Peer Effects and Social Networks in Education[*,†]

Antoni Calvó-Armengol[‡] Eleonora Patacchini[§] Yves Zenou[¶]

October 6, 2006

Abstract

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 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.

Keywords: Centrality measure, peer influence, network structure, school performance.

JEL Classification: A14, C31, C72, I21.

[*] We are grateful to Yann Bramoulle, David Genesove, Thierry Magnac 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).

[‡] ICREA and Universitat Autònoma de Barcelona, Department of Economics and Economic History, Edifici B, 08193 Bellaterra (Barcelona), Spain and CEPR. Email: antoni.calvo@uab.es

[§] Università di Roma "La Sapienza", Facoltà di Scienze Statistiche, Dipartimento di Contabilità Nazionale e Analisi dei Processi Sociali, P.le Aldo Moro, 5, 00185 Rome, Italy. E-mail: Eleonora.Patacchini@uniroma1.it

[¶] Research Institute of Industrial Economics, Box 55665, 102 15 Stockholm, Sweden, GAINS (Université du Maine) and CEPR. E-mail: yves.zenou@riie.se

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 theoretical model for peer effects that builds on the smallest unit of analysis for this cross influence, the dyad. The collection of dyadic bilateral relationships constitutes a social network, and our model relates 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

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

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 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). 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 his Katz-Bonacich centrality measure. We provide a closed-form expression for this Katz-Bonacich-Nash 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 that 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 his direct dyad partners. At equilibrium, though, each agent has to anticipate the actual behavior of his dyad partners to take on an optimal action himself. 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 conve-

[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. Jackson (2006) offers an exhaustive critical survey 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.

nient. For instance, the Nash-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, 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 (slightly modified) version of the so-called *spatial error model* (see, e.g. Anselin, 1988). Indeed, in the theoretical model peer effects results as 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. Recall that this parameter enters 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 the economics literature, here Nash equilibrium, and the sociology literature, here Katz-Bonacich

[4] Glaeser et al. (1996) study peer effecs in criminal behavior testing a specific model.

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.

2 A network model of peer effects

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

Individual outcomes result from the combination of individual characteristics and peer influence.

Consider some population of agents $\{1, ..., n\}$. Denote by y_i^0 the outcome of individual i absent of any peer influence. This is a function of idiosyncratic characteristics of the individual that we denote by $\mathbf{x}$. Formally, $y_i^0(\mathbf{x})$.

Suppose now that individuals exert some peer influence on each others. We assume that this peer influence is mediated by the network of friendship ties or peer relationships. 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. Therefore, i exerts a direct peer influence on j if and only if $g_{ij} = 1$. Because these influences are reciprocal, we assume that the network is symmetric, that is $g_{ij} = g_{ji}$. We also let $g_{ii} = 0$.

Denote by y_i the outcome of individual i that results from both idiosyncratic characteristics and peer influence. We assume that the peer influence acts as a multiplier on the behavior of the isolated individual. Denoting by $\pi_i(\mathbf{g}) \geq 0$ this multiplier, we have:

$$y_i(\mathbf{x}, \mathbf{g}) = [1 + \pi_i(\mathbf{g})]\; y_i^0(\mathbf{x}). \tag{1}$$

Clearly, an empty network (i.e. isolated agents) corresponds to no peer effects, and thus $\pi_i(\varnothing) = 0$, for all i.

Define $z_i(\mathbf{x}, \mathbf{g}) = \pi_i(\mathbf{g})\, y_i^0(\mathbf{x})$. Then, we can decompose additively individual behavior into an idiosyncratic part and a peer effect component that depends on the individual under consideration:

$$y_i(\mathbf{x}, \mathbf{g}) = \underbrace{y_i^0(\mathbf{x})}_{\text{idiosyncratic}} + \underbrace{z_i(\mathbf{x}, \mathbf{g})}_{\text{peer effect}}. \tag{2}$$

We assume that peer influence takes this particular linear form:

$$z_i = \mu g_i + \phi \sum_{j=1}^{n} g_{ij} z_j,\ i = 1, ..., n \tag{3}$$

where $\mu, \phi > 0$, and $g_i = \sum_{j=1}^{n} g_{ij}$ is the number of peer relationships entertained by i in the network $\mathbf{g}$. In words, the peer outcome of each player depends linearly on her level of exposure in the network of peers (g_i) and on the peer outcome of the players that influence her directly ($\sum_{j=1}^{n} g_{ij} z_j$). The influential part of ego is influenced by the influential part of her peers.

In fact, it can be shown (see the next section) that (3) is agent i's best response to the peer influence exerted by others for the following utility function:

$$u_i(\mathbf{y}^0, \mathbf{z}; \mathbf{g}) = u_i^0(\mathbf{y}^0) + \mu g_i z_i - \frac{1}{2} z_i^2 + \phi \sum_{j=1}^{n} g_{ij} z_i z_j. \tag{4}$$

This utility function is additively separable in the idiosyncratic component and the peer effects contribution. Besides, the peer effect contribution is linear quadratic. The idiosyncratic component $u_i^0(\mathbf{y}^0)$ introduces *idiosyncratic heterogeneity* that captures the intrinsic observable differences between individuals without taking into consideration peer effects. Examples of such heterogeneity are agent i's parents' education, neighborhood where he/she lives, age, sex, race, etc. The peer effect component is also heterogeneous, and this *peer heterogeneity* reflects the different locations of individuals in the friendship network $\mathbf{g}$. To be more precise, bilateral influences are captured by the following cross derivatives, for $i \neq j$:

$$\frac{\partial^2 u_i(\mathbf{y}^0, \mathbf{z}; \mathbf{g})}{\partial z_i \partial z_j} = \phi g_{ij} \geq 0. \tag{5}$$

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.

Note that the utility (4) 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. Equilibrium effort levels thus generally differ across agents in a manner that reflects the existing heterogeneity in friendship ties.

The exact mapping between network location and equilibrium outcome, though, is more intricate than simply counting direct network links. It is the purpose of next section.

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

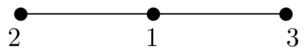


Figure 1. Three agents on a line.

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 he holds in the network that results from the dyads' overlap.

2.1 Analysis of the model

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 $\mathbf{G}^k = \mathbf{G}^{(k \overset{times}{\cdots})}\mathbf{G}$ 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:

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

Of course, we require that ϕ is small enough so that this infinite sum is well-defined.

Notice that $\sum_{k=0}^{+\infty} \phi^k\mathbf{G}^k = (\mathbf{I} - \phi\mathbf{G})^{-1}$, where $\mathbf{I}$ is the identity matrix. We can thus write the vector of Katz-Bonacich centralities in a more compact formula:

$$\mathbf{b}(\mathbf{g}, \phi) = (\mathbf{I}-\phi\mathbf{G})^{-1} \cdot (\phi\mathbf{G1}) . \tag{6}$$

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, $b_i(\mathbf{g}, \phi) \geq \phi g_i$.

Note also that (6) is well-defined for low enough values of ϕ, so that the infinite sum $\mathbf{1} + \phi\mathbf{G1} + \phi^2\mathbf{G}^2\mathbf{1} + \cdots$ converges to a finite value. 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,[5] the vector of Katz-Bonacich network centralities is:

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

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

Equilibrium behavior We now analyze the individual behavior characterized by (2) and (3). Recall that (3) can be interpreted as the best response in peer outcome for the payoffs (4). The individual behavior characterized by (2) and (3) thus corresponds to the Nash equilibrium of the game where agents choose their peer effort level $z_i \geq 0$ simultaneously.

It turns out that, when local network complementarities do not offset own decreasing marginal returns, the Nash equilibrium peer effort decision of each agent is uniquely defined and proportional to her Katz-Bonacich network centrality. We can thus map in a closed-form the ex ante idiosyncratic and peer heterogeneity into ex post heterogeneity in equilibrium outcomes. This is the purpose of the next result.

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 characterized by (2) and (3) is uniquely defined and given by:*

$$y_i^* = y_i^0 + \frac{\mu}{\phi} b_i(\mathbf{g},\phi). \tag{7}$$

Proof. See Appendix 1. ▮

In Appendix 1, we provide a theorem (Theorem 1) that generalizes the result of Proposition 1 for the case of an utility function with both local strategic complementarities (as in (4)) *and* global strategic substituabilities (see (12)). Observe that Theorem 1 in Appendix 1 is a generalization of Theorem 1 in Ballester et al. (2006) for the case of ex ante intrinsic heterogeneity, i.e. the idiosyncratic characteristics of each individual that we denote by $\mathbf{x}$. This generalization was important in order to be able to bring the model to the data.

The condition $\phi\omega(\mathbf{g}) < 1$ stipulates that network complementarities must be small enough than own concavity to prevent positive the feed-back loops triggered by such complementarities to escalate without bound. Network complementarities are measured by the compound index $\phi\omega(\mathbf{g})$, where ϕ refers to the intensity of each non-zero cross effect, whereas $\omega(\mathbf{g})$ refers to the pattern of such 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

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

complete network, where every agent is directly linked to every other agent, and is equal to $n-1$. A sufficient condition for the Nash-Katz linkage of Proposition 1 to hold for all networks is thus $\phi(n-1) < 1$.

Katz-Bonacich centrality is the right network index to account for equilibrium behavior. In (4), 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.[6] 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 $\phi\sqrt{2} < 1$, the individual equilibrium outcome is uniquely defined by:

$$y_1^* = y_1^0 + \mu \left(\frac{2 + 2\phi}{1 - 2\phi^2} \right)$$

$$y_2^* = y_2^0 + \mu \left(\frac{1 + 2\phi}{1 - 2\phi^2} \right)$$

$$y_2^* = y_3^0 + \mu \left(\frac{1 + 2\phi}{1 - 2\phi^2} \right)$$

The outcome of individual i depends both on the ex ante heterogeneity (y_i^0) and the 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 $y_i = y_i^0 + \mu\, g_i$. 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 y_i^0.

Network peer effects In this model, the structure of the social network and, in particular, the individual positions in such network, are the main explanatory variables for agents' behavior, together with idiosyncratic heterogeneity. This is the Nash-Bonacich linkage. In the education literature, for instance, social aspects as well as peer effects have been emphasized as important drivers for individual conduct,[7] 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. In the coming

[6] 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.

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

sections, we investigate the empirical relevance of this issue. The empirical measure of peer effects that is derived from our model thus differs substantially from previous work in this area.[8] Indeed, we are not looking at the impact of group peer effects on individual's activity in education. Instead, we consider the impact of the position of each individual in her network of peers (as measured by her Katz-Bonacich centrality) on education outcomes. In other words, the aim of our empirical analysis is to explore to which extent the structural position of each individual in a peer network, captured by her Katz-Bonacich centrality, accounts for her observed education outcome.

3 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).

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.[9] 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 friends nominations. Pupils were asked to identify their best friends from a school roster (up to five males and five females).[10] By matching the identification numbers of the friendship nominations to respondents' identification numbers, one can obtain information on the characteristics of nominated friends. Very importantly, one can also reconstruct the whole geometric structure of the friendship network. A link exists between two friends if at least one of the two individuals has identified the other as his/her best friend. For each school, we obtain all the networks of (best) friends.[11]

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

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

[10] The limit in the number of nominations is not binding. Less than 1% of the students in our sample show a list of ten best friends. On average, they declare to have 5.48 friends with a small dispersion around this mean value (standard deviation equal to 1.29).

[11] Note that, when an individual i identifies a best friend j who does not belong to the same school, the database does not include j in the network of i; it provides no information about j. Fortunately, in the large majority of cases

Education 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. The Crombach-α measure is then used to assess the quality of the derived variable. In our case, we obtain an α equal to 0.86 ($0 \leq \alpha \leq 1$) indicating that the different items incorporated in the index have considerable internal consistency. The mean is 2.34 and the standard deviation is equal to 2.11. The distribution of the index is not far from normal.[12]

By merging the in-home data to 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. Table 1 provides descriptive statistics on the students selected in this sample. 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*]

3.1 Descriptive evidence

Figures 2 displays the empirical distribution of the networks in our sample by their size (i.e. the number of network members).[13] It appears that the distribution is roughly normal and 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.

(more than 93%), best friends tend to be in the same school and thus are systematically included in the network.

[12] 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.

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

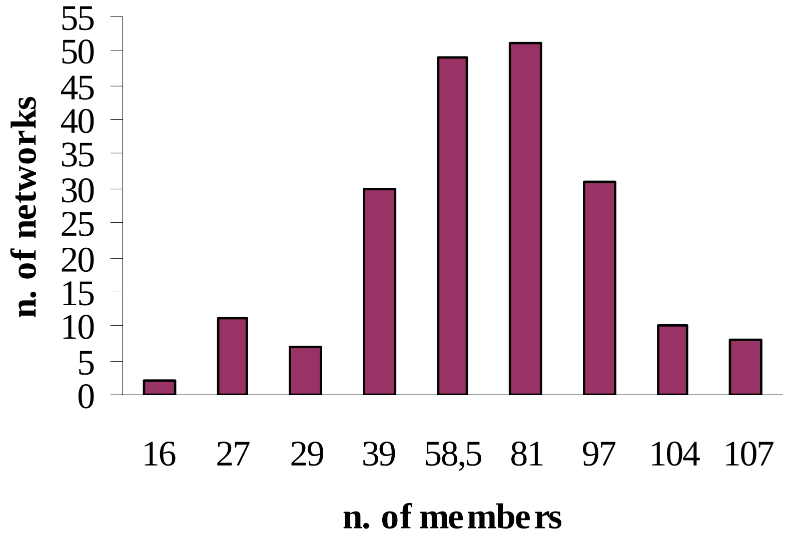


Figure 2. The empirical distribution of adolescent networks

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). This is a maximal network component, so that no student in the network has nominated any student outside the network. The largest network in our sample is almost seven times bigger and has 107 members.

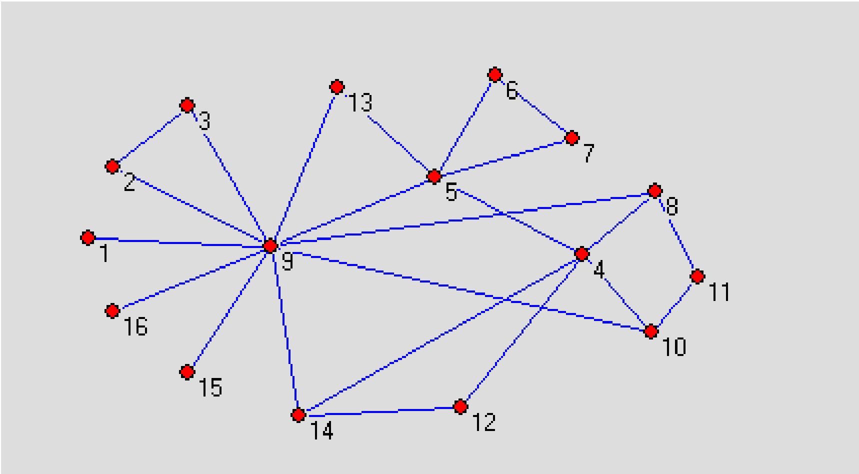


Figure 3. Smallest network of adolescents ($n = 16$)

4 Empirical strategy

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^*.

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 (2) and (3) is the following model:[14]

$$\begin{aligned} y_{i,\kappa} &= \sum_{m=1}^{M} \beta_m x_{i,\kappa}^m + \frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \gamma_m g_{ij,\kappa} x_{j,\kappa}^m + \eta_\kappa + \varepsilon_{i,\kappa}, \qquad (8) \\ \varepsilon_{i,\kappa} &= \mu g_{i,\kappa} + \phi \sum_{j=1}^{n_\kappa} g_{ij,\kappa} \varepsilon_{j,\kappa} + \upsilon_{i,\kappa}, \;\; i = 1, ..., n; \;\; \kappa = 1, ..., K, \end{aligned}$$

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,

[14] Notice that the structural equations (8) for $y_{i,\kappa}$ include the impact of contextual exogenous (given the network) variables, $\frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \gamma_m g_{ij,\kappa} x_{j,\kappa}^m$, thus departing slightly from the strict idiosyncratic individual heterogeneity highlighted in the theory. We do so because this specification allows us to match the standard empirical studies on peer effects and to focus on the endogenous effects that are generally captured by the error term.

neighborhood and school characteristics,[15] $g_{i,\kappa} = \sum_{j=1}^{n_\kappa} g_{ij,\kappa}$ 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 (8) 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 model is a variation of the Anselin (1988) spatial error model.[16]

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 (7) in Proposition 1). Indeed, our model allows us to decompose additively individual behavior into an idiosyncratic effect and a peer effect (see (2)), which boils down to the individual Katz-Bonacich centrality index.[17]

[15] A precise description of all these variables is contained in Appendix 3.

[16] 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 $n \times k$ matrix of observations on the exogenous variables associated to the $n \times 1$ regression coefficient vector $\boldsymbol{\beta}$, which do not include anything about the spatial topology, $\mathbf{D} = diag\left(1/g_1, ..., 1/g_n\right)$ 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.

[17] 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 the uniqueness of the equilibrium and the interiority of the solution. For that, we have estimated model (8) 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). This indicates that there is nothing unique about the discarded individuals.

4.1 The 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.

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 (8), 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.

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 (8) is, in matrix notation:

$$\boldsymbol{\varepsilon} = \mu \left[\mathbf{I} - \phi \mathbf{G}\right]^{-1} \mathbf{G1} + \left[\mathbf{I} - \phi \mathbf{G}\right]^{-1} \boldsymbol{\nu}. \tag{9}$$

We say that peer effects are identified if the structural parameters (μ, ϕ) uniquely determine the reduced-form coefficients in (9).

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. In particular, $g_i^{[2]}/g_i$ is the average connectivity of agent i's direct contacts.[18]

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 2. ∎

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 a simple property of the network, that amounts to checking that the $2 \times n$ matrix with column vectors $\mathbf{G1}$ and $\mathbf{G}^2\mathbf{1}$ is 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 (9)).

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 $\mathbf{G}^2\mathbf{1} = r^2\mathbf{1}$,[19] and so $g_i^{[2]}/g_i = r$, for all i. Identification fails in this case.[20]

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.

The role of specific controls Finally, the richness of the information provided by the AddHealth questionnaire on adolescents' behavior allow 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

[18] Indeed,

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

[19] If $\mathbf{G}$ is a regular (and symmetric) network with common node connectivity r, then $\mathbf{G}/r$ is a bi-stochastic matrix, with alls rows and columns that add up to one. Then, $(\mathbf{G}/r)^2 = \mathbf{G}^2/r^2$ is also a bi-stochastic matrix, implying that the rows and columns of $\mathbf{G}^2$ add up to r^2.

[20] 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

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.[21]

5 Empirical results

5.1 The Katz-Bonacich network centrality index

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.

Appendix 4 provides detailed statistical evidence that model (8) is appropriate and correctly specified. We can summarize this analysis as follows. 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. Although the set of explanatory variables included is wider that the one typically used in the estimation of an education production function (see Appendix 3), 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 (modified) spatial error model (8) derived from our theoretical set up is not rejected by the data.

Model (8) is estimated using the Maximum Likelihood approach. Different sets of controls have been used (see Appendix 3). 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

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

(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 ML estimation results for the model specification that includes the complete set of controls are reported in Table 2.[22] The estimated μ and ϕ are both positive and highly statistically significant. We then calculate the Katz-Bonacich measure (expression (6)) 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 variable on education outcomes that is predicted by the theory, i.e. $\widehat{\mu}/\widehat{\phi}$ (equation (7) in Proposition 1) is statistically significant[23] 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).

[*Insert Table* 2 *here*]

5.2 An alternative measure of network centrality: Betweenness

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 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.[24] 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.[25]

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

[23]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 statstical significance at the 5% level.

[24]See Freeman (1978/79) for an example of betweenness centrality, equal to the mean of the shortest-path distance between some given node and all other nodes that can be reached in the network.

[25]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.

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.

For robustness check, we test the explanatory value of an alternative centrality measure. We use betweenness centrality. This is a very popular network measure that, to our knowledge, lacks any behavioral (nor axiomatic) grounding. Freeman (1978/79) defines the betweenness centrality measure of agent i in a network component $\mathbf{g}_\kappa$ the following way:

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

where j and l denote two given agents in $\mathbf{g}_\kappa$. Friendships networks are, by definition, undirected networks, where relationships are reciprocal, $g_{ij,\kappa} = g_{ji,\kappa}$. For undirected networks, 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 is a parameter-free network measure. In our data, the normalized betweenness measure f_i^* has a mean equal to 0.45 and a standard deviation equal to 0.51.

Table 3 reports the estimation results obtained when using this alternative centrality measure as an additional explanatory variable in a OLS regression of individual outcomes on our set of observable individual characteristics, contextual and school-specific effects. It appears that this variable has statistically a non significant impact.

[*Insert Table* 3 *here*]

This result contrasts with the important role played by the Katz-Bonacich centrality index (Table 2). There are two main explanations for the discrepancy in the explanatory power of the betweenness centrality versus the Katz-Bonacich centrality.

The first reason is that the unique Nash equilibrium of a peer effects game is described exactly 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, betweenness centrality is, 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 betweenness centrality is a parameter-free network index. It only depends on the network geometry. Instead, Katz-Bonacich centrality depends both on the network topology and on the prevailing peer effect strength inside the group.

6 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 (8) for each network $\mathbf{g}$ separately,[26] thus using within network variation only. We obtain K 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 $\mathbf{g}$. These estimated values, $\widehat{\phi}_1, ..., \widehat{\phi}_K$, vary widely across the $K = 181$ school-peer networks. These differences are partly driven by the structural differences across such networks, as we examine below.

Figures 4a-4c plot the estimated $\widehat{\phi}_k$s (with confidence bands) against three different structural network measures.[27] 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 all possible ties present in a network. It ranges from 0 to 1 as networks get denser. Network asymmetry is the ratio of the highest to the lowest node connectivity in a network component. It is related to the variance of connectivities. We normalize it, so that it reaches 1 for the most asymmetric network in the sample. Network redundancy is the fraction of all transitive triads[28] 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.[29] Again, we normalize it.

[26] It has to be clear that when we estimated model (8) in Section 5.1, we obtained only *one* value of $\widehat{\phi}$ and not as in this section one ϕ per network component.

[27] The $\widehat{\phi}_k$s are normalized to be 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.

[28] 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 $i \to j$ and $j \to k$, then $i \to k$.

[29] See Jackson and Rogers (2004) for more details.

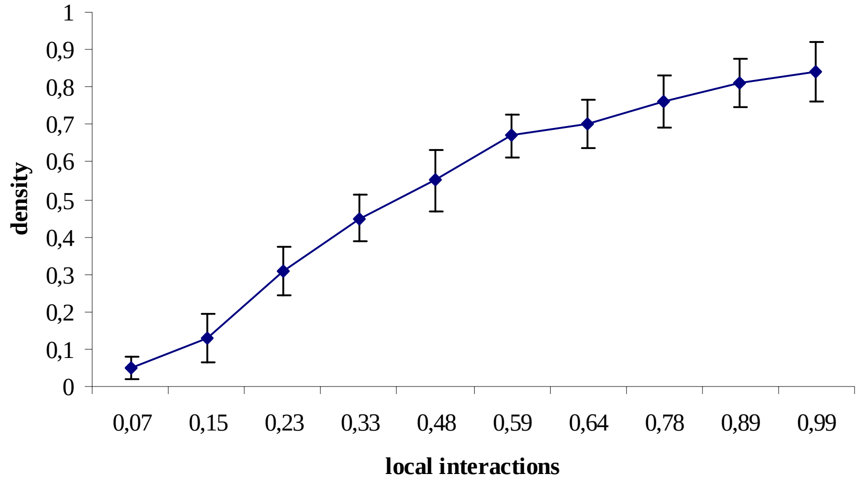


Figure 4a: Density in education networks

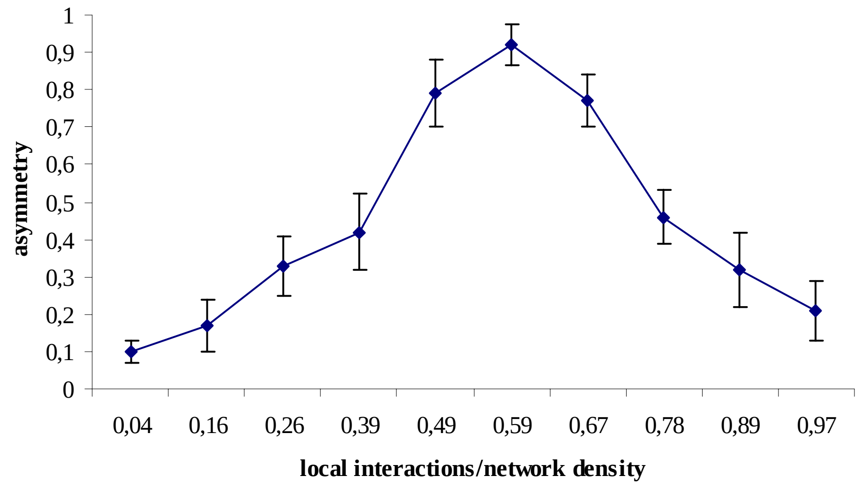


Figure 4b: Asymmetry in education networks

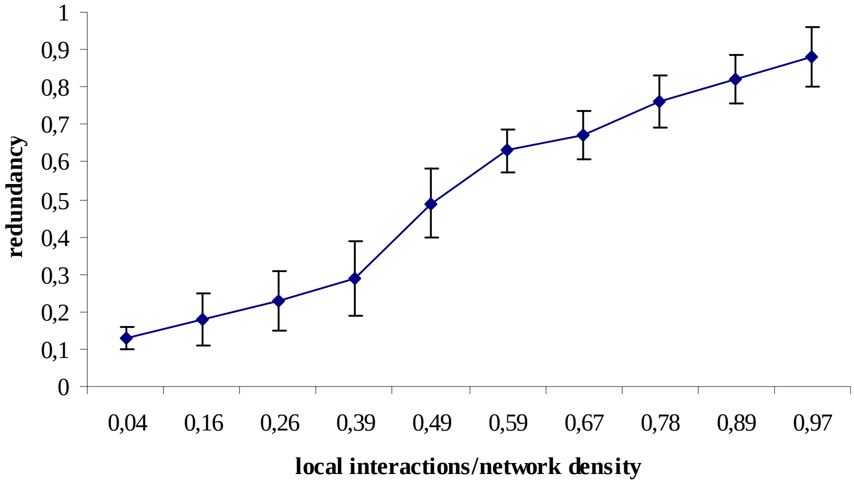


Figure 4c: Redundancy in education networks

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, 2004). 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.

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 (2006), "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. and R. Kranton (2006), "Public goods in networks," *Journal of Economic Theory*, forthcoming.

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

[11] Brock, W. and S.E. Durlauf (2001), "Discrete choice models with social interactions," *Review of Economic Studies*, 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] Cvetković D., Doob, M. and H. Sachs (1979) *Spectra of Graphs: Theory and Applications*, Academic Press: New York.

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

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

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

[17] 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.

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

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

[20] Jackson, M.O. (2006), "The Economics of Social Networks," Chapter 1 in Volume I of: *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, edited by Richard Blundell, Whitney Newey, and Torsten Persson, Cambridge University Press.

[21] Jackson, M.O. and B. Rogers (2004), "Search in the Formation of Large Networks: How Random Are Socially Generated Graphs?," Unpublished manuscript, Californian Institute of Technology.

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

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

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

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

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

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

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

$$u_i(\mathbf{z}) = \alpha_i z_i + \frac{1}{2}\sigma_{ii} z_i^2 + \sum_{j \neq i} \sigma_{ij} z_i z_j. \tag{10}$$

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 (10).

The model of peer influence in the text whose utility function is given by (4) 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.[30] 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:

$$\boldsymbol{\Sigma} = -\beta\mathbf{I} - \gamma\mathbf{J} + \lambda\mathbf{G}. \tag{11}$$

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 (2006) generalize the matrix substitutability shift $-\gamma\mathbf{J}$ to arbitrary rank one matrices.

[30] 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 (10) can now be rewritten as:

$$u_i(\mathbf{z}) = \alpha_i z_i - \frac{1}{2}(\beta - \gamma) z_i^2 - \gamma \sum_{j=1}^{n} z_i z_j + \lambda \sum_{j=1}^{n} g_{ij} z_i z_j, \text{ for all } i \in N. \tag{12}$$

The model of peer influence in the text whose utility function is given by (4) 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:*

$$\mathbf{w_u}(\mathbf{g}, a) = [\mathbf{I} - a\mathbf{G}]^{-1} \mathbf{u} = \sum_{p=0}^{+\infty} a^p \mathbf{G}^p \mathbf{u}. \tag{13}$$

Note that the Katz-Bonacich centrality $\mathbf{b}(\mathbf{g}, a)$ defined in (6) 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:

$$\mathbf{b}(\mathbf{g}, a) = \mathbf{w}_{a\mathbf{G1}}(\mathbf{g}, a). \tag{14}$$

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 (11).*

(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:*

$$\mathbf{z}^* = \frac{\alpha}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} \mathbf{w_1}(\mathbf{g}, \lambda/\beta). \tag{15}$$

(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:*

$$\mathbf{z}^* = \frac{1}{\beta}\left[\mathbf{w}_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta) - \frac{\gamma w_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta)}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} \mathbf{w_1}(\mathbf{g}, \lambda/\beta)\right]. \tag{16}$$

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 and 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 (16) boils down to (15). 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 $-\boldsymbol{\Sigma}$ has all its principal minors strictly positive, that is, $-\boldsymbol{\Sigma}$ is a $P-$matrix in the language of the linear complementarity problem (see Ballester and Calvó-Armengol, 2006). Nonetheless, the $P-$matrix condition does not guarantee that the equilibrium is interior (in which case it is given by the closed-form expression (16)). 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.[31]

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).

[31] Bramoullé and Kranton (2006) 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 (2006) 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 (16). We now check that this is indeed an interior equilibrium, that is, $z_i^* > 0$, for all $i \in N$, which is equivalent to:

$$w_{i,\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta) > \frac{\gamma w_{\boldsymbol{\alpha}}(\mathbf{g}, \lambda/\beta)}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} w_{i,\mathbf{1}}(\mathbf{g}, \lambda/\beta), \text{ for all } i \in N. \tag{17}$$

From (13), 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 (17) to hold is that a lower bound of the left-hand side is higher than an upper bound of the right-hand side, namely:

$$\underline{\alpha} > \overline{\alpha}\frac{\gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)}{\beta + \gamma w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} \Leftrightarrow \frac{\beta}{w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta)} > \gamma\left(\frac{\overline{\alpha}}{\underline{\alpha}} - 1\right). \tag{18}$$

By definition,

$$w_{\mathbf{1}}(\mathbf{g}, \lambda/\beta) = \sum_{p=0}^{+\infty}\left(\frac{\lambda}{\beta}\right)^p \mathbf{1}^t\mathbf{G}^p\mathbf{1}. \tag{19}$$

We know that $\omega(\mathbf{G}^p) = \omega(\mathbf{G})^p$, for all $p \geq 0$. Also, $\mathbf{1}^t\mathbf{G}^p\mathbf{1}/n$ is the average connectivity in the matrix $\mathbf{G}^p$ of paths of lenght p in the original network $\mathbf{G}$, which is smaller that its spectral radius $\omega(\mathbf{G})^p$ (Cvetković *et al.* 1979). Therefore, (19) 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 (18) to hold is thus:

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

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 (16) 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 here 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:

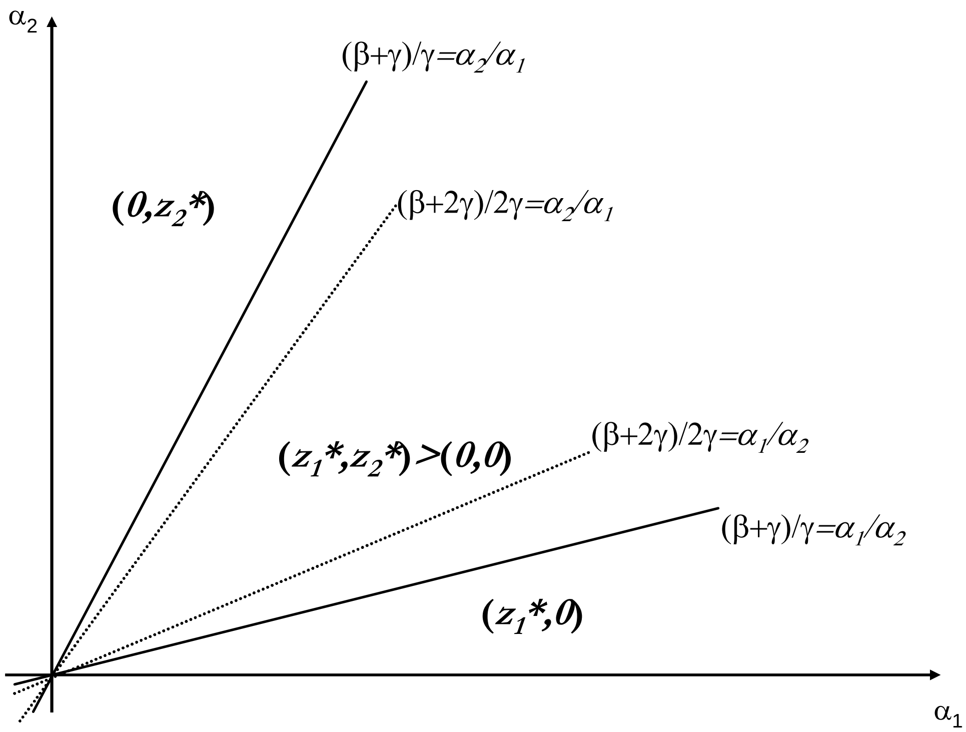
$$\frac{\partial u_i}{\partial z_i}(\mathbf{z}) = \mathbf{\Sigma}\mathbf{z}.$$

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 (16) 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 (12) with (4), 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 (14) for $a = \phi$. ∎

Appendix 2: Proof of Proposition 2

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

Consider two sets of structural parameters (μ, ϕ) and $\left(\mu', \phi'\right)$ leading to the same reduced form (9), 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 $\left[\mathbf{I} - \phi \mathbf{G}\right]\left[\mathbf{I} - \phi' \mathbf{G}\right]$. 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, $(\mu, \phi) = \left(\mu', \phi'\right)$ solves the previous system of linear equations, and this is the unique solution if and only if $\mu \neq 0$ and the matrix with column vectors $\mathbf{G1}$ and $\mathbf{G}^2\mathbf{1}$ has rank two, which is equivalent to $g_i^{[2]}/g_i \neq g_j^{[2]}/g_j$, for some $i \neq j$. ∎

Appendix 3: 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: respondents' 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: dummy taking value one if the respondent reports to try very hard to do his/her 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 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 (exogenous group characteristics).

School fixed effects

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

Appendix 4: Statistical details

In Table A1 we provide evidence that the statistical models defined by (8) is appropriate and correctly specified. Table A1 contains measures of statistical performance and the results of hypotheses tests. The results of the OLS estimation (with network fixed effects) of a traditional education production function where the individual school performance is explained as a function of a set of observable variables (those listed in Appendix 3, including school dummies) are used as a benchmark. They are reported in column two. Column three then shows the Maximum likelihood estimation results of our model-driven (modified) spatial error model (8). Specifically, the table reports the maximized log likelihood (LIK) and two likelihood based measures of goodness of fit: Akaike Information Criteria (AIC) and Schwartz Criterion (SC). A range of specification diagnostics follows. When estimating a classical regression model (column two), it consists of the Jarque-Bera test against non-normality (T_1), the Breusch-Pagan test against heteroskedasticity (T_2), a Lagrange Multiplier test on remaining spatial error autocorrelation (T_3) and a Lagrange Multiplier test on the spatial autoregressive coefficient (T_4). When estimating a spatial error model (column three), obviously we do not find the statistic T_1 (normality is assumed) and T_4 (there is no spatially lagged dependent variable included in the model specification), there is still a Breush-Pagan test against heteroskedasticity (T_2) and a test on the spatial error autoregressive coefficient (T_3), which is a Likelihood Ratio test in this case. In addition, we also find a Likelihood Ratio test (T_5) and a Wald test (T_6) on the common factor hypothesis. These last two tests verify if the coefficients soddisfy the restrictions needed to guarantee the consistency of the spatial error specification (see, e.g. Anselin, 1988). All the statistics are asymptotically distributed as a chi-squared. They differ in terms of degrees of freedom. The T_1 statistic presents two degree of freedom, both T_3 and T_4 statistics have one degree of freedom and T_2, T_5 and T_6 have as many degrees of freedom as the number of regressors in the model.[32]

Let us focus our attention on the analysis of the specification of model (8). Looking at the diagnostics in column two (classical regression model), the hypothesis of normality of the errors cannot be rejected (the T_1 statistic is not significant). This implies that the other misspecification tests (various Lagrange multiplier tests), that depend on the normality assumption, can be safely used. The T_2 statistic is not significant, providing no evidence of heteroskedasticity. On the contrary, both tests for spatial dependence (T_3 and T_4) are highly significant, indicating clearly the presence of spatial dependence ignored in the model. Although the T_4 statistic is slightly more significant than the T_3 statistic, there is no clear indication to conclude which is the proper alternative spatial model to use (spatial lag model or spatial error model). We choose the spatial error model on the basis of economic considerations (test of the theoretical model).

[32] For more details and a technical discussion of model validation in spatial regression models (measures of fit and specification diagnostics), see Anselin (1995).

Looking at column three (spatial error model), we can observe that the performance of the spatial model has been improved with respect to the standard regression model (column two) and it appears correctly specified. In fact, if we compare the values of LIK, AIC and SC for this spatial model with the ones reported in the second column (standard regression model), we can observe an increase in the value of LIK and a decrease in the value of AIC and SIC. This is consistent with an evidence that the fit of the model has been improved. Furthermore, the T_2 statistic is still not significant, providing evidence that there is no ignored heteroskedasticity in the model, the T_3 statistic is highly significant and neither T_5 nor T_6 are significant, indicating that the spatial error specification is appropriate.

TABLE A1: REGRESSION DIAGNOSTICS

	OLS standard regression model (with network fixed effects)	ML spatial error model (with network fixed effects)
LIK	−110.12	−103.78
AIC	264.24	251.55
SC	317.97	305.29
T_1	3.366 [0.1858]	—
T_2	26.734 [0.1798]	25.887 [0.2108]
T_3	8.88** [0.0029]	11.13*** [0.0008]
T_4	9.04*** [0.0026]	—
T_5	-	27.98 [0.1407]
T_6	-	29.16 [0.1102]

Notes:

- p-value in squared brackets

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
Friend contacts	0.89	0.12	0	1
Physical development	3.14	2.55	1	5
Self esteem	3.93	1.33	1	6

Table 2: Model (8) Maximum Likelihood estimation results on key variables
Dependent variable: school performance index

	ML (with network fixed effects)
Number of best friends (μ)	0.0314**
	(0.0149)
Peer effects (ϕ)	0.5667***
	(0.1433)
Individual socio-demographic variables	yes
Family background variables	yes
Protective factors	yes
Residential neighborhood variables	yes
Contextual effects	yes
School fixed effects	yes
$R^2 = 0.8987$	

Notes:
- Number of observations: 2,079,871 (11,491 pupils, 181 networks)
- Regressions are weighted to population proportions
- Standard errors in parentheses.
Coefficients marked with one (two) [three] asterisks
are significant at 10 (5) [1] percent level

Table 3: OLS estimation results on key variables
Dependent variable: school performance index

	OLS (with network fixed effects)
Beetweenness measure	0.0621
	(0.0698)
Individual socio-demographic variables	yes
Family background variables	yes
Protective factors	yes
Residential neighborhood variables	yes
Contextual effects	yes
School fixed effects	yes
$R^2 = 0.8001$	

Notes:
- Number of observations: 2,079,871 (11,491 pupils, 181 networks)
- Regressions are weighted to population proportions
- Standard errors in parentheses.
Coefficients marked with one (two) [three] asterisks
are significant at 10 (5) [1] percent level