43	0.22	0.42	*0.15*
Other races	0.25	0.43	0.23	0.42	*0.11*
Live with both parents	0.50	0.50	0.55	0.50	*0.01*
Parent education: college	0.43	0.50	0.36	0.48	*0.00*
Parent job: professional	0.46	0.50	0.40	0.49	*0.00*
Parent on welfare	0.07	0.25	0.10	0.30	*0.00*
Residential building quality	0.58	0.49	0.59	0.49	*0.23*
Single-family home	0.80	0.40	0.80	0.40	*0.49*
Number of Friends	3.35	2.39	2.87	2.21	*0.00*
Friends' Characteristics					
Age	15.43	1.59	15.40	1.58	*0.34*
Grade	3.45	1.55	3.49	1.57	*0.27*
Female	0.54	0.33	0.55	0.37	*0.44*
White American	0.51	0.43	0.55	0.46	*0.01*
African American	0.24	0.38	0.22	0.39	*0.15*
Other races	0.25	0.35	0.22	0.37	*0.02*
Live with both parents	0.52	0.35	0.56	0.38	*0.01*
Parent education: college	0.38	0.34	0.36	0.38	*0.21*
Parent job: professional	0.43	0.34	0.41	0.39	*0.09*
Parent on welfare	0.09	0.19	0.09	0.22	*0.36*
Residential building quality	0.58	0.33	0.61	0.38	*0.07*
Single-family home	0.81	0.28	0.80	0.32	*0.20*

Notes: The two-sample T- test p-value is reported in the last column.

FIGURE 1: WHO IS THE KEY PLAYER?