IZA

DISCUSSION PAPER SERIES

IZA DP No. 3859

Peer Effects and Social Networks in Education

Antoni Calvó-Armengol
Eleonora Patacchini
Yves Zenou

November 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Peer Effects and Social Networks in Education

Antoni Calvó-Armengol
Universitat Autònoma de Barcelona

Eleonora Patacchini
Università di Roma "La Sapienza" and IZA

Yves Zenou
Stockholm University, IFN, GAINS and IZA

Discussion Paper No. 3859
November 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3859
November 2008

ABSTRACT

Peer Effects and Social Networks in Education*

This paper studies whether structural properties of friendship networks affect individual outcomes in education. We first develop a model that shows that, at the Nash equilibrium, the outcome of each individual embedded in a network is proportional to her Katz-Bonacich centrality measure. This measure takes into account both direct and indirect friends of each individual but puts less weight to her distant friends. We then bring the model to the data by using a very detailed dataset of adolescent friendship networks. We show that, after controlling for observable individual characteristics and unobservable network specific factors, the individual's position in a network (as measured by her Katz-Bonacich centrality) is a key determinant of her level of activity. A standard deviation increase in the Katz-Bonacich centrality increases the pupil school performance by more than 7 percent of one standard deviation.

JEL Classification: A14, C31, C72, I21

Keywords: centrality measure, peer influence, network structure, school performance

Corresponding author:

Yves Zenou
Department of Economics
Stockholm University
106 91 Stockholm
Sweden
E-mail: yves.zenou@ne.su.se

* We are grateful to the editor Enrique Sentana and two anonymous referees for very constructive comments. We also thank Yann Bramoulle, David Genesove, Thierry Magnac, Federico Martellosio, Michele Pellizzari, and Barbara Petrongolo for helpful comments and suggestions. Financial support from the Fundación Ramón Areces, the Spanish Ministry of Science and Technology and FEDER under grant SEJ2005-01481ECON, and from the Barcelona Economics Program of XREA is gratefully acknowledged by Antoni Calvó-Armengol. Yves Zenou thanks the Marianne and Marcus Wallenberg Foundation for financial support.
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).
Antoni Calvó-Armengol passed away in November 2007. His friendship, creativity and talent will be deeply missed.

Yo soy yo y mi circunstancia, Ortega y Gasset

1 Introduction

It is commonly observed, both in ethnographic and empirical studies, that the behavior of individual agents is affected by that of their peers. This is particularly true in education, crime, labor markets, fertility, participation to welfare programs, etc.[1] The detection and measure of such peer effects is, however, a very difficult exercise. Two main assumptions, not always made explicit, usually accompany this detection and measure. First, peer effects are conceived as an average intra-group externality that affects identically all the members of a given group. Second, the group boundaries for such an homogeneous effect are often arbitrary, and at a quite aggregate level, in part due to the constraints imposed by the available disaggregated data. For instance, peer effects in crime are often measured at the neighborhood level using local crime rates, peer effects in school at the classroom or school level using average school achievements, etc.

In this paper, we propose a peer-effect model where we relate analytically equilibrium behavior to network location. Using a unique dataset of friendship networks from the National Longitudinal Survey of Adolescent Health (Add Health), we then test the empirical salience of our model predictions. We find that a standard deviation increase in our equilibrium measure of network location accounts for a 7 percent standard deviation increase in pupil school performance.

In what follows, we first describe the theoretical model of peer effects, and then the empirical measure stemming from our theoretical analysis.

Our model starts from two simple premises. First, individual outcome is additively separable into an idiosyncratic component and a peer effect component. This is true, for instance, if peer influence acts as a multiplier on the outcome of the isolated individual. Second, peer effects aggregate at the group level the collection of bilateral cross influences the members of this group may or may not exert on each other. Consistent with this approach, the smallest unit of analysis for peer effects is the dyad, a two-person group. The collection of active bilateral influences or dyads constitutes a social network. In this network, each player chooses an optimal level of activity.

Consider an individual connected by a network of peer influences. In this network, payoffs are interdependent, and each agent reaps complementarities from all her direct network peers. We compute the Nash equilibrium of this peer effect game when agents choose their peer effort simultaneously. Given that payoff complementarities are rooted in direct friendship ties, equilibrium decisions generally differ across agents, and in a manner that reflects the existing heterogeneity in friendship ties. Because of this heterogeneity in friendships ties, the equilibrium peer influence does not boil down to a common average externality exerted by a group of agents on all its members. Rather, this intra-group externality varies across group members depending on each agent particular

[1] Durlauf (2004) offers an exhaustive and critical survey.

location in the network of dyadic influences.[2]

The sociology literature abounds in network measures that assign to each node in a network a scalar associated with the geometric intricacies of the sub-network surrounding that particular node (Wasserman and Faust, 2004). When the analysis is restricted to linear-quadratic utility functions, it turns out that one (and only one) of such network measures captures exactly how each agent subsumes at equilibrium the network peer influence. This is the Katz-Bonacich network centrality, due to Katz (1953) and later extended by Bonacich (1987). The Katz-Bonacich network centrality counts, for each node in a given network, the *total* number of direct and indirect paths of any length in the network stemming from that node. Paths are weighted by a geometrically decaying factor (with path length). Therefore, the Katz-Bonacich centrality is not parameter free. It depends both on the network topology and on the value of this decaying factor. This has important implications for the empirical analysis.

Our main theoretical result establishes that the peer effects game has a unique Nash equilibrium where each agent strategy is proportional to her Katz-Bonacich centrality measure. We provide a closed-form expression for this Nash-Katz-Bonacich linkage. This equilibrium mapping between network structure and effort levels holds under a condition that involves the network eigenvalues. This condition guarantees that the level of network complementarities are low enough to prevent the corresponding positive feed-back loops in agent's efforts to escalate without bound. Under this condition, which is reminiscent but less demanding than standard dominance diagonal conditions in industrial organization, payoff functions are enough 'concave' so that interiority (and uniqueness) are obtained.[3]

One may wonder why the exact mapping between network location and equilibrium outcome is more intricate than simply counting direct network links, and also requires to account for weighted indirect network links. Recall, indeed, that the payoff interdependence is such that each agent only cares about the behavior of her direct dyad partners. At equilibrium, though, each agent has to anticipate the actual behavior of her dyad partners to take on an optimal action herself. For this reason, every dyad exerts a strategic externality on overlapping dyads, and the equilibrium effort levels of each agent must reflect this externality. As a matter of fact, the Katz-Bonacich centrality captures adequately how these dyads overlap boils down to an equilibrium (fixed point) pattern of decisions. At equilibrium, individual decisions emanate from all the existing network chains of direct and indirect contacts stemming from each node, which is a feature characteristic of Katz-Bonacich centrality.

That Nash equilibrium behavior can be exactly described by a network measure is very con-

[2] Calvó-Armengol and Jackson (2004) describes a network model of information exchange that opens the black-box of peer effects in drop-out decisions, that vary at equilibrium with network location. Calvó-Armengol and Zenou (2004) proposes a similar analysis for crime decisions. Goyal (2007) and Jackson (2007, 2008) offer exhaustive surveys of the growing literature on the economics of social networks.

[3] This unique equilibrium is also stable, and thus would naturally emerge from a tatônnement process.

venient. For instance, the Nash-Katz-Bonacich linkage has important implications both for comparative statics and for optimal network policy design (Ballester et al., 2006). Here, we explore its implications for empirical analysis by generalizing the model of Ballester et al. (2006) for the case of ex ante intrinsic heterogeneity, i.e. the observable characteristics of each individual, like e.g. her age, race, gender, parental education, etc. This generalization turns out to be non-trivial (see Proposition 1 and Theorem 1) and much more adequate for the empirical analysis.

We test the predictions of our peer effects model by using a very detailed and unique dataset of friendship networks from the National Longitudinal Survey of Adolescent Health (Add Health). We explore the role of network location for peer effects in education. We obtain a clear empirical prediction: the intensity of peer effects on education is well-accounted by the position of each individual in a network.

AddHealth contains rich information on friendship networks. Using the in-school friendship nominations data, we obtain a sample of 11,964 pupils distributed over 199 networks.

Our theoretical set up provides the behavioral foundation for the estimation of a version of the so-called *spatial error model* (see, e.g. Anselin, 1988). Indeed, in the theoretical model, peer-effect efforts are additively separated from the individual observable characteristics. A maximum likelihood approach thus produces an average estimate of the strength of the dyadic influences within the network. This parameter enters in the calculation of the Katz-Bonacich centralities, and corresponds to the decaying weight for path length.

The empirical issues that arise when measuring peer effects are tackled. Firstly, using the particular structure of social networks, we show to what extent it is possible to disentangle endogenous from exogenous (contextual) effects (Manski, 1993), thus identifying peer effects. Secondly, the richness of the information provided by the AddHealth data and the use of both within and between network variations allow us to control for issues stemming from endogenous network formation and unobserved individual, school and network heterogeneity that might affect our estimation results.

In economics, the influence of peers on education outcomes has been extensively studied. The standard approach is either to use instrumental variables (see e.g. Evans et al., 1992) or a natural experiment (see e.g. Angrist and Lavy, 1999; Sacerdote, 2001; Zimmerman, 2003) to single out a causal relationship. To the best of our knowledge, there are nearly no studies that have adopted a more structural approach to test a specific peer effect model in education.[4] This is what is done in the present paper. To be more precise, the novelty of our work is threefold. First, from a conceptual point of view, we stress the role of the structure of social networks in explaining individual behavior. Second, from a more operational point of view, we build a theoretical model of peer effects that envisions group influence as an equilibrium outcome, which aggregates the collection of active dyadic peer influences. The analysis of such model wedges a bridge between

[4] Glaeser et al. (1996) study peer effecs in criminal behavior testing a specific model. De Giorgi et al. (2007) provide empirical evidence on peer effects in the choice of college major.

the economics literature, here Nash equilibrium, and the sociology literature, here Katz-Bonacich centrality. Third, we conduct a direct empirical test of our model on the network structure of peer effects using a detailed dataset on friendship networks, AddHealth, with particular attention to the relevant econometric problems. More precisely, we characterize the exact conditions on the geometry of the peer network, so that the model is fully identified (see Durlauf, 2002, for a critical account of salient identification problems in the empirical analysis of social interactions).

In sociology, it has long been recognized that not only friends but also the *structure* of friendships ties are a determinant of individual behavior. The novelty, here, lies in the fact that we model explicitly individual incentives as tailored by the network of relationships, and conduct a full-fledged equilibrium analysis that relates topology to outcome. This equilibrium analysis then guides our empirical analysis. In particular, it singles out the Katz-Bonacich network centrality as the adequate topological index to explain outcomes. Besides, our analysis calls for exploiting both within and between network variations to explain behavior.

The study of peer effects is, however, complex and our analysis does not cover all its possibilities. There are potential limitations and we propose a variety of alternative theoretical set-ups and empirical tests. In particular, such alternative formulations highlight the fact that the impact of a student's effort on her outcomes due to peer effects is not so easily separated or identified. This paper presents one possible valid approach for the problematic identification and estimation of peer effects using a network structure.

2 A network model of peer effects

2.1 Model

We develop a network model of peer effects, where the network reflects the collection of active bilateral influences.

The network $N = \{1, \ldots, n\}$ is a finite set of agents. We keep track of social connections by a network $\mathbf{g}$, where $g_{ij} = 1$ if i and j are direct friends, and $g_{ij} = 0$, otherwise. Given that friendship is a reciprocal relationship, we set $g_{ij} = g_{ji}$. We also set $g_{ii} = 0$.

Preferences Denote by y_i^0 the effort of individual i absent of any peer influence, and by z_i the peer effort whose returns depend on others' peer efforts. Each agent i selects both efforts $y_i^0 \geq 0$ and $z_i \geq 0$, and obtains a payoff $u_i(\mathbf{y}^0, \mathbf{z}; \mathbf{g})$ that depends on the underlying network $\mathbf{g}$, in the following way:[5]

[5] In Appendix 1, we develop a model with a more general version of this utility function.

$$u_i(\mathbf{y}^0, \mathbf{z}; \mathbf{g}) = \theta_i\, y_i^0 - \frac{1}{2}\left(y_i^0\right)^2 + \mu g_i z_i - \frac{1}{2} z_i^2 + \phi \sum_{j=1}^{n} g_{ij} z_i z_j. \tag{1}$$

where $\phi > 0$, $\mu > 0$, and $g_i = \sum_{j=1}^n g_{ij}$ is the number of direct links of individual i. This utility function is additively separable in the idiosyncratic effort component and the peer effect contribution. The component θ_i introduces the *exogenous heterogeneity* that captures the *observable* differences between individuals. Examples of such heterogeneity are agent i's parents' education, neighborhood where she lives, age, sex, race, etc. but also the average characteristics of the individuals directly linked to i, i.e. average level of parental education of i's friends, etc. (*contextual effects*). To be more precise,

$$\theta_i(\mathbf{x}) = \sum_{m=1}^{M} \beta_m x_i^m + \frac{1}{g_i} \sum_{m=1}^{M} \sum_{j=1}^{n} \gamma_m g_{ij}\, x_j^m \tag{2}$$

where x_i^m is a set of M variables accounting for observable differences in individual, neighborhood and school characteristics of individual i, and β_m, γ_m are parameters.

The peer effect component is also heterogeneous, and this *endogenous heterogeneity* reflects the different locations of individuals in the friendship network $\mathbf{g}$ and the resulting effort levels. To be more precise, bilateral influences are captured by the following cross derivatives, for $i \neq j$:

$$\frac{\partial^2 u_i(\mathbf{y}^0, \mathbf{z}; \mathbf{g})}{\partial z_i \partial z_j} = \phi g_{ij} \geq 0. \tag{3}$$

When i and j are direct friends, the cross derivative is $\phi > 0$ and reflects a *strategic complementarity* in efforts. When i and j are not direct friends, this cross derivative is zero. In the context of education, $\phi > 0$ means that if two students are friends, i.e. $g_{ij} = 1$, and if j increases her effort (for example spends more time in studying and doing homework), then i will experience an increase in her (marginal) utility if she also increases her effort. Allowing ϕ to be less than zero would imply that under such a scenario student i will experience an increase in her (marginal) utility if she also decreases her effort. This would be true only under special circumstances (e.g., if there is some collective work involved), which are not the rules in the context under analysis.[6]

Note that the utility (1) is concave in own decisions, and displays decreasing marginal returns in own effort levels. Instead, we have complementarity of efforts across connected agents. Agents' equilibrium peer efforts $(z_1, ..., z_n)$ thus depend on the pattern of bilateral influences reflected in $\mathbf{g}$, and on the intensity of such bilateral influences, captured by ϕ. Given that complementarities are rooted in direct friendship ties, having more friends increases one's effort decision at equilibrium.

[6] The empirical results contained in Section 8, in fact, show that the estimates of ϕ are strictly positive for all the networks in our dataset. Strategic substituabilities in efforts, as captured by $\phi < 0$, would be more natural when there are strategic interactions in providing a public good that is non-excludable (see e.g. Bramoulle and Kranton, 2007).

Equilibrium effort levels thus generally differ across agents in a manner that reflects the existing heterogeneity in friendship ties.

Example 1 Consider the network $\mathbf{g}$ in Figure 1 with three agents.

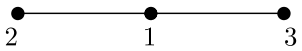


Figure 1. Three agents on a line.

This network results from the overlap between two different dyads with a common partner, agent 1. Agent 2 reaps direct complementarities from agent 1 in one dyad whom, in turn, reaps direct complementarities from agents 2 and 3 in both dyads. Thus, through the interaction with the central agent, peripheral agents end up reaping complementarities indirectly from each other. For this reason, the equilibrium decisions in each dyad cannot be analyzed independently of each other. Rather, each dyad exerts a strategic externality on the other one, and the equilibrium effort level of each agent reflects this externality, and the role each agent may play as a driver for the externality.

In what follows, we describe a network centrality measure that turns out to capture exactly how each agent subsumes these strategic externalities across dyads as a function of the location she holds in the network that results from the dyads' overlap.

2.2 Results

We first define a network centrality measure due to Katz (1953), and latter extended by Bonacich (1987), that proves useful to describe the equilibrium of the peer network model.

The Katz-Bonacich network centrality The Katz-Bonacich centrality measures the importance of a given node in a network. To assess how well located a node is, Katz proposed the following simple recurrent formula. To start with, every node i is assigned some initial value ϕg_i, proportional to its connectivity $g_i = \sum_{j=1}^{n} g_{ij}$. Here, $0 \leq \phi$ is some non-negative scalar. Then, this value is augmented by adding up the values of the nodes located one-link away from i, two-links away, and so on. A factor that decays with the distance discounts the contribution of all these nodes: the value of $k-$link away nodes is weighted by ϕ^{k-1}.

Given a network $\mathbf{g}$ and a scalar ϕ, we denote by $\mathbf{b}(\mathbf{g}, \phi)$ the vector whose coordinates correspond to the Katz-Bonacich centralities of all the network nodes.

A more formal expression for the recurrent formula defined above is as follows.

To each network $\mathbf{g}$, we associate its adjacency matrix $\mathbf{G} = [g_{ij}]$ that keeps track of the direct connections in $\mathbf{g}$. The kth power $\mathbf{G}^k = \mathbf{G}^{(k \ times)}_{\cdots}\mathbf{G}$ of this adjacency matrix then keeps track of indirect connections in $\mathbf{g}$.

More precisely, the coefficient in the (i, j) cell of $\mathbf{G}^k$ gives the number of paths of length k in $\mathbf{g}$ between i and j. Note that, by definition, a path between i and j needs not to follow the shortest possible route between those agents. For instance, when $g_{ij} = 1$, the sequence $ij \to ji \to ij$ constitutes a path of length three in $\mathbf{g}$ between i and j.

Denote by $\mathbf{1}$ the vector of ones. Then, $\mathbf{G1}$ is the vector of node connectivities, while the coordinates of $\mathbf{G}^k\mathbf{1}$ give the total number of paths of length k that emanate from the corresponding network node.

The vector of Katz-Bonacich centralities is thus:

$$\mathbf{b}(\mathbf{g}, \phi) = \phi\mathbf{G1} + \phi^2\mathbf{G}^2\mathbf{1} + \phi^3\mathbf{G}^3\mathbf{1} + \cdots = \sum_{k=0}^{+\infty} \phi^k \mathbf{G}^k \cdot (\phi\mathbf{G1}).$$

If ϕ is small enough, this infinite sum converges to a finite value, which is $\sum_{k=0}^{+\infty} \phi^k \mathbf{G}^k = (\mathbf{I} - \phi\mathbf{G})^{-1}$, where $\mathbf{I}$ is the identity matrix. We can then write the vector of Katz-Bonacich centralities as follows:

$$\mathbf{b}(\mathbf{g}, \phi) = (\mathbf{I} - \phi\mathbf{G})^{-1} \cdot (\phi\mathbf{G1}). \tag{4}$$

Observe that, by definition, the Katz-Bonacich centrality of a given node is zero when the network is empty. It is also null when $\phi = 0$, and is increasing and convex with ϕ. Finally, it is bounded from below by ϕ times the node connectivity, that is, $b_i(\mathbf{g}, \phi) \geq \phi g_i$.

It turns out that an exact strict upper bound for the scalar ϕ is given by the inverse of the largest eigenvalue of $\mathbf{G}$ (Debreu and Herstein, 1953). The largest eigenvalue of the adjacency matrix is also called spectral index of the network.

Example 1 (continued) Consider the network $\mathbf{g}$ in Figure 1. The corresponding adjacency matrix is,

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

while the vector of node connectivities is given by:

$$\mathbf{G1} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

The kth powers of $\mathbf{G}$ are then, for $k \geq 1$:

$$\mathbf{G}^{2k} = \begin{bmatrix} 2^k & 0 & 0 \\ 0 & 2^{k-1} & 2^{k-1} \\ 0 & 2^{k-1} & 2^{k-1} \end{bmatrix} \quad \text{and} \quad \mathbf{G}^{2k+1} = \begin{bmatrix} 0 & 2^k & 2^k \\ 2^k & 0 & 0 \\ 2^k & 0 & 0 \end{bmatrix}.$$

For instance, we deduce from $\mathbf{G}^3$ that there are exactly two paths of length three between agents 1 and 2, which are $12 \to 21 \to 12$ and $12 \to 23 \to 32$.

When ϕ is small enough,[7] the vector of Katz-Bonacich network centralities is:

$$\mathbf{b}(\mathbf{g},\phi) = \begin{bmatrix} b_1(\mathbf{g},\phi) \\ b_2(\mathbf{g},\phi) \\ b_3(\mathbf{g},\phi) \end{bmatrix} = \frac{\phi}{1-2\phi^2}\begin{bmatrix} 2+2\phi \\ 1+2\phi \\ 1+2\phi \end{bmatrix}. \tag{5}$$

Not surprisingly, the center (agent 1) is more central than the peripheral agents 2 and 3.

Equilibrium behavior We now characterize the Nash equilibrium of the game where agents choose their effort levels $y_i^0 \geq 0$ and $z_i \geq 0$ simultaneously. Each individual i maximizes (1) and we obtain the following best reply function for each $i = 1, ..., n$:

$$y_i^{0*}(\mathbf{x}) = \theta_i(\mathbf{x}) = \sum_{m=1}^{M} \beta_m x_i^m + \frac{1}{g_i}\sum_{m=1}^{M}\sum_{j=1}^{n} \gamma_m g_{ij}\, x_j^m \tag{6}$$

$$z_i^*(\mathbf{g}) = \mu g_i + \phi \sum_{j=1}^{n} g_{ij} z_j \tag{7}$$

The optimal exogenous and endogenous peer efforts are given by (6) and (7), and the individual outcome is the sum of these two different efforts, namely:

$$y_i^*(\mathbf{x},\mathbf{g}) = \underbrace{y_i^{0*}(\mathbf{x})}_{\text{idiosyncratic}} + \underbrace{z_i^*(\mathbf{g})}_{\text{peer effect}}. \tag{8}$$

In other words, we can decompose additively individual behavior into an exogenous part and an endogenous peer effect component that depends on the individual under consideration.

Denote by $\omega(\mathbf{g})$ the largest eigenvalue of the adjacency matrix $\mathbf{G} = [g_{ij}]$ of the network.

Proposition 1 *Suppose that $\phi\omega(\mathbf{g}) < 1$. Then, the individual equilibrium outcome is uniquely defined and given by:*

$$y_i^*(\mathbf{x},\mathbf{g}) = \theta_i(\mathbf{x}) + \frac{\mu}{\phi} b_i(\mathbf{g},\phi). \tag{9}$$

Proof. See Appendix 1. ▮

Let us rewrite (9) as

$$y_i^*(\mathbf{x},\mathbf{g}) = \left(1 + \frac{\mu}{\phi}\frac{b_i(\mathbf{g},\phi)}{\theta_i^0(\mathbf{x})}\right) y_i^{0*}(\mathbf{x}),$$

where $y_i^{0*}(\mathbf{x}) \equiv \theta_i(\mathbf{x})$. We see that the peer influence acts as a multiplier on the behavior of the isolated individual. The value of this multiplier varies across individuals as a function of

[7] Here, the largest eigenvalue of $\mathbf{G}$ is $\sqrt{2}$, and so the exact strict upper bound for ϕ is $1/\sqrt{2}$.

their location in the network $\mathbf{g}$, as captured by $b_i(\mathbf{g}, \phi)$. It also depends on the idiosyncratic characteristics of the individual through $\theta_i(\mathbf{x})$. When all the agents are isolated, there is no network, $b_i(\varnothing, \phi) = 0$, for all i, and so $y_i^*(\mathbf{x}, \varnothing) = y_i^{0*}(\mathbf{x})$. We have no multiplier peer effect and effort only depends on each individual i's idiosyncratic characteristic $y_i^{0*}(\mathbf{x}) \equiv \theta_i(\mathbf{x})$. Similarly, the multiplier peer effect disappears when $\mu = 0$, that is, there is no pure structural component on the peer effort (7).

The condition $\phi\omega(\mathbf{g}) < 1$ in Proposition 1 stipulates that network complementarities must be small enough compared to own concavity in order to prevent the positive feed-back loops triggered by such complementarities to escalate without bound.[8] Note that the condition $\phi\omega(\mathbf{g}) < 1$ does not bound the absolute values for these cross effects, but only their relative magnitude. Network complementarities are measured by the compound index $\phi\omega(\mathbf{g})$, where ϕ refers to the intensity of each non-zero cross effect, while $\omega(\mathbf{g})$ captures the population-wide pattern of these positive cross effects.

The largest eigenvalue increases with link addition, so that $\mathbf{g}' \supseteq \mathbf{g}$ implies $\omega(\mathbf{g}') \geq \omega(\mathbf{g})$. Therefore, the denser the network of local complementarities, the more stringent the condition in Proposition 1. The highest value for the largest eigenvalue is obtained for the complete network, where every agent is directly linked to every other agent, and is equal to $n - 1$.[9]

Katz-Bonacich centrality is the right network index to account for equilibrium behavior when the utility functions are linear-quadratic. In (1), the local payoff interdependence is restricted to direct network contacts. At equilibrium, though, this local payoff interdependence spreads all over the network through the overlap of direct friendship clusters.[10] Katz-Bonacich centrality precisely reflects how individual decisions feed into each other along any direct and indirect network path.

Example 1 (continued) Consider the network $\mathbf{g}$ in Figure 1. When $\phi < 1/\sqrt{2}$, the individual equilibrium outcome is uniquely defined by:

$$y_1^* = \theta_1 + \mu \left(\frac{2 + 2\phi}{1 - 2\phi^2} \right) \tag{10}$$

$$y_2^* = \theta_2 + \mu \left(\frac{1 + 2\phi}{1 - 2\phi^2} \right) \tag{11}$$

$$y_2^* = \theta_3 + \mu \left(\frac{1 + 2\phi}{1 - 2\phi^2} \right) \tag{12}$$

[8] We discuss in detail the implications of condition $\phi\omega(\mathbf{g}) < 1$ in Section 5.2 below.

[9] A sufficient condition for the Nash-Katz-Bonacich linkage of Proposition 1 to hold for all networks is thus $\phi(n - 1) < 1$.

[10] At equilibrium, i's effort decision depends on j's effort decision, for all j such that $g_{ij} = 1$. But j's effort decision depends, in turn, on k's effort decision, for all k such that $g_{jk} = 1$. Therefore, i's decision depends (indirectly) on k's decision, for all k located two-links away from i. And so on.

The outcome of individual i depends both on her ex ante heterogeneity ($y_i^{0*} \equiv \theta_i^0$) and her location in the network (as measured by her Katz-Bonacich centrality index). Thus, even if individual 1 is the most central player in the network and has the highest Katz-Bonacich centrality, she does not always obtain the highest outcome because of different ex ante heterogeneities.

This example allows us to highlight the different roles of ϕ and μ. Clearly, ϕ measures the intensity of the purely imitative (endogenous) effect of peers. Now fix $\phi = 0$. Then, $z_i = \mu\, g_i$, and thus $y_i^* = \theta_i^0 + \mu\, g_i$. In the absence of imitative peer effects, μ measures the impact of the investment in friendship ties (g_i) on the outcome y_i. In other words, this is an additional structural measure to add to the idiosyncratic heterogeneity of workers θ_i.

Alternative formulation of the model In the utility function (1) each individual i chooses two different effort levels, y_i^0 and z_i. This set-up is not essential. We can in fact obtain similar results by assuming the following utility function with only one type of effort z_i:

$$u_i(\mathbf{z};\mathbf{g}) = \mu g_i z_i - \frac{1}{2} z_i^2 + \phi \sum_{j=1}^{n} g_{ij} z_i z_j. \tag{13}$$

The best reply function for each individual i is clearly given by (7) and, using Proposition 1, we can characterize the Nash equilibrium as follows:

$$z_i^*\left(\mathbf{g}\right) = \frac{\mu}{\phi} b_i\left(\mathbf{g}, \phi\right)$$

In this formulation, $z_i^*\left(\mathbf{g}\right)$ is the optimal effort level provided by each individual i, which is now influenced by the effort provided by her peers and her location in the network. In practice, we are not interested in the effort but in the outcome of effort, i.e. the educational achievement obtained by each student. We assume the educational achievement of each individual i, denoted by $y_i^*\left(\mathbf{x}, \mathbf{g}\right)$, is given by

$$y_i^*\left(\mathbf{x}, \mathbf{g}\right) = \theta_i(\mathbf{x}) + z_i^*\left(\mathbf{g}\right)$$

where $\theta_i(x)$ is defined by (2). This means that the educational achievement of each student i depends on her effort $z_i^*\left(\mathbf{g}\right)$ but also on her and her friends' idiosyncratic characteristics $\theta_i(x)$. Both models lead to the same equation (9).

A more general model of network peer effects In the model presented here, we have used a rather specific expression for network peer effects where only local (strategic) complementarities were present. In Appendix 1, we develop a more general model of network peer effects, including in particular global (strategic) substituabilities. We show (i) how any linear-quadratic utility with peer effects can be transformed into a utility with an explicit social network and can be decomposed into different effects related to the different ex ante heterogeneities, (ii) the existence and uniqueness of a Nash equilibrium (Theorem 1). It turns out that, even though the expression is more complicated, the Nash equilibrium effort decision of each agent is uniquely defined and proportional to her weighted Bonacich network centrality. Observe that the condition that guarantees that the Nash equilibrium exists, is unique, interior and can be written with respect to the Bonacich centrality measure, is not anymore given by $\phi\omega(\mathbf{g}) < 1$ but by a slightly more complicated expression (see part (b) of Theorem 1). Observe also that Theorem 1 in Appendix 1 is a generalization of Theorem 1 in Ballester et al. (2006) for the case of ex ante exogenous heterogeneity. This generalization was crucial to be able to bring the model to the data.

Discussion In this model, the structure of the social network and, in particular, the individual positions in such network, are the main explanatory variables of agents' behavior, together with idiosyncratic heterogeneity. This is the Nash-Katz-Bonacich linkage. In the education literature, for instance, social aspects as well as peer effects have been emphasized as important drivers for individual conduct,[11] but seldom from a network perspective.

The novelty of our model lies precisely on the fact that network structural properties become the cornerstone for understanding the influence of peers on individual behavior. Indeed, when the analysis is restricted to linear-quadratic utility functions, it turns out that the Katz-Bonacich centrality index accounts for peer effects in networks. In the coming sections, we investigate the empirical relevance of this issue. The empirical measure of peer effects, which is derived from our model, thus differs substantially from previous works in this area.[12] Indeed, we are not looking at the impact of group peer effects on individual's educational achievement. Our behavioral model elaborates on the premise that the effects of (exogenous) individual and peer characteristics on individual outcomes can be distinguished from (endogenous) peer effects in networks. The aim of our empirical analysis is to test whether and to what extent heterogeneities in terms of idiosyncratic characteristics and location in the network affect differently the educational outcomes of students. We can therefore compare our results to that of previous studies that only use the first type of heterogeneity (i.e. idiosyncratic characteristics) and we can evaluate the importance of network effects on educational outcomes. In particular, we will be able to determine how the location of each adolescent in a network affects her educational outcomes, once the effects of her characteristics and her friends' characteristics have been accounted for.

3 Data and descriptive evidence

3.1 Data

Our analysis is made possible by the use of a unique database on friendship networks from the National Longitudinal Survey of Adolescent Health (AddHealth).

[11] Akerlof (1997) provides a general discussion on these issues.

[12] See, for example, Topa (2001) for an example of an empirical measure of network effects.

The AddHealth database has been designed to study the impact of the social environment (i.e. friends, family, neighborhood and school) on adolescents' behavior in the United States by collecting data on students in grades 7-12 from a nationally representative sample of roughly 130 private and public schools in years 1994-95. Every pupil attending the sampled schools on the interview-day is asked to compile a brief questionnaire (in-school data) containing questions on respondents' demographic and behavioral characteristics, education, family background and friendship. The AddHealth website describes surveys and data in details.[13] This sample contains information on 90,118 students. In a second phase of the survey, a subset of adolescents selected from the rosters of the sampled schools is then asked to compile a longer questionnaire containing more sensitive individual and household information (in-home and parental data). This sample contains information on 20,745 students.

Friendship networks AddHealth contains unique detailed information on friendship relationships. This information proves crucial for our analysis. The friendship information is based upon actual friend nominations. Pupils were asked to identify their best friends from a school roster (up to five males and five females).[14] Importantly, one can then reconstruct the whole geometric structure of the friendship network. We assume that friendships relationships are reciprocal, i.e. a link exists between two friends if at least one of the two individuals has identified the other as his/her best friend.[15] For each school, we obtain all the networks of (best) friends. By matching the identification numbers of the friendship nominations to respondents' identification numbers, one can obtain information on the characteristics of nominated friends.[16]

Educational achievements The in-home questionnaire contains detailed information on the grade achieved by each student in mathematics, history and social studies and science, ranging from D or lower to A, the highest grade (re-coded 1 to 4). Following the standard approach in the sociological literature to derive quantitative information on a topic using qualitative answers to a battery of related questions, we calculate a school performance index for each respondent.[17] The

[13]http://www.cpc.unc.edu/projects/addhealth

[14]The limit in the number of nominations is not binding, not even by gender. Less than 1 percent of the students in our sample show a list of ten best friends, less than 3 percent a list of five males and roughly 4 percent name five females. On average, they declare to have 5.48 friends with a small dispersion around this mean value (the standard deviation is equal to 1.29). The corresponding figures for male and female friends are 2.78 (with standard deviation equal to 1.85) and 3.76 (with standard deviation equal to 1.04).

[15]This assumption will be relaxed in Section 8.

[16]Note that, when an individual i identifies a best friend j who does not belong to the surveyed schools, the database does not include j in the network of i; it provides no information about j. However, in the large majority of cases (more than 93%), best friends tend to be in the same school and thus are systematically included in the network.

[17]This is a standard factor analysis, where the factor loadings of the different variables are used to derive the total score. The Crombach-α measure is then used to assess the quality of the derived index. In our case, we obtain an

mean is 2.34 and the standard deviation is equal to 2.11.[18]

By merging the in-home data with the in-school friendship nominations data and by excluding the individuals that report a non valid answer to the target questions, we obtain a final sample of 11,964 pupils distributed over 199 networks. Appendix 2 details the information available on the students selected in this sample. Table 1 provides the corresponding descriptive statistics. It reveals that, for instance, the average student is in grade 9, has spent more than 3 years in the school, is fairly motivated in education, with a good relationship with teachers, whose parents have a level of education higher than high school degree and lives in a fairly well kept building. The variables indicating the interaction with friends and parents show a high involvement in friends' relations and a high level of parental care.

[*Insert Table* 1 *here*]

3.2 Descriptive evidence

Figures 2 displays the empirical distribution of the networks in our sample by their size (i.e. the number of network members).[19] It appears that most friendship networks have between 30 and 90 members. The minimum number of friends in a network is 16, while the maximum is 107. The mean and the standard deviation of network size are 60.42 and 24.48, respectively.

α equal to 0.86 ($0 \leq \alpha \leq 1$) indicating that the different items incorporated in the index have considerable internal consistency.

[18] The empirical analysis has also been performed separately for each subject. The qualitative results (i.e. the evidence on the important role of individual position in the network on education outcome) remain unchanged.

[19] The histogram shows on the horizontal axes the percentiles of the empirical distribution of network members corresponding to the percentages 1, 5, 10, 25, 50, 75, 90, 95, 100 and in the vertical axes the number of networks having number of members between the i and $i-1$ percentile.

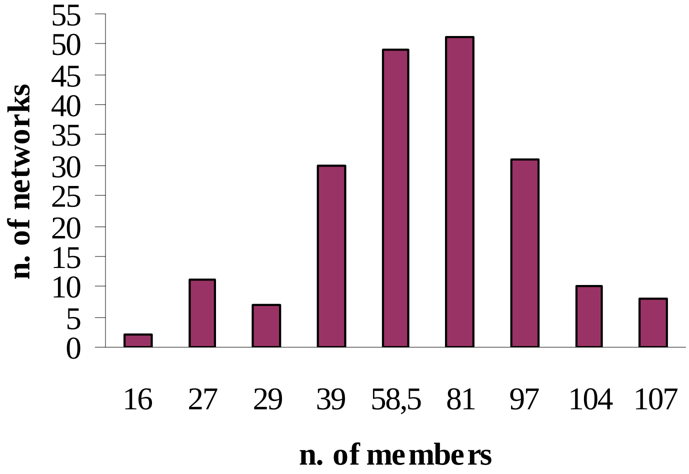


Figure 2. The empirical distribution of adolescent networks

Figure 3 depicts a friendship school network with 16 pupils, which is the smallest networks in our sample. In this network, the most connected student (number 9) has ten direct friends, and the least connected students (numbers 1, 15 and 16) have only one direct friend. Not surprisingly, agent 9 has also the highest Bonacich centrality measure (equal to 3.40) while agents 1, 15 and 16 have the lowest one (equal to 1.28). The largest network in our sample is almost seven times bigger and has 107 members.

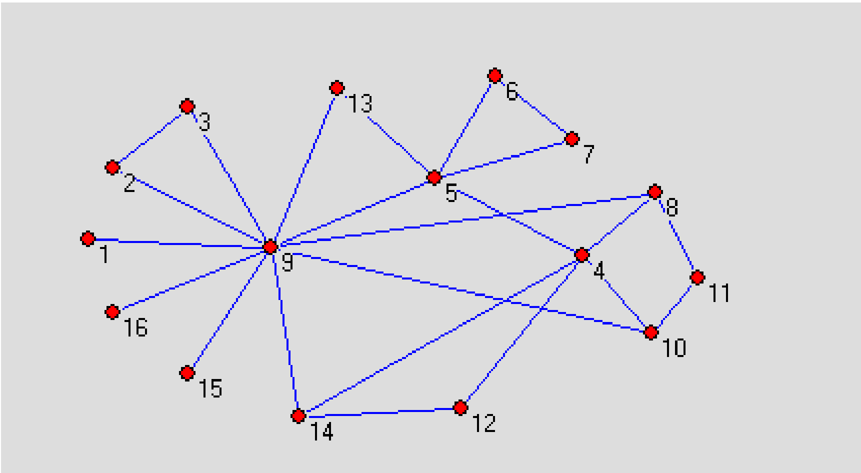


Figure 3. Smallest network of adolescents ($n = 16$)

4 Empirical strategy and identification of peer effects

Guided by Proposition 1, we wish to measure the actual empirical relationship between $b_i(\mathbf{g}, \phi)$ and the observed effort level y_i^*.

4.1 Empirical strategy

Assume that there are K network components in the economy. Network components are maximally connected networks, that satisfy the two following conditions. First, two agents in a network component $\mathbf{g}_\kappa$ are either directly linked, or are indirectly linked through a sequence of agents in $\mathbf{g}_\kappa$ (this is the requirement of connectedness). Second, two agents in different network components $\mathbf{g}_\kappa$ and $\mathbf{g}_{\kappa'}$ cannot be connected through any such sequence (this is maximality). Note that $\sum_{\kappa=1}^{K} n_\kappa = n$.

The empirical counterpart of (6) and (7) is the following model:

$$\begin{aligned} y_{i,\kappa} &= \sum_{m=1}^{M} \beta_m x_{i,\kappa}^m + \frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \gamma_m g_{ij,\kappa} x_{j,\kappa}^m + \eta_\kappa + \varepsilon_{i,\kappa}, \qquad (14) \\ \varepsilon_{i,\kappa} &= \mu g_{i,\kappa} + \phi \sum_{j=1}^{n_\kappa} g_{ij,\kappa} \varepsilon_{j,\kappa} + \upsilon_{i,\kappa}, \;\; i = 1, ..., n; \;\; \kappa = 1, ..., K, \end{aligned}$$

where $y_{i,\kappa}$ is the individual i's level of activity (educational achievement) in the network component $\mathbf{g}_\kappa$, $x_{i,\kappa}^m$ is a set of M control variables accounting for observable differences in individual, neighborhood and school characteristics,[20] $g_{i,\kappa} = \sum_{j=1}^{n_\kappa} g_{ij,\kappa}$ is the number of direct links of i, $\sum_{j=1}^{n_\kappa} \left(g_{ij,\kappa} x_{j,\kappa}^m \right) / g_{i,\kappa}$ is the set of the average values of the M controls of i's direct friends (i.e. contextual effects), and η_κ is an (unobserved) network-specific component (constant over individuals in the same network), which might be correlated with the regressors.

The second equation of (14) describes the process of $\varepsilon_{i,\kappa}$, which is the residual of individual i's level of activity in the network $\mathbf{g}_\kappa$ that is *not* accounted for neither by individual heterogeneity and contextual effects nor by (unobserved) network-specific components. Here, $\sum_{j=1}^{n_\kappa} g_{ij,\kappa} \varepsilon_{i,\kappa}$ is the spatial lag term and ϕ is the spatial autoregressive parameter. Observe that, consistently with the theoretical model, spatial dependence is incorporated in the regression disturbance term only. This

[20] A precise description of all these variables is contained in Appendix 2.

model is a variation of the Anselin (1988) spatial error model.[21,22]

Using the Maximum Likelihood approach (see, e.g. Anselin, 1988), we estimate jointly $\widehat{\boldsymbol{\beta}}$, $\widehat{\boldsymbol{\gamma}}$, $\widehat{\phi}$, $\widehat{\mu}$. These values are then used to measure the relative importance of individual characteristics, $\widehat{\beta}_1, ..., \widehat{\beta}_m$ (e.g. parental education, school and neighborhood quality), contextual effects, $\widehat{\gamma}_1, ..., \widehat{\gamma}_m$ (e.g. average parental education of each individual's best friends, etc.), and the individual Katz-Bonacich centrality index, $\widehat{\phi}$ and $\widehat{\mu}$, in shaping individuals' behavior (equation (9) in Proposition 1). Indeed, our model allows us to decompose additively individual behavior into an idiosyncratic effect and a peer effect (see (8)), which boils down to the individual Katz-Bonacich centrality index.[23]

4.2 Identification of peer effects

The assessment of the effects of peer pressure on individual behavior, i.e. the identification of endogenous social effects, is typically characterized by econometric issues, that render the identification and the measurement of these effects problematic. The crucial (well-known) issues are the endogenous sorting of individuals into groups and the reflection problem (Manski, 1993). Let us explain how we tackle each of them in turn.

---

[21] In matrix notation, we have:

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{DGX}\gamma + \boldsymbol{\eta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \mu\mathbf{G1} + \phi\mathbf{G}\boldsymbol{\varepsilon} + \boldsymbol{\nu}, \end{aligned}$$

where $\mathbf{y}$ is a $n \times 1$ vector of observations on the dependent (decision) variable, $\mathbf{X}$ is a $n \times M$ matrix of observations on the exogenous variables associated to the $M \times 1$ regression coefficient vector $\boldsymbol{\beta}$, $\mathbf{D} = diag\left(1/g_1, ..., 1/g_n\right)$ is a $n \times n$ matrix, $\boldsymbol{\eta}$ is a $n \times n$ diagonal matrix of network fixed effects, with diagonal cells taking the same value within each network component, $\mathbf{1}$ is a $n \times 1$ vector of ones, $\mathbf{G}$ is a $n \times n$ spatial weight matrix that formalizes the network structure of the agents (with elements g_{ij} equal to 0 if $i = j$ or if i and j are not connected, and equal to a constant otherwise), ϕ is the spatial autoregressive parameter, and $\boldsymbol{\nu}$ is a vector of random error terms. This is the fixed-effects panel counterpart of the Anselin (1988) spatial error model where an exogenous variable ($\mathbf{G1}$) has been added in the error process.

[22] Our empirical results remain qualitatively unchanged when working with a row-standardized adjacency matrix, i.e. if we normalize the error term (second equation of (14)) by $g_{i,\kappa} = \sum_{j=1}^{n_\kappa} g_{ij,\kappa}$, the number of direct links of i. Because a row-standardized matrix implies that the largest eigenvalue is 1, we present the analysis using this approach to ease the interpretation of the results. The identification condition (Proposition 2) is also substantially unchanged (see Bramoullé et al., 2006).

[23] To be consistent with the theoretical model, we need to discard networks whose associated ϕ do not satisfy the condition $\phi\omega(\mathbf{g}) < 1$ of Proposition 1, where $\omega(\mathbf{g})$ is the largest eigenvalue of the adjacency matrix associated to network $\mathbf{g}$. This guarantees that the Katz-Bonacich index is well-defined and the uniqueness of the equilibrium as well as the interiority of the solution. For that, we have estimated model (14) for each network $\mathbf{g}$ separately, thus obtaining 199 different estimates of ϕ. We find that only 18 networks fail to satisfy this condition (less than 10% of the total), with a total number of 473 discarded people. We obtain a final sample of 11,491 individuals distributed over 181 networks. Descriptive statistics on this sample do not differ significantly from those on the whole sample (contained in Table 1). We describe in more detail our estimation strategy in Section 4.2 and discuss more extensively the importance of this condition $\phi\omega(g) < 1$ in Section 5.2 below.

The role of network fixed effects In most cases individuals sort into groups non-randomly. For example, kids whose parents are low educated or worse than average in unmeasured ways would be more likely to sort with low human capital peers. If the variables that drive this process of selection are not fully observable, potential correlations between (unobserved) group-specific factors and the target regressors are major sources of bias. The use of network fixed effects, also referred to as correlated effects or network unobserved heterogeneity, proves useful in this respect. Assume, indeed, that agents self-select into different groups in a first step, and that link formation takes place within groups in a second step. Then, as Bramoullé et al. (2006) observe, if link formation is uncorrelated with the observable variables, this two-step model of link formation generates network fixed effects. Assuming additively separable group heterogeneity, a within group specification is able to control for these correlated effects. In other words, we use the model specification (14), which has a network-specific component η_κ of the error term, and adopt a traditional (pseudo) panel data fixed effects estimator, namely, we subtract from the individual-level variables the network average.

Observe that our particularly large information on individual (observed) variables should reasonably explain the process of selection into networks. Then, the inclusion of network fixed effects acts as a further control to account for a possible sorting based on unobservables. Let us now document to what extent such an approach accounts for self-selection into peer-groups. Table 2 reports the estimated correlations between individual and peer-group averages (i.e., averages over best friends) of variables that are commonly believed to induce self-selection into teenagers' friendship groups, once the influence of our extensive set of controls (Appendix 2) and network-fixed effects are washed out. The estimated correlation coefficients reported in Table 2 are not statistically significant for all variables. This indicates that, conditionally on individual and network characteristics, linking decisions are uncorrelated with observable variables.

[*Insert Table* 2 *here*]

The role of peer groups with individual level variation While a network fixed effects estimation allows us to distinguish endogenous effects from correlated effects, it does not necessary estimate the causal effect of peers' influence on individual behavior. A second and more subtle issue has to be tackled. In the standard framework, individuals interact in groups, that is individuals are affected by all others in their group and by none outside the group. As a consequence, in a peer group, everyone's behavior affects the others, so that we cannot distinguish if a group member's action is the cause or the effect of peers' influence, which is the well-known reflection problem (Manski, 1993). In our network framework, instead, the reference group is the number of friends each individual has and groups do overlap. Because peer groups are individual specific, this issue is eluded. Let us be more precise. The reduced-form equation corresponding to the spatial error term in (14) is, in matrix notation:

$$\boldsymbol{\varepsilon} = \mu \left(\mathbf{I} - \phi \mathbf{G}\right)^{-1} \mathbf{G1} + \left(\mathbf{I} - \phi \mathbf{G}\right)^{-1} \boldsymbol{\nu}. \tag{15}$$

We say that peer effects are identified if the structural parameters (μ, ϕ) uniquely determine the reduced-form coefficients in (15).

Bramoullé et al. (2006) provide general results on the identification of peer effects through social networks via variations of the linear-in-means model (see, also, Laschever, 2005, and Lin, 2005). Using a similar approach, we show that identification is granted in our model under a mild condition that involves the structure of one link and two-links away network contacts. Recall that $\mathbf{G1} = [g_i]$ is the vector of node connectivities, while $\mathbf{G}^2\mathbf{1} = \left[g_i^{[2]}\right]$ gives the total number of two-link away contacts in the network.[24] In particular, $g_i^{[2]}/g_i$ is the average connectivity of agent i's direct contacts.[25]

Proposition 2 *Suppose that $\phi\omega(\mathbf{g}) < 1$ and $\mu \neq 0$. Peer effects are identified if and only if $g_i^{[2]}/g_i \neq g_j^{[2]}/g_j$ for at least two agents i and j.*

Proof. See Appendix 3. ▮

In words, peer effects are identified if we can find two agents in the economy that differ in the average connectivity of their direct friends. This is a simple property of the network, that amounts to checking that the $2 \times n$ matrices with column vectors $\mathbf{G1}$ and $\mathbf{G}^2\mathbf{1}$ are of rank two. Note also that the condition $\mu \neq 0$ is very natural in this setting because otherwise there are no peer effects at all (see equation (15)). Take for example the network described in Figure 1. Then, it is easy to verify that: $1 = g_1^{[2]}/g_1 \neq g_2^{[2]}/g_2 = 1/2$.

Although not very demanding, this condition still rules out some network architectures, as it requires a minimum level of heterogeneity in the network connectivities. As an extreme case, consider a regular network, where all agents have the same number of links, say r. Formally, $\mathbf{G1} = r\mathbf{1}$. Then, it is readily checked that $\mathbf{G}^2\mathbf{1} = r^2\mathbf{1}$, and so $g_i^{[2]}/g_i = r$, for all i. Identification fails in this case.[26]

Consequently, this identification result imposes that there is some irregularities in the network so that at least two individuals do not have the same number of links. In general, in the real-world and in particular in our data, no network is regular and the identification requirement is always satisfied. Indeed, peer-groups are individual specific and individuals belong to more than one group.

[24] This is only true if a contact is counted (potentially) multiple times.

[25] Indeed,

$$\frac{g_i^{[2]}}{g_i} = \frac{1}{g_i}\sum_{j=1}^{n} g_{ij}^{[2]} = \frac{1}{g_i}\sum_{j=1}^{n}\sum_{k=1}^{n} g_{ik}g_{kj} = \frac{1}{g_i}\sum_{k=1}^{n} g_{ik}\sum_{j=1}^{n} g_{kj} = \frac{1}{g_i}\sum_{k=1}^{n} g_{ik}g_{k}.$$

[26] Note that our identification condition is weaker than the one in Bramoullé et al. (2006), where contextual effects and endogenous effects can be separated from each other in a linear-in-local-means model provided that the matrices $\mathbf{I}$, $\mathbf{G}$, and $\mathbf{G}^2$ are independent from each other

The role of specific controls Finally, the richness of the information provided by the AddHealth questionnaire on adolescents' behavior allows us to find proxies for typically unobserved individual characteristics that may be correlated with our variable of interest. For example one might argue that more self-confident and (very likely) more successful students at school are contacted by a larger number of friends, thus showing a higher value of the Katz-Bonacich measure. Therefore, we deal with unobservable individual characteristics correlated with the Katz-Bonacich measure that may cause education outcomes not directly caused by the centrality measure. To control for differences in leadership propensity across adolescents, we include an indicator of self-esteem because more successful students are likely to consider themselves as more intelligent than their peers, and an indicator of the level of physical development compared to the peers. Also, we attempt to capture differences in attitude towards education and parenting by including indicators of the student's motivation in education and parental care.

Similar arguments can be put forward for the existence of possible correlations between our centrality measure and unobservable school characteristics affecting structure and/or quality of school-friendship networks in analyzing students' school performance. Because the AddHealth survey interviews all children within a school, we estimate our model conditional on school fixed effects (i.e. we incorporate in the estimation *school dummies*). This approach enables us to capture the influence of school level inputs (such as teachers and students quality and possibly the parents' residential choices), so that only the variation in the Katz-Bonacich measure (across students in the same school) would be exploited.[27]

Estimation strategy Let us summarize our estimation strategy, i.e. how we discard the networks and how we obtain our (average) estimate of peer effects ϕ (and μ). First, we estimate our empirical model defined by equation (14) for each network in our data set, i.e. we only exploit the cross-sectional variation in each network (*within-network variation*), thus obtaining 199 different estimated values of ϕ (and μ), that we denote by $\widehat{\phi}_\kappa$ for each network $\kappa = 1, ..., 199$. Second, for each network g_κ, we calculate its largest eigenvalue $\omega(g_\kappa)$ and check which network does not satisfy the condition $\phi_\kappa < 1/\omega(g_\kappa)$. We then stack the networks that satisfy this condition and estimate model (14) by running a pseudo-panel data estimation (i.e. using both *within and between-network variations*), which allows us to control for network-specific unobservable factors (captured by the term η_κ), thus obtaining an average estimate of ϕ and μ in our dataset. From those estimated of μ and ϕ, we can finally calculate the Katz-Bonacich centrality for each individual and look at the relevance of network position in shaping individual outcomes.

[27]The introduction of student-grade or student-year of attendance dummies does not change qualitatively the results on our target variable.

5 Empirical results and discussion

Let us now test our theoretical model by investigating whether our model-driven measure of peer effects, namely the Katz-Bonacich network centrality index, matters in explaining individual outcomes.

5.1 Empirical results

We start by estimating a traditional regression model where the individual school performance is explained as a function of a set of observable individual characteristics (first equation of model (14)). Although the set of explanatory variables included is wider that the one typically used in the estimation of an education production function (see Appendix 2), the standard OLS results with diagnostics for spatial effects show that (i) there is still a substantial part of the variance that is not explained and (ii) there is a strong evidence of spatial correlation in the residuals. Our model claims that the position and the peer effects of links in a network are important factors, thus providing an economic behavioral foundation for the estimation of a spatial model. Indeed, we find that the spatial error model (14) derived from our theoretical set up is not rejected by the data.[28]

Model (14) is estimated using the Maximum Likelihood approach. Different sets of controls have been used (see Appendix 2). We start by including standard individuals' characteristics and behavioral factors (i.e., socio-demografic factors, family background, motivation in education and a proxy for individual ability, namely mathematics score). Then, we gradually introduce protective factors (i.e., relationship with teachers, social exclusion, school attachment, friends attachment, parental care) and residential neighborhood characteristics. The corresponding average characteristics of direct friends aiming at capturing the quality of social interactions are included in all specifications (these variables are referred to as contextual effects). Finally, we also attempt to control for unobservable individual and school characteristics that may be correlated with our variable of interest by adding a proxy of self-esteem, an indicator of the level of physical development compared to the peers and school dummies.

The Maximum Likelihood estimation results for the model specification that includes the complete set of controls are reported in Table 3.[29] The estimated μ and ϕ are both positive and highly statistically significant. We then calculate the Katz-Bonacich measure (expression (4)) by fixing the value of ϕ at the point estimate $\widehat{\phi}$. The derived Katz-Bonacich measures range from 0.32 to 3.48, with an average of 1.65 and a standard deviation of 2.79. The estimated impact of this

[28] A variety of measures of statistical performance show that model (14) improves the statistical fit to the data with respect to the one where only the first equation of model (14) is considered. Standard hypotheses tests provide evidence that model (14) is appropriate and correctly specified. The details of such a statistical analysis are available upon request.

[29] The estimated effects of the control variables are qualitatively the same across all model specifications and in line with the expectations. The estimation results for all the model specifications are available upon request.

variable on education outcomes that is predicted by the theory, i.e. $\widehat{\mu}/\widehat{\phi}$ (equation (9) in Proposition 1) is statistically significant[30] and non negligible in magnitude. Specifically, we find that a one-standard deviation increase in the Katz-Bonacich index translates into roughly 7 percent of a standard deviation in education outcome, whereas for instance this effect is about 17 percent for parental education (which is higher, but comparable).[31]

[*Insert Table* 3 *here*]

5.2 Discussion

Let us be more explicit about the meaning and the role of the condition $\phi\omega(g) < 1$ in Proposition 1. Remember that such a condition is needed both for characterizing the Nash equilibrium in terms of the Katz-Bonacich centrality measure and for the existence and uniqueness of the equilibrium as well as for the interiority of the solution. This is quite appealing since only one condition guarantees all these results. Let us now give some more intuition. Condition $\phi\omega(g) < 1$ requires that the parameter for own-concavity 1 (i.e. $\left|\frac{\partial^2 u_i}{\partial z_i^2}\right|$) is high enough to counter the payoff complementarity, measured by $\phi\omega(g)$. The scalar ϕ measures the level of positive cross effects (i.e. $\left|\frac{\partial^2 u_i}{\partial z_i \partial z_j}\right|$), while $\omega(g)$ captures the population-wide pattern of these positive cross effects.[32] In other words, $\phi\omega(\mathbf{G})$ accounts both for the size of complementarities, captured by ϕ, and for their pattern, captured by $\omega(\mathbf{G})$. Indeed, for a fixed number of agents and links, the largest eigenvalue of a network is a measure of its *regularity*. A higher eigenvalue corresponds to an irregular hub-like structure, whereas a lower eigenvalue (still for the same number of agents and links) refers to a more regular network. Observe that the eigenvalue condition $\phi\omega(g) < 1$ does not bound the *absolute* values for these cross effects, but only their *relative* magnitude. The correlations between any two units (i.e., students' complementarities) can be as large as one when $\phi\omega(g) < 1$. Roughly speaking, this condition only ensures that the effect of the friends of friends is lower than the one of direct friends. In other words, it guarantees that the weights ϕs in (4) are decreasing with path length.

When the condition $\phi\omega(g) < 1$ does not hold, then there is no upper bound on network complementarities and therefore the positive feed-back loops in the network of complementarities can

[30] Because we deal with a non-linear transformation, the standard error is calculated using the *deltha method*. The associated $t-$test value is equal to 2.11, which denotes statistical significance at the 5% level.

[31] This analysis has also been performed by gender. The point estimates of our target parameters, $\widehat{\mu}$ and $\widehat{\phi}$, are very similar between males and females and not statistically different. These results are available upon request.

[32] A standard condition for existence and uniqueness in the economics literature of interactions (especially the Cournot model; see e.g. Kolstad and Mathiesen, 1987) is diagonal dominance which, in its simplest version, requires that:

$$\sum_{j \neq i} \left| \frac{\partial^2 u_i}{\partial z_i \partial z_j} \right| < \frac{\partial^2 u_i}{\partial z_i^2} \text{ , for all } i = 1, ..., n$$

Diagonal dominance is a stronger requirement than our condition $\phi\omega(g) < 1$.

trigger an unbounded escalation of efforts, and never reach an equilibrium level. To be more precise, from a theoretical viewpoint, when $\phi\omega(g) < 1$ does not hold:

(i) We cannot characterize the Nash equilibrium since the Katz-Bonacich centrality measure is not anymore defined, i.e. the infinite sum $\sum_{k=0}^{+\infty} \phi^k \mathbf{G}^k$ does not converge to a finite value and therefore cannot be equal to $(\mathbf{I} - \phi\mathbf{G})^{-1}$;

(ii) The existence of equilibrium becomes an issue since the strategy space is unbounded. In that case, our game with complementarities becomes related to the literature on supermodular games (see Topkis, 1979, and Vives, 2005). In this literature, to obtain existence of equilibrium, the strategy space of a supermodular game is a bounded lattice. In our case, to obtain existence, we need to bound the strategy space in some arbitrary way but this is obviously not satisfactory.[33]

(iii) Even if existence is guaranteed, uniqueness does not always follows. In fact, as it is well known in the literature on supermodular games, multiple equilibria are rather the rule than the exception.

From an empirical point of view, we could relax the condition $\phi\omega(g) < 1$ and introduce another one that bounds the strategy space, so that the existence of equilibrium can be guaranteed. Let us therefore relax this condition on the eigenvalue and not restrict the analysis when complementarities are small. Let us bound the strategy space in such a game rather naturally by simply acknowledging the fact that students have a time constraint and allocate their time between leisure and school work. In that case, multiple equilibria will certainly emerge, which is a plausible outcome in the school setting. From a theoretical viewpoint, Proposition 1 is not valid anymore and only equations (6) and (7) will be considered. The empirical model is exactly the same (and defined by (14)) with the only difference that we now run the regression on all the 199 networks and not on 181 networks (i.e. we do not discard the networks that do not satisfy the eigenvalue condition). When we run such a regression, we obtain that both point estimates and standard errors of both μ and ϕ are very similar to the ones reported in Table 3. The estimates remain highly significant (at the same level of statistical significance) showing point estimates only slightly lower in magnitude.[34]

To conclude, the condition on eigenvalue $\phi\omega(g) < 1$ is not that demanding since it does not change very much the results of the empirical analysis. First, only 9 percent of the networks do not satisfy this condition, which means that we switch from a sample of 11,964 pupils distributed over 199 networks to a sample of 11,491 individuals distributed over 181 networks. Second, when we compare the descriptive statistics of the sample before (with 199 networks) and after (with 181 networks), we find they do not differ significantly. This indicates that there is nothing unique

[33]Imposing an arbitrary bound on a strategy space need not be an innocuous modelling choice. Indeed, while this arbitrary bound solves equilibrium existence concerns, when the resulting lattice of equilibria does not boil down to a single outcome, the structure of this equilibrium lattice turns out to depend critically on the arbitrary choice of this upper bound.

[34]The estimated values of μ and ϕ are now respectively given by 0.0301 (with a standard error of 0.0140) and 0.5352 (with a standard error of 0.1366).

about the discarded individuals/networks. Finally, when we estimate our model with 199 networks and with 181 networks, the estimation results are very close, with point estimates having slightly lower values in the former case. In this case, however, we cannot interpret our results in terms of Bonacich centrality measures, thus losing the interesting result of our paper about the impact of network topology on individual outcomes.

6 Alternative formulations

In Section 2, we have developed a theoretical network model and then test it in Section 5. Our model is not rejected by the data. However, the topic is so complex that such a model does not cover all the possible issues at stake. In this section, we would like to highlight the limitations of our analysis and propose some extensions.

One concern with our model is that we arbitrarily separate idiosyncratic effects from peer effects (see equation (8)) and assume that there are two different efforts y_i and z_i for each of these effects. This is obviously a strong assumption. In Appendix 4, we develop an alternative network model with individuals' heterogeneity and peer effects where these two components are not separated and only one effort is considered. It turns out that the equilibrium effort of each individual i is equal to her *weighted* Katz-Bonacich centrality index (see (29)), where the weights are her idiosyncratic characteristics and that of her friends. This approach implies that in practice, since the idiosyncratic heterogeneity θ_i of individual i is multidimensional (see Appendix 2, when θ_i can be defined by the gender, race, age, education of parents, etc.), there are as many Katz-Bonacich centrality indices as idiosyncratic characteristics θ_i. Furthermore, testing the equation derived from this model (i.e. equation (30)) will not allow us to distinguish between the impact of network location and that of individual idiosyncratic characteristics on individual educational achievement.

An other potential problem is that we assume that both (individual and peer-oriented) efforts contribute *positively* to an individual's total school effort and thus enhance an individual's educational outcome. It is possible that a peer-oriented effort can detract from school outcomes, that is, "doing homework with friends" is actually not productive school effort and can lead to lower total effort than doing homework alone and therefore can lead to lower educational attainment. In that case, equation (8) should be replaced by:

$$y_i^*(\mathbf{x},\mathbf{g}) = \underbrace{y_i^{0*}(\mathbf{x})}_{\text{idiosyncratic}} - \underbrace{z_i^*(\mathbf{g})}_{\text{peer effect}}$$

where we "subtract" rather than "add" the two effort terms. In this formulation, $y_i^{0*}(\mathbf{x}) > y_i^*(\mathbf{x},\mathbf{g})$ so that doing effort alone leads to a higher outcome than doing effort alone and with peers. It should be clear that, in that case, the theoretical analysis is qualitatively the same where the individual equilibrium outcome is uniquely defined by: $y_i^*(\mathbf{x},\mathbf{g}) = \theta_i(\mathbf{x}) - \frac{\mu}{\phi} b_i(\mathbf{g},\phi)$ and not anymore by (9).

If we keep the assumption of local complementarities, then the difference between the two models lies on the sign of μ. It is thus an empirical matter whether μ is positive or negative. When we perform a panel estimation of our empirical model (14) and obtain an average estimate of μ, we find that the sign is positive (value reported in Table 3). This indicates that, on average, peer effects increase rather than decrease total outcome. However, when running our empirical model on each of the networks separately, thus obtaining 199 different estimates of μ, we find that, in a few cases (i.e. 6 percent of the networks), the estimate of μ is negative. These are cases were, indeed, "peer-oriented effort can detract from school outcomes".

A last potential problem with our analysis is that it is assumed that a student's idiosyncratic characteristics is not related to her peer effect outcomes. For example, suppose a student works hard at mathematics. This will have a direct effect on the individual's outcome and an effect from and to peers. Moreover, the effect on peers could depend on the individual's idiosyncratic efforts. The effort of a student from the bottom of the socio-economic ladder might have more effect than the effort of someone at the top. In order to address these issues, two alternative models can be proposed. First, we can change the utility function (1) in the following way:[35]

$$u_i(\mathbf{y}^0, \mathbf{z}; \mathbf{g}) = \theta_i \, y_i^0 - \frac{1}{2}\left(y_i^0\right)^2 + \mu \sum_{j=1}^{n} g_{ij}\theta_j z_i - \frac{1}{2} z_i^2 + \phi \sum_{j=1}^{n} g_{ij} z_i \theta_j z_j,$$

so that i's *friends' characteristics* θ_j directly affect peer effect outcomes but keep the separability equation (8). We can solve this model and estimate it, and as for the model in Appendix 4, we obtain a Katz-Bonacich index that accounts for both idiosyncratic characteristics and location in the network,[36] and it is not possible to disentangle between these two effects. Second, we can keep the same utility function (1) but relax the separability equation (8) and assume instead:

$$y_i^* \left(\mathbf{x}, \mathbf{g}\right) = y_i^{0*} \left(\mathbf{x}\right) \; z_i^* \left(\mathbf{g}\right)$$

We can once more solve the model and estimate it.[37] We obtained a weighted Katz-Bonacich and it is again not possible to evaluate the impact of peer effects stemming from network topology and the one arising from individual idiosyncratic characteristics.

To conclude, when we abandon the separability assumption, then, in any model, it will difficult to separate peer effects due to the individual location in the network (i.e. the "pure" Katz-Bonacich index) from the ones arising from individual idiosyncratic characteristics. This highlights the fact that the impact of a student's effort on his/her outcomes due to peer effects is not so easily separated or identified. In this paper, we propose one way to do it but this has obviously its limitations.

[35] We could even have a more general utility function with the last term being $\phi \sum_{j=1}^{n} g_{ij}\theta_i z_i \theta_j z_j$, so that both *own* and friend characteristics directly affect peer effect outcomes. This will, however, complicate the analysis without changing the main results.

[36] See the "Second Alternative Analysis" in Appendix 4 for the complete analysis.

[37] See the "Third Alternative Analysis" in Appendix 4 for the complete analysis.

7 Alternative measures of unit centrality

Our model of social network interactions puts forward the Katz-Bonacich centrality as the relevant network measure to account for peer effect outcomes. Over the past years, social network theorists have proposed a number of centrality measures to account for the variability in network location across agents (Wasserman and Faust, 1994). Roughly, these indices encompass two dimensions of centrality, connectivity (or closeness) and betweenness. The simplest index of connectivity is the number of direct links stemming from each node in the network. Instead, betweenness indexes derive from the number of optimal paths across (or from) every node.[38]

While these measures are mainly geometric in nature, our theory provides a behavioral foundation to the Katz-Bonacich centrality measure (and only this one) that coincides with the unique Nash equilibrium of a non-cooperative peer effects game on a social network. The Katz-Bonacich centrality is an index of connectivity since it counts the number of *any* path stemming from a given node, not just the optimal paths. Our main finding in Section 5 is that the Katz-Bonacich index exerts a statistically significant and non negligible in magnitude effect on individual behavior: a one standard deviation increase in the Katz-Bonacich index translates into roughly a 7 percent of a standard deviation increase in educational outcome. What is the comparable figure for other standard individual centrality measures as opposed to the Katz-Bonacich centrality?

In this section, we test the explanatory value of the three most widely used centrality measures: degree, closeness and betweenness centralities.

Degree centrality The individual-level *degree centrality* is simply each individual's number of direct friends:

$$\delta_i(\mathbf{g}_\kappa) = g_i = \sum_{j=1}^{n} g_{ij}$$

To compare networks of different sizes, this measure is normalized to be in an interval from 0 to 1, where values 0 and 1 indicate the smallest and the highest possible centrality. To do so, we divide $\delta_i(\mathbf{g}_\kappa)$ by the maximum possible number of friends individual i can have (i.e. $n_\kappa - 1$ individuals in the network component κ if everyone is directly connected to individual i), and express the result as a proportion:

$$\delta_i^*(\mathbf{g}_\kappa) = \frac{g_i}{n_\kappa - 1} = \frac{\sum_{j=1}^{n} g_{ij}}{n_\kappa - 1}$$

In our data, the normalized degree centrality index $\delta_i^*(\mathbf{g}_\kappa)$ has a mean equal to 0.35 and a standard deviation equal to 0.18.

[38] See Borgatti (2003) for a discussion on the lack of a systematic criterium to pick up the "right" network centrality measure for each particular situation.

Closeness centrality The standard measure of *closeness* centrality of individual i is given by:

$$c_i(\mathbf{g}_\kappa) = \frac{1}{\sum_j d_{ij,\kappa}}$$

where d_{ij} is the geodesic distance (length of the shortest path)[39] between individuals i and j. As a result, the closeness centrality of individual i is the inverse of the sum of geodesic distances from i to the $n-1$ other individuals (i.e. the reciprocal of its "farness"). Compared to degree centrality, the closeness measure takes into account not only direct connections among individuals but also indirect connections. However, compared to the Katz-Bonacich centrality, the closeness measure assumes a weight of one to each indirect connection, whereas the Katz-Bonacich centrality uses weights that depend on the strength of social interaction within the network.

Again, in order to be able to compare different networks, we can define a relative closeness centrality measure as:

$$c_i^*(\mathbf{g}_\kappa) = \frac{n_\kappa - 1}{\sum_j d_{ij,\kappa}}$$

where $n_\kappa - 1$ is the maximum possible distance between two individuals in network component κ. This measure takes values between 0 and 1. Its highest value is when individual i is directly connected to all other $n_\kappa - 1$ individuals in network component κ since, in that case, $\sum_j d_{ij,\kappa} = n_\kappa - 1$ and therefore $c_i^*(\mathbf{g}_\kappa) = 1$. The mean and the standard deviation in our data of this normalized index are 0.49 and 0.27.

Betweenness centrality Freeman (1978/79) defines the betweenness centrality measure of agent i in a network component $\mathbf{g}_\kappa$ in the following way:

$$f_i(\mathbf{g}_\kappa) = \sum_{j<l} \frac{\text{number of shortest paths between } j \text{ and } l \text{ through } i \text{ in } \mathbf{g}_\kappa}{\text{number of shortest paths between } j \text{ and } l \text{ in } \mathbf{g}_\kappa}$$

where j and l denote two given agents in $\mathbf{g}_\kappa$. A normalized version of this measure is:

$$f_i^*(\mathbf{g}_\kappa) = \frac{f_i(\mathbf{g}_\kappa)}{(n_\kappa - 1)(n_\kappa - 2)/2},$$

where n_κ is the size of the network $\mathbf{g}_\kappa$. Note that betweenness centrality, as the degree and closeness centrality measures, is a parameter-free index. In our data, the normalized betweenness measure $f_i^*(\mathbf{g}_\kappa)$ has a mean equal to 0.45 and a standard deviation equal to 0.51.

Table 4 reports the estimation results obtained when using these alternative centrality measures as an additional explanatory variable in a OLS regression of individual outcomes on our set of observable individual characteristics, contextual and network-specific effects.

[39] The length of a shortest path is the smallest k such that there is at least one path of length k from i to j. Therefore we can find the length by computing $\mathbf{G}$, $\mathbf{G}^2$, $\mathbf{G}^3$, ..., until we find the first k such that the (i, j)th entry of $\mathbf{G}^k$ is not zero.

These results contrast with the important role played by the Katz-Bonacich centrality index (Table 3). Indeed, out of the three measures, only degree centrality shows a statistically slightly significant impact (i.e., at the 10% significance level). Furthermore, when such an effect is translated in terms of standard deviations, its impact on educational outcomes is not even one third of the effect exerted by the Katz-Bonacich centrality index (roughly 2.1 percent versus 7 percent). If the Katz-Bonacich measure is seen as a natural extension of the degree centrality index, this finding is quite intuitive since one expects the Katz-Bonacich measure to better capture network externalities and subtleties in network structure that a local measure like degree necessarily ignores. However, this simple reasoning does not explain why closeness centrality (which accounts also for indirect connections) and betweenness centrality do not show any significant effect.

There are in fact two main explanations for the discrepancy in the explanatory power of the different centrality measures used above (i.e. degree, betweenness and closeness centralities) versus the Katz-Bonacich centrality.

The first reason is that the unique Nash equilibrium of a peer effects game with linear-quadratic utility functions is exactly described by the Katz-Bonacich centrality network measure. The Katz-Bonacich centrality is therefore not an arbitrary network measure that tries to describe the structural role of network positioning on individual behavior in the presence of local complementarities. Rather, it results from a positive analysis that maps network topology to equilibrium behavior. Instead, all the other centrality measures are, to our knowledge, just an ad hoc choice for a network measure that tries to grasp how topology shapes behavior, with no a priori connection with the sort of complementarities in decisions characteristic of peer effects.

The second reason is that all these centrality measures are parameter-free network indices. They only depend on the network geometry. Instead, Katz-Bonacich centrality depends both on the network topology and on the prevailing peer effect strength inside the group.

[*Insert Table* 4 *here*]

8 Directed networks

So far, we have only considered *undirected* networks, i.e. we have assumed that friendship relationships are reciprocal, $g_{ij,\kappa} = g_{ji,\kappa}$. Our data, however, make it possible to know exactly who nominates whom in a network.[40] Indeed, 14 percent of relationships in our dataset are not reciprocal.

In order to see how robust is our analysis, we now exploit the directed nature of the network data. Of course, the interpretation of centrality is now different since centrality contributions only

[40] See, for example, Galeotti and Muller (2007), who use the directed nature of friendship relationships to study the impact of friends in the final-year class of high school on subsequent labor market success.

flow in one direction on the directed links. We would like to see if our results change significantly under such a specification.

We follow the approach of Wasserman and Faust (1994, pages 205-210) who define the Katz-Bonacich centrality measure for directed networks. As they put it: "Centrality indices for directional relations generally focus on choices made, while prestige indices generally examine choices received, both direct and indirect". Since the Katz-Bonacich index falls into the category of prestige indices (see pages 206-208), the links in the directed network are defined by the choices received, i.e. when someone is nominated as a friend by someone else rather than when he/she nominates a friend. Indeed, in our data the correlation of each agent's *received* nominations with our indicators of leadership propensity is quite high, although it is far from perfect. Specifically, it amounts to 0.71 and to 0.75 for our indicators of self-esteem and level of physical development compared to the peers.

In the language of graph theory, in a directed graph, a link has two distinct ends: a head (the end with an arrow) and a tail. Each end is counted separately. The sum of head endpoints count toward the *indegree* and the sum of tail endpoints count toward the *outdegree*. Formally, we denote a link from i to j as $g_{ij} = 1$ if j has nominated i as his/her friend, and $g_{ij} = 0$, otherwise. The indegree of student i, denoted by g_i^+, is the number of nominations student i receives from other students, that is $g_i^+ = \sum_j g_{ij}$. The outdegree of student i, denoted by g_i^-, is the number of friends student i nominates, that is $g_i^- = \sum_j g_{ji}$. We consider only the indegree to define the Katz-Bonacich centrality measure. Observe that, by definition, the adjacency matrix $\mathbf{G} = [g_{ij}]$ is now *asymmetric*.

From a theoretical point of view, it is easily verified that in the proof of Theorem 1 (and thus of Proposition 1), the symmetry of G does not play any explicit role. One only needs to define $\omega(g)$ as the spectral radius of the adjacency matrix G and not as the largest eigenvalue, since the eigenvalues of an asymmetric matrix are in general complex and cannot be ordered by magnitude. Using such a definition, Proposition 1 holds true, that is the Nash-Katz-Bonacich linkage holds for any matrix G, under the condition $\phi\omega(g) < 1$. Our theoretical analysis is therefore totally unchanged. We can define the Katz-Bonacich centrality measure $b(g, \phi)$ exactly as in (4).

Turning to the empirical analysis, the last column in Table 3 reports the results of the estimation of model (14) when the directed nature of the network data is taken into account (i.e., with this alternative specification of $\mathbf{G}$). When the estimated effect of the Katz-Bonacich measure is calculated, we find that it is still statistically significant, at an higher significance level (1 percent versus 5 percent), and only slightly lower in magnitude (5.6 percent versus 7 percent). Therefore the results do not change substantially.

9 Peer effects and network structure

To conclude this paper, we would like to present some findings on the relationship between peer effects and the network topology. For that, we estimate model (14) for each network $\mathbf{g}_\kappa$ separately, thus using within network variation only. We obtain $K = 181$ different estimates of ϕ, $\widehat{\phi}_1, ..., \widehat{\phi}_K$. The estimated value $\widehat{\phi}_\kappa$ measures the strength of each existing bilateral influence in the network $\mathbf{g}_\kappa$. It turns out that all the $\widehat{\phi}_k$s are strictly positive, thus supporting our theoretical assumption (3) about strategic complementarity in individual efforts. These differences are partly driven by the structural differences across such networks, as we examine below.

Figures 4a-4c plot three different structural network measures against the estimated $\widehat{\phi}_k$s (with 5% confidence bands).[41] These measures are density (Figure 4a), asymmetry (Figure 4b) and redundancy (Figure 4c). In Figure 4b and 4c, the $\widehat{\phi}_k$s are divided by the network density.

Network density is simply the fraction of ties present in a network over all possible ones. It ranges from 0 to 1 as networks get denser. Network asymmetry is measured using the variance of connectivities. We normalize it, so that it reaches 1 for the most asymmetric network in the sample. Network redundancy or clustering is the fraction of all transitive triads[42] over the total number of triads. It measures the probability with which two of i's friends know each other. Redundancy, or clustering, is much higher in social networks than in randomly generated graphs.[43] Again, we normalize it.

[41] The $\widehat{\phi}_k$s are between 0 and 1, divided in ten intervals and averaged over each interval. The mean values in each interval are displayed on the horizontal axes, while the average structural properties of the corresponding networks are reported on the vertical axes. The confidence bands are based on the derived standard errors of the average estimated levels of $\widehat{\phi}_k$s in each interval, assuming independency of the $\widehat{\phi}_k$s across the different networks.

[42] A triad is the subgraph on three individuals, so that when studying triads, one has to consider the threesome of individuals and all the links between them. A triad involving individuals i, j, k is transitive if whenever $i \rightarrow j$ and $j \rightarrow k$, then $i \rightarrow k$.

[43] See Jackson and Rogers (2007) for more details.

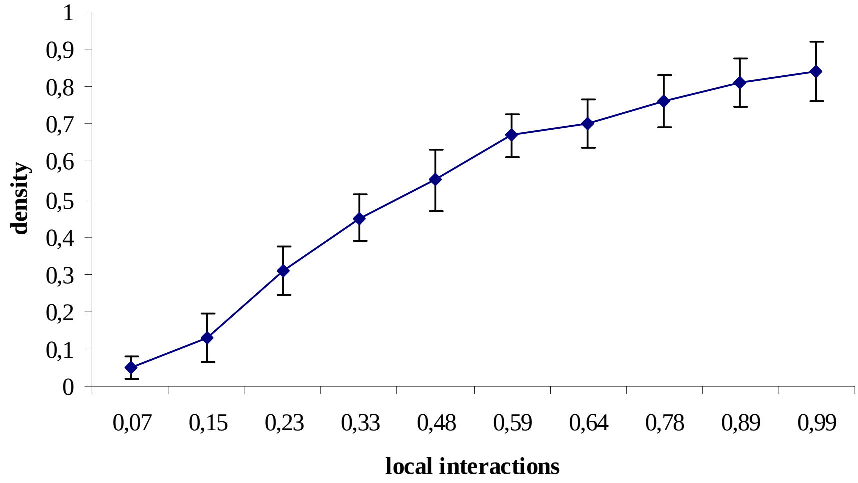


Figure 4a. Density in education networks

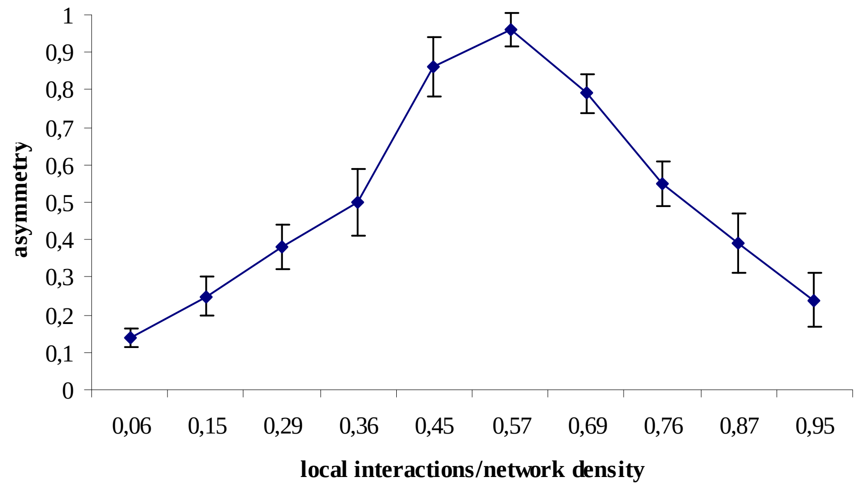


Figure 4b. Asymmetry in education networks

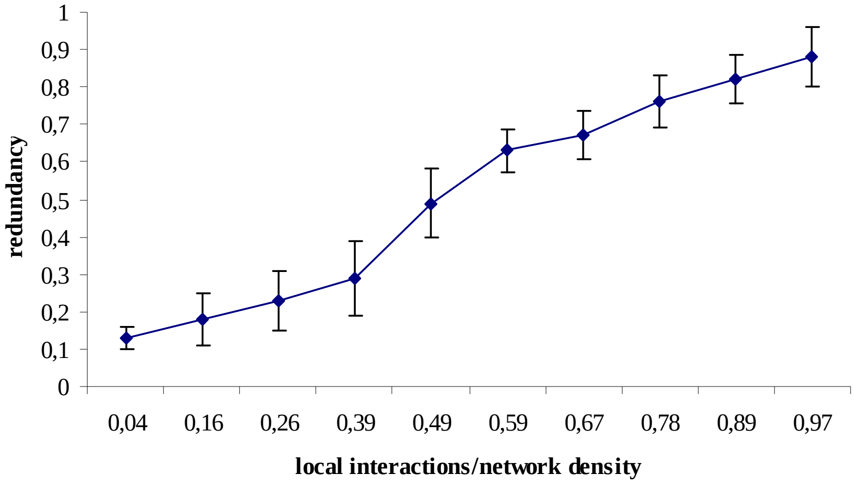


Figure 4c. Redundancy in education networks

Figure 4a shows that the strength of bilateral influences increases steadily with network density for low values, and remains roughly unchanged for higher values. Therefore, richer networks are a sign of stronger dyadic cross effects, at least until roughly 60% of all possible networks links are created. Figure 4b shows that network asymmetry has a non-trivial impact on the intensity of peer effects. Highly distributed and symmetric networks are compatible with both very low and very high values of the peer-to-density ratio, while highly centralized and asymmetric networks are always synonymous of an average value of peer effects. Finally, Figure 4c shows that link redundancy, or clustering, has a strong positive impact on the strength of bilateral influences above a minimum threshold value.

Altogether, these figures suggest that peer effects are strong in moderately dense networks displaying a highly skewed connectivity distribution and a high level of clustering. This is, in fact, the footprint of most real-life large scale social networks (Jackson and Rogers, 2007). Peer effects can also be strong in dense and distributed networks with high clustering. Instead, peer effects are always low in sparse and distributed networks with low clustering. High clustering, therefore, is a necessary condition for strong peer effects.

10 Concluding remarks

In this paper, we propose a model that studies the impact of network location on educational outcomes. Our main theoretical result establishes that the peer effects game has a unique Nash equilibrium where each agent strategy is proportional to her Katz-Bonacich centrality measure. The Katz-Bonacich network centrality counts, for each node in a given network, the total number of direct and indirect paths of any length in the network stemming from that node. Paths are weighted by a geometrically decaying factor (with path length).

We then test the predictions of our model by using a very detailed and unique dataset of friendship networks from the National Longitudinal Survey of Adolescent Health (Add Health). We explore the role of network location for peer effects in education. We first characterize the exact conditions on the geometry of the peer network, so that the model is fully identified. We then show that, after controlling for observable individual characteristics and unobservable network specific factors, the individual's position in a network (as measured by her Katz-Bonacich centrality) is a key determinant of her level of activity. A standard deviation increase in the Katz-Bonacich centrality increases the pupil school performance by more than 7 percent of one standard deviation.

There are a number of possible extensions of this work. First, our theoretical analysis is restricted to linear-quadratic utility functions. It would be challenging to go beyond this case and see if the Katz-Bonacich-Nash linkage stills hold. Second, in this paper, we consider a utility function where peer effects are local aggregates so that it is the sum of peers' efforts that affects one's utility. It would be interesting to consider local average instead so that it would be the average effort of peers that matter. In that case, we would study how important is conformism in explaining educational outcomes. Finally, other outcomes than education could be studied. It is indeed well documented that social networks are important in the labor market, in criminal activities, in smoking behaviors, etc. It would be interesting to investigate if the location of an individual in a network of friends, as measured by her Katz-Bonacich index, has also a crucial impact on these outcomes.

References

[1] Akerlof, G.A. (1997), "Social distance and social decisions," *Econometrica* 65, 1005-1027.

[2] Angrist, J. and V. Lavy (1999), "Using Maimonides' rules to estimate the effect of class size on scholastic achievement," *Quarterly Journal of Economics* 114, 533-575.

[3] Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Boston: Kluwer Academic Publishers.

[4] Anselin, L. (1995), *SpaceStat Version 1.80. User's Guide*, Morgantown: Regional Research Institute, West Virginia University.

[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. and A. Calvó-Armengol (2007), "Interaction patterns with hidden complementarities," Unpublished manuscript, Universitat Autònoma de Barcelona.

[7] Bonacich, P. (1987), "Power and centrality: A family of measures," *American Journal of Sociology* 92, 1170-1182.

[8] Borgatti, S.P. (2003), "The key player problem," in *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, ed. by R. Breiger, K. Carley and P. Pattison, Committee on Human Factors, National Research Council, pp. 241-252.

[9] Bramoullé, Y., Djebbari, H. and B. Fortin (2006), "Identification of peer effects through social networks," Unpublished manuscript, Université de Laval.

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

[11] Brock, W. and S.E. Durlauf (2001), "Discrete choice models with social interactions," *Review of Economic Studies*, 68, 235–260.

[12] Calvó-Armengol, A. and M.O. Jackson (2004), "The effects of social networks on employment and inequality," *American Economic Review* 94, 426-454.

[13] 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, 939-958.

[14] Cvetković D., Doob, M. and H. Sachs (1979) *Spectra of Graphs: Theory and Applications*, Academic Press: New York.

[15] Debreu, G. and I.N. Herstein (1953), "Nonnegative square matrices," *Econometrica* 21, 597-607.

[16] De Giorgi, G., Pellizzari, M. and S. Redaelli (2007), "Be as careful of the books you read as of the company you keep. Evidence on peer effects in educational choices," IZA Discussion Papers 2833, Institute for the Study of Labor (IZA).

[17] Durlauf, S.E. (2002), "On the empirics of social capital," *Economic Journal* 112, F459-F479.

[18] Durlauf, S.E. (2004), "Neighborhood effects," In: J.V. Henderson and J-F. Thisse (Eds.), *Handbook of Regional and Urban Economics Vol. 4*, Amsterdam: Elsevier Science, pp. 2173-2242.

[19] Evans, W.N, Oates, W.E and R.M. Schwab (1992), "Measuring peer group effects: A study of teenage behavior," *Journal of Political Economy* 100, 966-991.

[20] Freeman, L.C. (1978/79), "Centrality in social networks: Conceptual clarification," *Social Networks* 1, 215-239.

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

[22] Galeotti, A. and G. Muller (2007), "Friendship relations in the school class and adult economic attainment," Unpublished manuscript, University of Essex.

[23] Goyal, S. (2007), *Connections: An Introduction to the Economics of Networks*, Princeton: Princeton University Press.

[24] Jackson, M.O. (2007) "The economics of social networks," In: R. Blundell, W.K. Newey, and T. Persson (Eds.), *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress, Vol. I*, Cambridge: Cambridge University Press, pp. 1-56.

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

[26] Jackson, M.O. and B.W. Rogers (2007), "Meeting strangers and friends of friends: How random are socially networks?" *American Economic Review* 97, 890-915.

[27] Katz, L. (1953), "A new status index derived from sociometric analysis," *Psychometrika* 18, 39-43.

[28] Kolstad, C.D. and L. Mathiesen (1987), "Necessary and sufficient conditions for uniqueness of a Cournot equilibrium," *Review of Economic Studies* 54, 681-690.

[29] Laschever, R. (2005), "The doughboys network: social interactions and labor market outcomes of World War I veterans," Unpublished manuscript, Northwestern University.

[30] Lin, X. (2005), "Peer effects and students' academic achievement: An application of spatial autoregressive model with group unobservables," Unpublished manuscript, Ohio State University.

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

[32] Sacerdote, B. (2001), "Peer effects with random assignment: Results from Dartmouth roomates," *Quarterly Journal of Economics* 116, 681-704.

[33] Topa, G. (2001), "Social interactions, local spillovers and unemployment," *Review of Economic Studies* 68, 261-295.

[34] Topkis, D. (1979), "Equilibrium points in nonzero-sum $n-$person submodular games," *SIAM Journal of Control and Optimization* 17, 773-787.

[35] Vives, X. (2005), "Complementarities and games: New developments," *Journal of Economic Literature* 43, 437-479.

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

[37] Zimmerman, D. (2003), "Peer effects in academic outcomes: Evidence from a natural experiment," *Review of Economics and Statistics* 85, 9-23.

Appendix 1: Proof of Proposition 1

We describe and analyze a more general network game with linear quadratic payoffs. Proposition 1 then follows as an immediate corollary of the equilibrium characterization for this game provided below in Theorem 1. We focus on the game where agents choose optimally z_i since for the game when they choose y_i^0, there are no interactions between agents, and the outcome is thus trivial.

$N = \{1, \ldots, n\}$ is a finite set of agents. Each agent $i \in N$ selects $z_i \geq 0$. Payoffs are:

$$u_i(\mathbf{z}) = \alpha_i z_i + \frac{1}{2}\sigma_{ii} z_i^2 + \sum_{j \neq i} \sigma_{ij} z_i z_j. \tag{16}$$

Let $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)$ and $\boldsymbol{\Sigma} = [\sigma_{ij}]$. We analyze the game $\Gamma(\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ with players in N, strategy space $\mathbb{R}_+$ for each player, and payoff (16).

The model of peer influence in the text whose utility function is given by (1) corresponds to a game $\Gamma(\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ where $\alpha_i = \mu\, g_i$, $\sigma_{ii} = -1$, and $\sigma_{ij} = \phi g_{ij}$, for all i and j.

More precisely, we focus on games $\Gamma(\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ such that $\boldsymbol{\alpha} > \mathbf{0}$ (that is, $\alpha_i > 0$, for all $i \in N$) and $\sigma_{ii} < \min\{0, \min\{\sigma_{ij} : j \neq i\}\}$, for all $i \in N$. We further assume that $\sigma_{ii} = \sigma_{11}$, for all $i \in N$. This is without loss of generality. Indeed, let $\mathbf{D} = diag(1, \sigma_{11}/\sigma_{22}, ..., \sigma_{11}/\sigma_{nn})$. This is a diagonal matrix with a strictly positive diagonal. It is readily checked that the Nash equilibria of $\Gamma(\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ and that of $\Gamma(\mathbf{D}\boldsymbol{\alpha}, \mathbf{D}\boldsymbol{\Sigma})$ coincide, where the diagonal terms of $\mathbf{D}\boldsymbol{\Sigma}$ are all equal to σ_{11}.

This model is analyzed in Ballester et al. (2006), who focus primarily on the case where $\boldsymbol{\alpha}$ is a diagonal vector (the general case is covered in Remark 2). Here, we provide a new and intuitive equilibrium existence and uniqueness condition for a general $\boldsymbol{\alpha}$, as well as closed-form equilibrium payoffs.

Following Ballester et al. (2006), let $\underline{\sigma} = \min\{\sigma_{ij} \mid i \neq j\}$, $\overline{\sigma} = \max\{\sigma_{ij} \mid i \neq j\}$, $\gamma = -\min\{\underline{\sigma}, 0\} \geq 0$, $\lambda = \overline{\sigma} + \gamma \geq 0$. We assume that $\lambda > 0$, which is a generic property.[44] Let $g_{ij} = (\sigma_{ij} + \gamma)/\lambda$, for $i \neq j$, and set $g_{ii} = 0$. By construction, $0 \leq g_{ij} \leq 1$. Let $\mathbf{G} = [g_{ij}]$, a zero-diagonal non-negative square matrix interpreted as the adjacency matrix of a network. Finally, let $\sigma = -\beta - \gamma$, where $\beta > 0$. Given that $\sigma < \min\{\underline{\sigma}, 0\}$, this is without loss of generality.

Let $\mathbf{I}$ be the identity matrix and $\mathbf{J}$ the matrix of ones. We obtain the following additive decomposition of the interaction matrix:

$$\boldsymbol{\Sigma} = -\beta\mathbf{I} - \gamma\mathbf{J} + \lambda\mathbf{G}. \tag{17}$$

This decomposition separates own-concavity effects $-\beta\mathbf{I}$ from global substitutability effects $-\gamma\mathbf{J}$ and local (network) complementarity effects $+\lambda\mathbf{G}$. We refer the reader to Ballester et al. (2006) for more details on this additive decomposition. Ballester and Calvó-Armengol (2007) generalize the matrix substitutability shift $-\gamma\mathbf{J}$ to arbitrary rank one matrices.

[44] Indeed, $\lambda = 0$ if and only if $\underline{\sigma} = \overline{\sigma}$, and this is a set of measure zero in $\mathbb{R}^2$.

Following this decomposition, payoffs (16) can now be rewritten as:

$$u_i(\mathbf{z}) = \alpha_i z_i - \frac{1}{2}(\beta - \gamma) z_i^2 - \gamma \sum_{j=1}^{n} z_i z_j + \lambda \sum_{j=1}^{n} g_{ij} z_i z_j, \text{ for all } i \in N. \tag{18}$$

The model of peer influence in the text whose utility function is given by (1) is such that $\alpha_i = \mu\, g_i$, $\beta = 1$, $\gamma = 0$ and $\lambda = \phi$.

Definition 1 *Given a vector* $\mathbf{u} \in \mathbb{R}_+^n$*, and* $a \geq 0$ *a small enough scalar, we define the vector of* $\mathbf{u}$*-weighted centrality of parameter* a *in the network* $\mathbf{g}$ *as:*

$$\mathbf{w}_{\mathbf{u}}(\mathbf{g}, a) = \left(\mathbf{I} - a\mathbf{G}^{-1}\right)\mathbf{u} = \sum_{p=0}^{+\infty} a^p \mathbf{G}^p \mathbf{u}. \tag{19}$$

Note that the Katz-Bonacich centrality $\mathbf{b}(\mathbf{g}, a)$ defined in (4) corresponds to the $\mathbf{u}$-weighted centrality with $\mathbf{u} = a\mathbf{G1}$ (where $\mathbf{1}$ is the vector of ones), that is, the vector $\mathbf{u}$ is a times the node connectivities $\mathbf{G1}$. Formally:

$$\mathbf{b}(\mathbf{g}, a) = \mathbf{w}_{a\mathbf{G1}}(\mathbf{g}, a). \tag{20}$$

Denote by $\omega(\mathbf{G})$ the largest eigenvalue of $\mathbf{G}$. For all vector $\mathbf{u} \in \mathbb{R}^n$, let $u = u_1 + ... + u_n$. We have the following result.

Theorem 1 *Consider a game* $\Gamma(\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ *with* $\boldsymbol{\alpha} > \mathbf{0}$ *and* $\boldsymbol{\Sigma}$ *decomposed additively as in (17).*

(a) *Suppose first that* $\boldsymbol{\alpha} = \alpha\mathbf{1}$*. Then,* $\Gamma(\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ *has a unique Nash equilibrium in pure strategies if and only if* $\beta > \lambda\omega(\mathbf{G})$*. This equilibrium* $\mathbf{z}^*$ *is interior and given by:*

$$\mathbf{z}^* = \frac{\alpha}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} \mathbf{w}_{\mathbf{1}}(\mathbf{g}, \lambda/\beta). \tag{21}$$

(b) *Suppose now that* $\boldsymbol{\alpha} \neq \alpha\mathbf{1}$*. Let* $\overline{\alpha} = \max\{\alpha_i \mid i \in N\}$ *and* $\underline{\alpha} = \min\{\alpha_i \mid i \in N\}$*, with* $\overline{\alpha} > \underline{\alpha} > 0$*. If* $\beta > \lambda\omega(\mathbf{G}) + n\gamma(\overline{\alpha}/\underline{\alpha} - 1)$*, then* $\Gamma(\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ *has a unique Nash equilibrium in pure strategies* $\mathbf{z}^*$*, which is interior and given by:*

$$\mathbf{z}^* = \frac{1}{\beta}\left[\mathbf{w}_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta) - \frac{\gamma w_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta)}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} \mathbf{w}_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)\right]. \tag{22}$$

Before proving this result, a number of comments are in order.

First, when $\boldsymbol{\alpha} = \alpha\mathbf{1}$, the equilibrium existence, uniqueness (and interiority) condition is independent of γ, the global level of substitutabilities, and only depends on the own concavity term β and the network of local complementarities $\lambda\mathbf{G}$. The condition $\beta > \lambda\omega(\mathbf{G})$ sets an upper bound on network complementarities. This upper bound guarantees that the positive feed-back loops in the network of complementarities do not trigger an unbounded escalation of efforts, but rather reach an equilibrium level. Notice that $\lambda\omega(\mathbf{G})$ accounts both for the size of complementarities, λ, and for their pattern, $\mathbf{G}$. Theorem 1(a) is established in Ballester et al. (2006).

Second, when $\boldsymbol{\alpha} = \alpha\mathbf{1}$, the equilibrium closed-form expression (22) boils down to (21). Indeed, notice that $\mathbf{w}_{\alpha\mathbf{1}}(\mathbf{g}, \lambda/\beta) = \alpha\mathbf{w}_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)$, and the identity then follows by simple algebra. The sufficient existence, uniqueness and interiority equilibrium condition in Theorem 1(b) also boils down to the necessary and sufficient existence, uniqueness and interiority equilibrium condition in Theorem 1(a).

Third, for general $\boldsymbol{\alpha}$, a necessary and sufficient condition for equilibrium existence and uniqueness is that $-\mathbf{\Sigma}$ has all its principal minors strictly positive, that is, $-\mathbf{\Sigma}$ is a $P-$matrix in the language of the linear complementarity problem (see Ballester and Calvó-Armengol, 2007). Nonetheless, the $P-$matrix condition does not guarantee that the equilibrium is interior (in which case it is given by the closed-form expression (22)). Besides, the $P-$matrix property is computationally very demanding and economically nonintuitive. Altogether, this motivates the sufficient condition in Theorem 1(b), which is derived from that in Theorem 1(a), but imposes a more stringent requirement on $\beta, \lambda, \mathbf{G}$ as the right-hand side of the inequality is now augmented by $n\gamma(\overline{\alpha}/\underline{\alpha} - 1) \geq 0$. In words, everything else equal, the higher the discrepancy $\overline{\alpha}/\underline{\alpha}$ of marginal payoffs at the origin, the lower the level of network complementarities $\lambda\omega(\mathbf{G})$ compatible with a unique and interior Nash equilibrium.

Notice that, absent of any payoff cross effect, the individual maximization problem has a unique solution $\alpha_i/(\beta - \gamma)$ that increases in α_i. Players with lower marginal payoffs α_i at the origin thus exhaust their marginal returns with a lower effort level than players with higher marginal payoffs.

In the presence of payoff complementarities, the player with the highest marginal payoff $\overline{\alpha}$ thus reaps "more" complementarities from her network peers and may want to increase her effort level without bound, unless the strength of the available complementarities is low enough. Theorem 1(b) sets precisely this upper bound.

Symmetrically, in the presence of payoff substitutabilities, the player with the lowest marginal payoff $\underline{\alpha}$ may want to free-ride on her network peers and decrease her effort level to zero, unless the strength γ of such substitutabilities is low enough. Again, Theorem 1(b) sets this upper bound.

To summarize, the condition in Theorem 1(b) bounds local complementarities $\lambda\omega(\mathbf{G})$, global substitutabilities γ and marginal payoff differences $\overline{\alpha}/\underline{\alpha}$ such that players have no incentives to increase their effort level without bound, neither to free-ride on their network peers by decreasing them down to zero. A unique and interior equilibrium is then achieved.[45]

Proof of Theorem 1: Part (a) is Theorem 1 in Ballester et al. (2006). The necessary part derives from Corollary 1 in Ballester and Calvó-Armengol (2006). We prove part (b).

[45] Bramoullé and Kranton (2007) present a public good network game and characterize the geometric pattern of free riders in the network as a function of its geometry. See Ballester and Calvó-Armengol (2007) for a connection between this network public good game and the network game with quadratic payoffs analyzed here.

Suppose that the game $\Gamma(\boldsymbol{\alpha}, \boldsymbol{\Sigma})$ has an interior equilibrium, which is obtained by solving $\partial u_i / \partial y_i(\mathbf{y}^*) = 0$, for all $i \in N$. The equilibrium conditions in matrix form are:

$$-\boldsymbol{\Sigma}\mathbf{z}^* = [\beta \mathbf{I} + \gamma \mathbf{J} - \lambda \mathbf{G}]\,\mathbf{z}^* = \boldsymbol{\alpha},$$

Notice that $\mathbf{J}\mathbf{z}^* = z^* \mathbf{1}$. We thus rewrite the equilibrium conditions as:

$$\beta \mathbf{z}^* = [\mathbf{I} - \lambda/\beta \mathbf{G}]^{-1}(\boldsymbol{\alpha} - \gamma z^* \mathbf{1}) = \mathbf{w}_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta) - \gamma z^* \mathbf{w}_{\mathbf{1}}(\mathbf{g}, \lambda/\beta).$$

Multiplying to the left by $\mathbf{1}^t$ and solving for z^* gives:

$$z^* = \frac{w_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta)}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)}.$$

Plugging back into the previous equation gives (22). We now check that this is indeed an interior equilibrium, that is, $z_i^* > 0$, for all $i \in N$, which is equivalent to:

$$w_{i,\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta) > \frac{\gamma w_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta)}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} w_{i,\mathbf{1}}(\mathbf{g}, \lambda/\beta), \text{ for all } i \in N. \tag{23}$$

From (19), we deduce that:

$$\overline{\alpha} w_{i,\mathbf{1}}(\mathbf{g}, \lambda/\beta) \geq w_{i,\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta) \geq \underline{\alpha} w_{i,\mathbf{1}}(\mathbf{g}, \lambda/\beta), \text{ for all } i \in N,$$

implying, in particular, that $\overline{\alpha} w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta) \geq w_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta)$.

A sufficient condition for (23) to hold is that a lower bound of the left-hand side is higher than an upper bound of the right-hand side, namely:

$$\underline{\alpha} > \overline{\alpha} \frac{\gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} \Leftrightarrow \frac{\beta}{w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} > \gamma \left(\frac{\overline{\alpha}}{\underline{\alpha}} - 1 \right). \tag{24}$$

By definition,

$$w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta) = \sum_{p=0}^{+\infty} \left(\frac{\lambda}{\beta} \right)^p \mathbf{1}^t \mathbf{G}^p \mathbf{1}. \tag{25}$$

We know that $\omega(\mathbf{G}^p) = \omega(\mathbf{G})^p$, for all $p \geq 0$.[46] Also, $\mathbf{1}^t \mathbf{G}^p \mathbf{1}/n$ is the average connectivity in the matrix $\mathbf{G}^p$ of paths of length p in the original network $\mathbf{G}$, which is smaller than that its spectral radius $\omega(\mathbf{G})^p$ (Cvetković *et al.* 1979). Therefore, (25) leads to the following inequality:

$$w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta) \leq n \sum_{p=0}^{+\infty} \left(\frac{\lambda}{\beta} \right)^p \omega(\mathbf{G})^p = \frac{n\beta}{\beta - \lambda \omega(\mathbf{G})}.$$

A sufficient condition for (24) to hold is thus:

$$\beta - \lambda \omega(\mathbf{G}) > n\gamma \left(\frac{\overline{\alpha}}{\underline{\alpha}} - 1 \right).$$

[46] Observe that $\omega(\mathbf{G}^p) = \omega(\mathbf{G})^p$ is true for both a *symmetric* and an *asymmetric* adjacency matrix $\mathbf{G}$ as long as $\mathbf{G}$ has non-negative entries $g_{ij} \geq 0$. This follows from the Perron-Frobenius theorem.

Clearly, this interior equilibrium is unique. ▮

The next example illustrates Theorem 1.

When $n = 2$, symmetric cross effects correspond either to substitutability or to complementarity, but not both. Formally, $\gamma\lambda = 0$. We analyze the cases $\gamma = 0$ and $\lambda = 0$ separately.

Example with $n = 2$ and $\gamma = 0$ The interaction matrix is:

$$\mathbf{\Sigma} = \begin{bmatrix} \beta & -\lambda \\ -\lambda & \beta \end{bmatrix}.$$

The equilibrium existence and uniqueness necessary and sufficient condition in Theorem 1(a) becomes $\beta > \lambda$. Indeed, equilibrium conditions for an interior equilibrium (i.e., zero marginal payoffs for both players) are:

$$\begin{bmatrix} \beta & -\lambda \\ -\lambda & \beta \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}.$$

It is readily checked that this system has a unique positive solution if and only if $\beta > \lambda$, given by:

$$\begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix} = \frac{1}{\beta^2 - \lambda^2} \begin{bmatrix} \beta\alpha_1 + \lambda\alpha_2 \\ \lambda\alpha_1 + \beta\alpha_2 \end{bmatrix},$$

which corresponds to (22) when $\gamma = 0$ and

$$\mathbf{G} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Example with $n = 2$ and $\lambda = 0$ The interaction matrix is now:

$$\mathbf{\Sigma} = \begin{bmatrix} \beta + \gamma & \gamma \\ \gamma & \beta + \gamma \end{bmatrix}.$$

When $\alpha_1 = \alpha_2$, the equilibrium condition in Theorem 1(a) is trivially satisfied.

Suppose that $\alpha_1 > \alpha_2$. The sufficient condition for equilibrium existence, uniqueness and interiority in Theorem 1(b) is $(\beta + 2\gamma)/2\gamma > \alpha_1/\alpha_2$. Instead, we show that the equilibrium existence, uniqueness and interiority is obtained 7here if and only if $\alpha_2 (\beta + \gamma)/\gamma > \alpha_1/\alpha_2$, highlighting the fact that Theorem 1(b) is only sufficient but not necessary. Beyond this simple example with only $n = 2$ players, we believe that the fact that the condition in Theorem 1(b) is too stringent is compensated by its full generality and economic appeal.

An effort profile $\mathbf{z}^* = (z_1^*, z_2^*) \in \mathbb{R}_+^2$ is a pure strategy Nash equilibrium of $\Gamma(\boldsymbol{\alpha}, \mathbf{\Sigma})$ if and only if:

$$\begin{aligned} \frac{\partial u_i}{\partial z_i}(\mathbf{z}^*) &= 0, \text{ for all } i = 1,2 \text{ such that } z_i^* > 0 \\ \frac{\partial u_i}{\partial z_i}(\mathbf{z}^*) &\leq 0, \text{ for all } i = 1,2 \text{ such that } z_i^* = 0. \end{aligned}$$

Notice that marginal payoffs are:

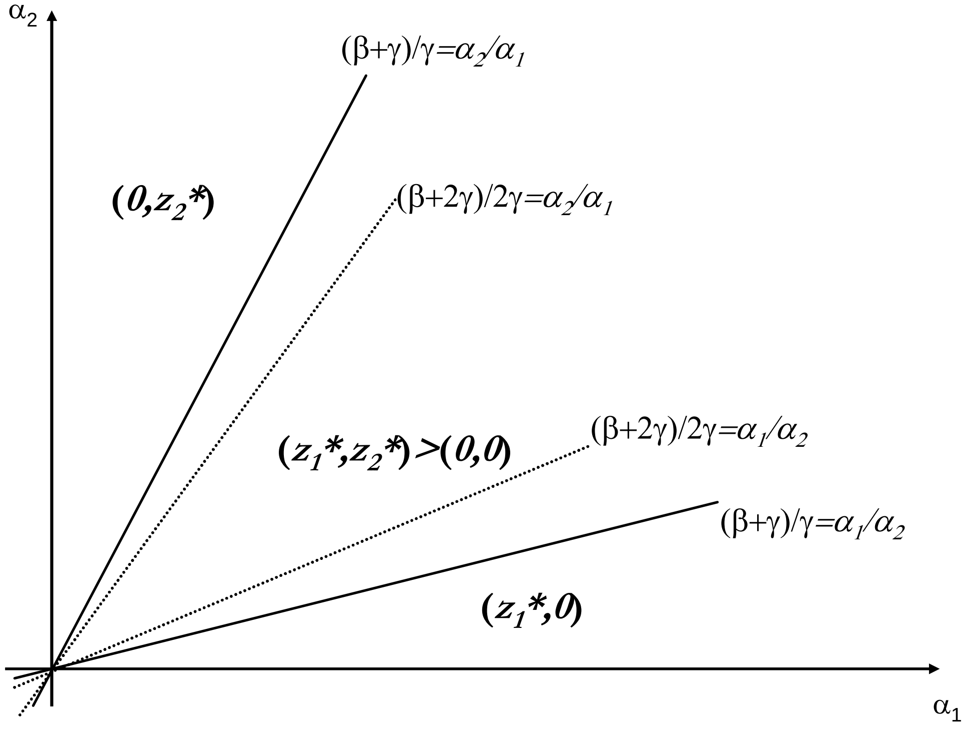
$$\frac{\partial u_i}{\partial z_i}(\mathbf{z}) = \mathbf{\Sigma z}.$$

The equilibrium conditions thus boil down to a system of inequalities. Straight algebra leads to the following equilibrium characterization, when $\alpha_1 > \alpha_2$:

$$\mathbf{z}^* = \begin{cases} \left(\frac{\alpha_1}{\beta+\gamma}, 0\right), \text{ if } (\beta+\gamma)/\gamma > \alpha_1/\alpha_2 \\ \frac{1}{(\beta+\gamma)^2-\gamma^2}\left[(\beta+\gamma)\alpha_1 - \gamma\alpha_2, -\gamma\alpha_1 + (\beta+\gamma)\alpha_2\right], \text{ otherwise} \end{cases}.$$

The case when $\alpha_1/\alpha_2 \geq (\beta+\gamma)/\gamma$ corresponds to (22) for $\lambda = 0$.

The next figure shows the regions for corner and interior equilibria. The dashed line corresponds to the sufficient condition in Theorem 1(b).



Proof of Proposition 1: The peer effects game is $\Gamma(\mu\mathbf{G1}, \mathbf{I}-\phi\mathbf{G})$. Confronting (18) with (1), we deduce that $\alpha_i = \mu\, g_i$, $\gamma = 0$, $\beta = 1$, $\lambda = \phi$. According to part (b) of Theorem 1, the solution is given by:

$$
\begin{aligned}
\mathbf{z}^* &= \mathbf{w}_{\mu\mathbf{G1}}(\mathbf{g},\phi) \\
&= \mu\,\mathbf{w}_{\mathbf{G1}}(\mathbf{g},\phi) \\
&= \frac{\mu}{\phi}\,\mathbf{w}_{\phi\mathbf{G1}}(\mathbf{g},\phi) \\
&= \frac{\mu}{\phi}\,\mathbf{b}(\mathbf{g},\phi)
\end{aligned}
$$

where the last equality is obtained using (20) for $a = \phi$. ∎

Appendix 2: Description of control variables

Individual socio-demographic variables

Female: dummy variable taking value one if the respondent is female.

Race: race of respondent, coded as 3-category dummies (white, the reference group, Black or African American and other races).

Age: respondent age measured in years.

Health status: response to the question “In the last month, how often did a health or emotional problem cause you to miss a day of school”, coded as 0= never, 1=just a few times, 2= about once a week, 3= almost every day, 4= every day.

Religion practice: response to the question: “In the past 12 months, how often did you attend religious services”, coded as 1= never, 2= less than once a month, 3= once a month or more, but less than once a week, 4= once a week or more.

School attendance: number or years the respondent has been a student at the school.

Student grade: grade of the student in the current year.

Organized social participation: dummy taking value one if the respondent participate in any clubs, organizations, or teams at school in the school year.

Motivation in education: response to the question: “how much do you try hard to do your school work well”, coded as 1=I never try at all, 2=I don’t try very hard, 3=I try hard enough, but not as hard as I could, 4=I try very hard to do my best.

Self esteem: response to the question: “Compared with other people your age, how intelligent are you”, coded as 1 = moderately below average, 2= slightly below average, 3= about average, 4= slightly above average, 5= moderately above average, 6= extremely above average.

Physical development: response to the question: “How advanced is your physical development compared to other boys/girls your age”, coded as 1= I look younger than most, 2= I look younger than some, 3= I look about average, 4= I look older than some, 5= I look older than most.

Family background variables

Household size: number of people living in the household.

Public assistance: dummy taking value one if either the father or the mother receives public assistance, such as welfare.

Mother working: dummy taking value one if the mother works for pay.

Two married parent family: dummy taking value one if the respondent lives in a household with two parents (both biological and non biological) that are married.

Single parent family: dummy taking value one if the respondent lives in a household with only one parents (both biological and non biological).

Parental education: schooling of (biological or non-biological) parent that is living with the child, coded as 1=never went to school, 2= not graduate from high school, 3= high school graduate, 4=graduated from college or a university, 5= professional training beyond a four-year college. If both parents are in the household, the education of the father is considered.

Parent age: mean value of the age of the parents (biological or non-biological) living with the child.

Parent occupation: closest description of the job of (biological or non-biological) parent that is living with the child, coded as 9-category dummies (doesn't work without being disables, the reference group, manager, professional or technical, office or sales worker, manual, military or security, farm or fishery, retired, other). If both parents are in the household, the occupation of the father is considered.

Protective factors

Parental care: dummy taking value one if the respondent reports that the (biological or non-biological) parent that is living with her/him or at least one of the parents (if both are in the household) cares very much about her/him.

Relationship with teachers: dummy taking value one if the respondent reports to have trouble getting along with teachers at least about once a week, since the beginning of the school year.

School attachment: composite score of three items derived from the questions: "How much do you agree or disagree that a) you feel close to people at your school, b) you feel like you are part of your school, c) you are happy to be at your school", all coded as 1= strongly agree, 2= agree, 3=neither agree nor disagree, 4= disagree, 5= strongly disagree. (Crombach-alpha =0.75).

Social exclusion: response to the question: "How much do you feel that adults care about you", coded as 1= very much, 2= quite a bit, 3= somewhat, 4= very little, 5= not at all.

Friend attachment: dummy taking value one if the respondent reports that he/she feels that his/her friends cares very much about him/her

Friend involvement: response to the question: "During the past week, how many times did you just hang out with friends", coded as 0= not at all, 1=1 or 2 times, 2=3 or 4 times, 3=5 or more times.

Residential neighborhood variables

Neighborhood quality: interviewer response to the question "How well kept are most of the buildings on the street", coded as 1= very poorly kept (needs major repairs), 2= poorly kept (needs minor repairs), 3= fairly well kept (needs cosmetic work), 4= very well kept.

Residential building quality: interviewer response to the question "How well kept is the building in which the respondent lives", coded as 1= very poorly kept (needs major repairs), 2= poorly kept (needs minor repairs), 3= fairly well kept (needs cosmetic work), 4= very well kept.

Neighborhood safety: dummy variable taking value if the interviewer felt concerned for his/her safety when he/she went to the respondent's home.

Residential area type: interviewer's description of the immediate area or street (one block, both sides) where the respondent lives, coded as 6-category dummies (rural, the reference group, suburban, urban - residential only, commercial properties - mostly retail, commercial properties - mostly wholesale or industrial, other).

Contextual effects

Average values of all the control variables over the respondent's direct friends (peer-group characteristics).

School fixed effects

Dummy variable taking value one if the school is the one attended by respondent.

Appendix 3: Proof of Proposition 2

We follow the proof methodology of Bramoullé et al. (2006).

Consider two sets of structural parameters (μ, ϕ) and (μ', ϕ') leading to the same reduced form (15), that is:

$$\mu \left(\mathbf{I} - \phi \mathbf{G}\right)^{-1} \mathbf{G1} = \mu' \left(\mathbf{I} - \phi' \mathbf{G}\right)^{-1} \mathbf{G1}.$$

We multiply both sides by $(\mathbf{I} - \phi \mathbf{G}) \left(\mathbf{I} - \phi' \mathbf{G}\right)$. Noticing the commutativity of all the matrices with each other, we obtain, after rearranging terms:

$$\left(\mu - \mu'\right) \mathbf{G1} + \left(\mu' \phi - \mu \phi'\right) \mathbf{G}^2 \mathbf{1} = \mathbf{0}.$$

Clearly, $(\mu, \phi) = (\mu', \phi')$ solves the previous system of linear equations, and this is the unique solution if and only if $\mu \neq 0$ and the matrices with column vectors $\mathbf{G1}$ and $\mathbf{G}^2\mathbf{1}$ have rank two, which is equivalent to $g_i^{[2]}/g_i \neq g_j^{[2]}/g_j$, for some $i \neq j$. ∎

Appendix 4: Alternative analyses

First alternative analysis

Theoretical model Denote by y_i the effort of individual i. Each agent i selects efforts $y_i \geq 0$ and obtains a payoff $u_i(\mathbf{y};\mathbf{g})$ that depends on the underlying network $\mathbf{g}$, in the following way:

$$u_i(\mathbf{y};\mathbf{g}) = [\mu g_i + \theta_i(\mathbf{x})]\ y_i - \frac{1}{2}y_i^2 + \phi \sum_{j=1}^{n} g_{ij} y_i y_j \tag{26}$$

where $\phi > 0$, $\mu > 0$, and $g_i = \sum_{j=1}^{n} g_{ij}$ is the number of direct links of individual i. As in the model in the main text, $\theta_i(\mathbf{x})$ introduces the *exogenous heterogeneity* that captures the *observable* differences between individuals and is still defined by (2). We now define a *new* Katz-Bonacich centrality, which is similar to (4), but where the heterogeneity stems from both the θs and the connectivities μg_i. Indeed, using (19) in Definition 1, the vector of $\mathbf{u}$-*weighted centrality* of parameter ϕ in the network $\mathbf{g}$ was defined as:

$$\mathbf{w_u}(\mathbf{g},\phi) = \left(\mathbf{I} - \phi \mathbf{G}^{-1}\right)\mathbf{u} = \sum_{p=0}^{+\infty} \phi^p \mathbf{G}^p \mathbf{u}$$

In the present model, the Katz-Bonacich centrality $\mathbf{b}(\mathbf{g},\phi)$ corresponds to the $\mathbf{u}$-weighted centrality with $\mathbf{u} = \mu\mathbf{G1} + \boldsymbol{\theta}$ (where $\mathbf{1}$ is the vector of ones and $\boldsymbol{\theta}$ is a vector of heterogeneities θs). Formally, define a *weighted* Katz-Bonacich centrality measure as follows:

$$\mathbf{b}(\mu\mathbf{G1} + \boldsymbol{\theta}, \mathbf{g}, \phi) = (\mathbf{I} - \phi\mathbf{G})^{-1} \cdot (\mu\mathbf{G1} + \boldsymbol{\theta}) = \mathbf{w}_{\mu\mathbf{G1}+\boldsymbol{\theta}}(\mathbf{g},\phi) \tag{27}$$

Interestingly, for each agent i,

$$b_i(\mu\mathbf{G1} + \boldsymbol{\theta}, \mathbf{g}, \phi) = \sum_{p=0}^{+\infty}\sum_{j=1}^{n} \phi^p\ (\mu g_j + \theta_j(\mathbf{x}))\ g_{ij}^{[p]}$$

is the sum of all paths in $\mathbf{g}$ starting from i weighted by the two heterogeneities $\mu g_j + \theta_j(\mathbf{x})$, where paths of length p are weighted by the geometrically decaying factor ϕ^p.

We can now characterize the Nash equilibrium of the game where agents choose their effort levels $y_i \geq 0$ simultaneously. Each individual i maximizes (26) and we obtain the following best reply function for each $i = 1, ..., n$:

$$\begin{aligned} y_i^*(\mathbf{x},\mathbf{g}) &= \mu g_i + \theta_i(\mathbf{x}) + \phi \sum_{j=1}^{n} g_{ij} y_j \\ &= \mu g_i + \sum_{m=1}^{M} \beta_m x_i^m + \frac{1}{g_i}\sum_{m=1}^{M}\sum_{j=1}^{n} \gamma_m g_{ij}\, x_j^m + \phi \sum_{j=1}^{n} g_{ij} y_j \end{aligned} \tag{28}$$

The optimal endogenous peer efforts are given by (28). Denote by $\omega(\mathbf{g})$ the largest eigenvalue of the adjacency matrix $\mathbf{G} = [g_{ij}]$ of the network.

Proposition 3 *Suppose that $\phi\omega(\mathbf{g}) < 1$. Then, the individual equilibrium outcome is uniquely defined and given by:*

$$y_i^* (\mathbf{x}, \mathbf{g}) = b_i (\mu \mathbf{G1} + \boldsymbol{\theta}, \mathbf{g}, \phi). \tag{29}$$

where $b_i (\mu \mathbf{G1} + \boldsymbol{\theta}, \mathbf{g}, \phi)$ is defined by (27).

Proof. Confronting (18) with (26), we deduce that $z_i = y_i$, $\alpha_i = \mu g_i + \theta_i(\mathbf{x})$, $\gamma = 0$, $\beta = 1$, $\lambda = \phi$. According to part (b) of Theorem 1, the solution is given by:

$$\begin{aligned} \mathbf{y}^* &= \mathbf{w}_{\mu \mathbf{G1}+\boldsymbol{\theta}} (\mathbf{g}, \phi) \\ &= \mathbf{b}(\mu \mathbf{G1} + \boldsymbol{\theta}, \mathbf{g}, \phi) \end{aligned}$$

where the last equality is obtained using (27). ∎

Empirical model For $i = 1, ..., n$; $\kappa = 1, ..., K$, the empirical counterpart of (28) is the following model:

$$y_{i,\kappa} = \mu g_{i,\kappa} + \sum_{m=1}^{M} \beta_m x_{i,\kappa}^m + \frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \gamma_m g_{ij,\kappa} x_{j,\kappa}^m + \phi \sum_{j=1}^{n_\kappa} g_{ij,\kappa} y_{j,\kappa} + \eta_\kappa + \nu_{i,\kappa} \tag{30}$$

where the notation of model (14) applies. In the spatial econometric literature, this is referred to as the *spatial lag model.* In matrix notation, and adding network fixed-effects, (30) can be written as:

$$\mathbf{y} = \mu \mathbf{G1} + \mathbf{X}\boldsymbol{\beta} + \mathbf{DGX}\boldsymbol{\gamma} + \phi \mathbf{Gy} + \boldsymbol{\eta} + \boldsymbol{\nu}$$

where $\mathbf{y}$ is a $n \times 1$ vector of observations on the dependent (decision) variable, $\mathbf{X}$ is a $n \times M$ matrix of observations on the exogenous variables associated to the $M \times 1$ regression coefficient vector $\boldsymbol{\beta}$, $\mathbf{D} = diag (1/g_1, ..., 1/g_n)$ is a $n \times n$ matrix, $\boldsymbol{\eta}$ is a $n \times n$ diagonal matrix of network fixed effects, with diagonal cells taking the same value within each network component, $\mathbf{1}$ is a $n \times 1$ vector of ones, $\mathbf{G}$ is a $n \times n$ spatial weight matrix that formalizes the network structure of the agents, ϕ is the spatial autoregressive parameter, and $\boldsymbol{\nu}$ is a vector of random error terms.

The reduced-form equation is now given by:

$$\mathbf{y} = \mu [\mathbf{I} - \phi \mathbf{G}]^{-1} \mathbf{G1} + [\mathbf{I} - \phi \mathbf{G}]^{-1} \mathbf{X}\boldsymbol{\beta} + [\mathbf{I} - \phi \mathbf{G}]^{-1} \mathbf{DGX}\boldsymbol{\gamma} + [\mathbf{I} - \phi \mathbf{G}]^{-1} (\boldsymbol{\eta} + \boldsymbol{\nu}). \tag{31}$$

Again, using Proposition 2, peer effects are identified if and only if $g_i^{[2]}/g_i \neq g_j^{[2]}/g_j$ for at least two agents i and j.

Observe that spatial lag models can be naturally interpreted in terms of Katz-Bonacich centrality measure. Indeed, when such a measure is defined as in (27), it is straightforward to see that the parameter estimates give the effect of two different Katz-Bonacich centrality measures on $\mathbf{y}$. The estimation of μ will give the impact of a pure location effect (the Katz-Bonacich index $\mu \left[\mathbf{I} - \phi\mathbf{G}\right]^{-1} \mathbf{G1}$ is weighted by the number of direct friends in the network) on educational outcome while the estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ will measure the impact of a mix of location and individual characteristic effects on y. Indeed, the Katz-Bonacich index $\left[\mathbf{I} - \phi\mathbf{G}\right]^{-1} \boldsymbol{\theta}$ accounts for both idiosyncratic characteristics and location in the network, and it is not possible to disentangle between the two effects. This approach implies that in practice, since the idiosyncratic heterogeneity θ_i of individual i is multidimensional (see Appendix B, when θ_i can be defined by the gender, race, age, education of parents, etc.), there are as many Katz-Bonacich centrality indices as idiosyncratic characteristics θ_i.

The spatial error model used in the text, instead, implies that the impact of idiosyncratic characteristics and of the endogenous peer effects can be additively separated and allows us to appreciate the relative merit of each component.

From an empirical point of view, when the spatial autoregressive parameter ϕ is not large (as in our case where we are equipped with a long list of controls that already explain a substantial part of the spatial association) the spatial error and the spatial lag model specifications might not be statistically different. The choice between the two models should be motivated on the theoretical ground.

Second alternative analysis

In our analysis, it is assumed that a student's idiosyncratic characteristics is not related to her peer effect outcomes. For example, suppose a student works hard at mathematics. This will have a direct effect on the individual's outcome and an effect from and to peers. Moreover, the effect on peers could depend on the individual's idiosyncratic efforts. The effort of a student from the bottom of the socio-economic ladder might have more effect than the effort of someone at the top. In order to address these issues, two alternative models can be proposed. First, we can change the utility function (1) in the following way:[47]

$$u_i(\mathbf{y}^0, \mathbf{z}; \mathbf{g}) = \theta_i \, y_i^0 - \frac{1}{2}\left(y_i^0\right)^2 + \mu \sum_{j=1}^{n} g_{ij}\theta_j z_i - \frac{1}{2} z_i^2 + \phi \sum_{j=1}^{n} g_{ij} z_i \theta_j z_j, \tag{32}$$

so that i's *friends' characteristics* θ_j directly affect peer effect outcomes but keep the separability equation (8). Observe that we have also added θ_j in the term μg_i (we have now $\mu g_i \theta_j z_i$ instead of

[47]We could even have a more general utility function with the last term being $\phi \sum_{j=1}^{n} g_{ij}\theta_i z_i \theta_j z_j$, so that both *own* and friend characteristics directly affect peer effect outcomes. This will, however, complicate the analysis without changing the main results.

$\mu g_i z_i)$[48] in order to be consistent with the last term $\phi \sum_{j=1}^{n} g_{ij} z_i \theta_j z_j$. In that case, the first order condition in z_i gives:

$$z_i^* (\mathbf{g}) = \mu \sum_{j=1}^{n} g_{ij} \theta_j + \phi \sum_{j=1}^{n} g_{ij} \theta_j z_j$$

instead of (7). Now, it is easy to see that the "peer effect outcome of student i", i.e. $z_i^* (\mathbf{g})$, is a function of not only the efforts of his/her friends z_j but also of "the idiosyncratic characteristics" of his/her friends, i.e. θ_j. So, for example if θ_j captures the social status of j's parents, then if two students j and k who are friends to i have different social statuses, they will have different impacts on $z_i^* (\mathbf{g})$. Indeed, if we only focus on the peer effort term, then individual j will have an impact of $\phi \theta_j z_j$ while the other will have an impact of $\phi \theta_k z_k$. In other words, as you said, "the effort of a student from the bottom of the socio-economic ladder might have more effect than the effort of someone at the top".

So if we solve the model with a utility function given by (32), then the two first order conditions are:

$$y_i^{0*} (\mathbf{x}) = \theta_i (\mathbf{x}) = \sum_{m=1}^{M} \beta_m x_i^m + \frac{1}{g_i} \sum_{m=1}^{M} \sum_{j=1}^{n} \gamma_m g_{ij} \, x_j^m \tag{33}$$

$$z_i^* (\mathbf{g}) = \mu \sum_{j=1}^{n} g_{ij} \theta_j (\mathbf{x}) + \phi \sum_{j=1}^{n} g_{ij} \theta_j (\mathbf{x}) \, z_j \tag{34}$$

and total outcome is now given by:

$$\begin{aligned} y_i^* (\mathbf{x}, \mathbf{g}) &= y_i^{0*} (\mathbf{x}) + z_i^* (\mathbf{g}) \\ &= \theta_i (\mathbf{x}) + \mu \sum_{j=1}^{n} g_{ij} \theta_j (\mathbf{x}) + \phi \sum_{j=1}^{n} g_{ij} \theta_j (\mathbf{x}) \, z_j \end{aligned}$$

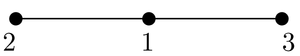
Observe that equation (34) can be written in matrix form as follows:

$$\mathbf{z} = \mu \mathbf{G} \mathbf{D}_{\boldsymbol{\theta}} \mathbf{1} + \phi \mathbf{G} \mathbf{D}_{\boldsymbol{\theta}} \mathbf{z}$$

where $\mathbf{1}$ is a $(n, 1)$ vector of 1 and $\mathbf{D}_{\boldsymbol{\theta}} = diag\,(\theta_1, ..., \theta_n)$ is a $n \times n$ matrix. Rewriting this equation gives:

$$\mathbf{z}^* = [\mathbf{I} - \phi \mathbf{G} \mathbf{D}_{\boldsymbol{\theta}}]^{-1} \mu \mathbf{G} \mathbf{D}_{\boldsymbol{\theta}} \mathbf{1}$$

Define $\mathbf{G}_{\boldsymbol{\theta}} = \mathbf{G} \mathbf{D}_{\boldsymbol{\theta}}$, which is equivalent to the adjacency matrix where each 1 (which defines a link) has been replaced by the idiosyncratic characteristic of the person with whom the individual is friend with. For example, if we take the following network



[48] Remember that $g_i = \sum_{j=1}^{n} g_{ij}$.

then

$$\mathbf{G}_{\boldsymbol{\theta}}=\mathbf{G}\mathbf{D}_{\boldsymbol{\theta}}=\begin{bmatrix} 0 & \theta_2 & \theta_3 \\ \theta_1 & 0 & 0 \\ \theta_1 & 0 & 0 \end{bmatrix}$$

Then, we have:

$$\begin{aligned}\mathbf{z}^* &= (\mathbf{I}-\phi\mathbf{G}_{\boldsymbol{\theta}})^{-1}\,\mu\mathbf{G}_{\boldsymbol{\theta}}\mathbf{1}\\ &= \frac{\mu}{\phi}(\mathbf{I}-\phi\mathbf{G}_{\boldsymbol{\theta}})^{-1}\,\phi\mathbf{G}_{\boldsymbol{\theta}}\mathbf{1}\\ &= \frac{\mu}{\phi}\mathbf{b}(\mathbf{g}_{\boldsymbol{\theta}},\phi).\end{aligned}$$

Thus, Proposition 1 still holds and we have the following result: Suppose that $\phi\omega(\mathbf{g}_{\boldsymbol{\theta}})<1$, then, the individual equilibrium outcome is uniquely defined and given by:

$$y_i^*(\mathbf{x},\mathbf{g})=\theta_i(\mathbf{x})+\frac{\mu}{\phi}b_i(\mathbf{g}_{\boldsymbol{\theta}},\phi).$$

The empirical equivalent of (33) and (34) are given by:

$$\begin{aligned}y_{i,\kappa} &= \sum_{m=1}^{M}\beta_m x_{i,\kappa}^m+\frac{1}{g_{i,\kappa}}\sum_{m=1}^{M}\sum_{j=1}^{n_\kappa}\gamma_m g_{ij,\kappa}x_{j,\kappa}^m+\eta_\kappa+\varepsilon_{i,\kappa},\\ \varepsilon_{i,\kappa} &= \mu\sum_{j=1}^{n_\kappa}g_{ij,\kappa}\theta_{j,\kappa}+\phi\sum_{j=1}^{n_\kappa}g_{ij,\kappa}\theta_{j,\kappa}\varepsilon_{j,\kappa}+\upsilon_{i,\kappa},\ \text{ for } i=1,...,n;\kappa=1,...,K,\end{aligned}$$

where

$$\theta_{j,\kappa}=\sum_{m=1}^{M}\beta_m x_{i,\kappa}^m+\frac{1}{g_{i,\kappa}}\sum_{m=1}^{M}\sum_{j=1}^{n_\kappa}\gamma_m g_{ij,\kappa}x_{j,\kappa}^m$$

In matrix notation, we have:

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\boldsymbol{\beta}+\mathbf{DGX}\boldsymbol{\gamma}+\boldsymbol{\eta}+\boldsymbol{\varepsilon}\\ \boldsymbol{\varepsilon} &= [\mathbf{I}-\phi\mathbf{G}_{\boldsymbol{\theta}}]^{-1}\,\mu\mathbf{G}_{\boldsymbol{\theta}}\mathbf{1}+[\mathbf{I}-\phi\mathbf{G}_{\boldsymbol{\theta}}]^{-1}\,\boldsymbol{\nu},\end{aligned}$$

where $\mathbf{D}=diag\,(1/g_1,...,1/g_n)$ is a $n\times n$ matrix. Of course, as in the first alternative model described above, the Katz-Bonacich index $[\mathbf{I}-\phi\mathbf{G}_{\boldsymbol{\theta}}]^{-1}\,\mu\mathbf{G}_{\boldsymbol{\theta}}\mathbf{1}$ accounts for both idiosyncratic characteristics and location in the network, and it is not possible to disentangle between the two effects. This approach also implies that in practice, since the idiosyncratic heterogeneity θ_i of individual i is multidimensional, there are as many Katz-Bonacich centrality indices as idiosyncratic characteristics θ_i.

Third alternative analysis

We now keep the same utility function defined by (1) but that we *relax the separability assumption* (8), that is the educational outcome is not additive in the two efforts but *multiplicative* and given by:

$$y_i^*(\mathbf{x},\mathbf{g}) = y_i^{0*}(\mathbf{x})\;\; z_i^*(\mathbf{g})$$

In that case, since the first conditions are still defined by (6) and (7), we have

$$y_i^*(\mathbf{x},\mathbf{g}) = \theta_i(\mathbf{x})\,\mu g_i + \theta_i(\mathbf{x})\,\phi\sum_{j=1}^{n} g_{ij} z_j$$

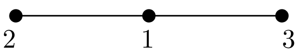
In matrix form, this can be written as:

$$\mathbf{y} = \mu\mathbf{D}_{\boldsymbol{\theta}}\mathbf{G}\mathbf{1} + \phi\mathbf{D}_{\boldsymbol{\theta}}\mathbf{G}\mathbf{y}$$

where, as above, $\mathbf{1}$ is a $(n,1)$ vector of 1 and $\mathbf{D}_{\boldsymbol{\theta}} = diag\,(\theta_1,...,\theta_n)$ is a $n\times n$ matrix. Rewriting this equation gives:

$$\mathbf{y}^* = [\mathbf{I} - \phi\mathbf{D}_{\boldsymbol{\theta}}\mathbf{G}]^{-1}\,\mu\mathbf{D}_{\boldsymbol{\theta}}\mathbf{G}\mathbf{1}$$

Define $\mathbf{G}_{\boldsymbol{\theta}}^T = \mathbf{D}_{\boldsymbol{\theta}}\mathbf{G}$, which is the *transpose* of the matrix $\mathbf{G}_{\boldsymbol{\theta}}$ where each 1 (which defines a link) has been replaced by the idiosyncratic characteristic of herself and not the person with whom the individual is friend with. For example, if we take the following network



then

$$\mathbf{G}_{\boldsymbol{\theta}}^T = \mathbf{D}_{\boldsymbol{\theta}}\mathbf{G} = \begin{bmatrix} 0 & \theta_1 & \theta_1 \\ \theta_2 & 0 & 0 \\ \theta_3 & 0 & 0 \end{bmatrix}$$

Then, we have:

$$\begin{aligned} \mathbf{y}^* &= \left[\mathbf{I} - \phi\mathbf{G}_{\boldsymbol{\theta}}^T\right]^{-1}\mu\mathbf{G}_{\boldsymbol{\theta}}^T\mathbf{1} \\ &= \frac{\mu}{\phi}\left[\mathbf{I} - \phi\mathbf{G}_{\boldsymbol{\theta}}^T\right]^{-1}\phi\mathbf{G}_{\boldsymbol{\theta}}^T\mathbf{1} \\ &= \frac{\mu}{\phi}\mathbf{b}(\mathbf{g}_{\boldsymbol{\theta}}^T,\phi). \end{aligned}$$

Thus we can have exactly the same theoretical analysis as for the other model with the only difference that we are dealing with $\mathbf{G}_{\boldsymbol{\theta}}^T$ and not $\mathbf{G}_{\boldsymbol{\theta}}$. The difference lies in the empirical analysis.

For $i = 1, ..., n$; $\kappa = 1, ..., K$, the empirical counterpart is the following model:

$$\begin{aligned} y_{i,\kappa} &= \left(\sum_{m=1}^{M} \beta_m x_{i,\kappa}^m + \frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \gamma_m g_{ij,\kappa} x_{j,\kappa}^m \right) \mu g_{i,\kappa} \\ &+ \left(\sum_{m=1}^{M} \beta_m x_{i,\kappa}^m + \frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \gamma_m g_{ij,\kappa} x_{j,\kappa}^m \right) \phi \sum_{j=1}^{n_\kappa} g_{ij,\kappa} y_{j,\kappa} + \eta_\kappa + \nu_{i,\kappa} \end{aligned}$$

In the spatial econometric literature, this is referred to as the *spatial lag model*. In matrix notation, we have:

$$\mathbf{y} = \left[\mathbf{I} - \phi \mathbf{G}_{\boldsymbol{\theta}}^T \right]^{-1} \mu \mathbf{G}_{\boldsymbol{\theta}}^T \mathbf{1} + \left[\mathbf{I} - \boldsymbol{\phi} \mathbf{G}_{\boldsymbol{\theta}}^{\mathbf{T}} \right]^{\mathbf{-1}} (\boldsymbol{\eta} + \boldsymbol{\nu})$$

Again, it will difficult to identify the Katz-Bonacich index and it will not be possible to disentangle between the two effects (location and idiosyncratic effects). The other problem is the fact that the idiosyncratic heterogeneity θ_i of individual i is multidimensional which leads to a multidimensional Katz-Bonacich index.

To conclude, when we abandon the separability assumption, then, in any model, it will difficult to separate peer effects due to individual location in the network (i.e. the "pure" Katz-Bonacich index) from the ones stemming from individual idiosyncratic characteristics.

Table 1. Descriptive statistics

	Mean	St. Dev.	Min	Max
Female	0.41	0.35	0	1
Black or African American	0.17	0.31	0	1
Other races	0.12	0.15	0	1
Age	15.29	1.85	10	19
Religion practice	3.11	1.01	1	4
Health status	3.01	1.77	0	4
School attendance	3.28	1.86	1	6
Student grade	9.27	3.11	7	12
Organized social participation	0.62	0.22	0	1
Motivation in education	2.23	0.88	1	4
Relationship with teachers	0.12	0.34	0	1
Social exclusion	2.26	1.81	1	5
School attachment	2.59	1.76	1	5
Parental care	0.69	0.34	0	1
Household size	3.52	1.71	1	6
Two married parent family	0.41	0.57	0	1
Single parent family	0.23	0.44	0	1
Public assistance	0.12	0.16	0	1
Mother working	0.65	0.47	0	1
Parental education	3.69	2.06	1	5
Parent age	40.12	13.88	33	75
Parent occupation manager	0.11	0.13	0	1
Parent occupation professional or technical	0.09	0.21	0	1
Parent occupation office or sales worker	0.26	0.29	0	1
Parent occupation manual	0.21	0.32	0	1
Parent occupation military or security	0.09	0.12	0	1
Parent occupation farm or fishery	0.04	0.09	0	1
Parent occupation retired	0.06	0.09	0	1
Parent occupation other	0.11	0.16	0	1

Table 1. Descriptive statistics (continued)

	Mean	St. Dev.	Min	Max
Neighborhood quality	2.99	2.02	1	4
Residential building quality	2.95	1.85	1	4
Neighborhood safety	0.51	0.57	0	1
Residential area suburban	0.32	0.38	0	1
Residential area urban - residential only	0.18	0.21	0	1
Residential area commercial properties - retail	0.12	0.15	0	1
Residential area commercial properties - industrial	0.13	0.18	0	1
Residential area type other	0.19	0.25	0	1
Friend attachment	0.49	0.54	0	1
Friend involvement	1.88	1.56	0	3
Physical development	3.14	2.55	1	5
Self esteem	3.93	1.33	1	6

Table 2. Correlation between individual and peer-group level observables

Variable	OLS
Parental education	0.0926
	(0.1054)
Parental care	−0.1427
	(0.1976)
Mathematics score	0.0515
	(0.0612)
Motivation in education	0.1632
	(0.1950)
School attachment	0.0435
	(0.0748)
Social exclusion	−0.0807
	(0.0899)
Individual socio-demographic variables	yes
Family background variables	yes
Protective factors	yes
Residential neighborhood variables	yes
Contextual effects	yes
School fixed effects	yes

Notes:
- Number of observations: 2,079,871 (11,491 pupils, 181 networks)
- Network fixed-effects OLS estimates are reported. They are within-group estimates where individuals are grouped by networks
- Standard errors in parentheses
- None of the coefficients is statistically significant at any conventional level
- The listed control variables are defined in Appendix 2
- Regressions are weighted to population proportions

Table 3. Model (14): Maximum Likelihood estimation results on key variables
Dependent variable: school performance index

	Undirected networks	Directed networks
Number of best friends (μ)	0.0314**	0.0323**
	(0.0149)	(0.0152)
Peer effects (ϕ)	0.5667***	0.5505***
	(0.1433)	(0.1247)
Individual socio-demographic variables	yes	yes
Family background variables	yes	yes
Protective factors	yes	yes
Residential neighborhood variables	yes	yes
Contextual effects	yes	yes
School fixed effects	yes	yes
R^2	0.8987	0.8905

Notes:
- Number of observations: 2,079,871 (11,491 pupils, 181 networks)
- Network fixed-effects estimates are reported. They are within-group estimates where individuals are grouped by networks
- Standard errors in parentheses
- Coefficients marked with one (two) [three] asterisks are significant at 10 (5) [1] percent level
- The listed control variables are defined in Appendix 2
- Regressions are weighted to population proportions

Table 4. Explanatory power of different unit centrality measures
Dependent variable: school performance index

	OLS	OLS	OLS
Degree centrality	0.2508*	-	-
	(0.1475)		
Closeness centrality	-	0.2892	-
		(0.2599)	
Beetweenness centrality	-	-	0.0621
			(0.0698)
Individual socio-demographic variables	yes	yes	yes
Family background variables	yes	yes	yes
Protective factors	yes	yes	yes
Residential neighborhood variables	yes	yes	yes
Contextual effects	yes	yes	yes
School fixed effects	yes	yes	yes
R^2	0.7958	0.8202	0.8001

Notes:
- Number of observations: 2,079,871 (11,491 pupils, 181 networks)
- Network fixed-effects OLS estimates are reported. They are within-group estimates where individuals are grouped by networks
- Standard errors in parentheses
- Coefficients marked with one (two) [three] asterisks are significant at 10 (5) [1] percent level
- The listed control variables are defined in Appendix 2
- Regressions are weighted to population proportions