# DISCUSSION PAPER SERIES

No. 9149

## STUDENT NETWORKS AND LONG-RUN EDUCATIONAL OUTCOMES: THE STRENGTH OF STRONG TIES

Eleonora Patacchini, Edoardo Rainone and Yves Zenou

**LABOUR ECONOMICS**

Centre for Economic Policy Research

www.cepr.org

# STUDENT NETWORKS AND LONG-RUN EDUCATIONAL OUTCOMES: THE STRENGTH OF STRONG TIES

**Eleonora Patacchini, University of Rome and CEPR**
**Edoardo Rainone, University of Rome**
**Yves Zenou, Stockholm University and CEPR**

This Discussion Paper is issued under the auspices of the Centre's research programme in **LABOUR ECONOMICS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

# ABSTRACT

## Student Networks and Long-Run Educational Outcomes: The Strength of Strong Ties*

The aim of this paper is to investigate and understand the effect of high-school friends on years of schooling. We develop a simple network model where students first choose their friends and then decide how much effort they put in education. The empirical salience of the model is tested using the four waves of the AddHealth data by looking at the impact of school peers nominated in the first two waves in 1994-1995 and in 1995-1996 on the educational outcome of teenagers reported in the fourth wave in 2007-2008 (when adult). We find that there are strong and persistent peer effects in education but peers tend to be influential only when they are strong ties (friends in both wave I and II) and not when they are weak ties (friend in one wave only). We also find that this is not true in the short run since both weak and strong ties tend to influence current grades.

Eleonora Patacchini
University of Rome La Sapienza
Box 83 - RM62
P.le Aldo Moro 5
00185 Roma
ITALY

Email:
eleonora.patacchini@uniroma1.it

Edoardo Rainone
University of Rome La Sapienza
P.le Aldo Moro 5
00185 Roma
ITALY

Email: edoardo.rainone@uniroma1.it

Yves Zenou
Department of Economics
Stockholm University
10691 Stockholm
SWEDEN

Email: yves.zenou@ne.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=126027

# 1    Introduction

The influence of peers on education outcomes has been widely recognized both in economics and sociology (Sacerdote, 2011). The extremely difficult task is to disentangle neighborhood effects from peer effects and there is no consensus on the importance of peer effects on own achievement in this literature (see the recent literature surveys by Durlauf, 2004, Ioannides and Topa, 2010, and Ioannides, 2011). The constraints imposed by the available disaggregated data force many studies to analyze peer effects in education at a quite aggregate and arbitrary level, such as at the high school (Evans et al., 1992), the census tract (Brooks-Gun et al., 1993), and the ZIP code (Datcher, 1982; Corcoran et al., 1992) where individuals reside. Besides, the mechanisms by which the peer effects affect education is unclear. In this paper, we focus on the long-run effects of high-school peers on years of schooling and try to pin down the mechanisms behind these effects.

To be more precise, we develop a simple network model where students first choose their friends and then decide how much effort they put in education. This model helps us understand how students make friends and how peers affect each other in the long run. Indeed, individuals create friendship relationships when they are at school anticipating the impact of their friendship choices on their future educational outcomes. The key factors that affect this relationship are the costs of having friends, the strength of interactions in the network, and the characteristics of each individual involved in the friendship relationship.

From an empirical point of view, very little is known about the effect of school peers on the long-run outcomes of teenagers. This is primarily due to the absence of information on peers together with long-run outcomes of individuals in most of the existing data. Our analysis is made possible by using an unique dataset of friendship networks in the United States, the Addhealth data, which allows us to test the peer effect aspect of the model, i.e. how friends in school affect future educational outcomes. For that, we use the four waves of the AddHealth data by looking at the impact of school friends nominated in the first two waves in 1994-1995 and in 1995-1996 on own educational outcome (when adult) reported in the fourth wave in 2007-2008 (measured by the number of completed years of full time education). We exploit four unique features of the AddHealth data: *(i)* the nomination-based friendship information, which allows us to reconstruct the precise geometry of social contacts, *(ii)* the directed nature of the nominations to measure precisely peer groups, *(iii)* the nomination order, which enables us to consider heterogenous influences within peer groups, *(iv)* the longitudinal dimension, which provides a temporal interval between friends' nomination and educational outcomes.

To the best of our knowledge, this is the first paper that exploits this comprehensive set of information to assess peer effects in education in this long-run perspective.

We find that there are *strong and persistent peer effects in education*. In other words, the "quality" of friends (in terms of future educational achievement) from high school has a positive and significant impact on own future education level. In terms of magnitude, we find that a standard deviation increase in peers' aggregate years of education (roughly two more high-school graduate friends) translates into roughly a 10 percent increase of a standard deviation in the individual's education attainment (roughly 3.5 more months of education). This is a strong effect, especially given our long list of controls and the fact that friendship networks might have changed over time. It is even stronger when the peer influence is allowed to be heterogenous according to the order of nomination. The influence of peers at school seems to be carried over time. We also analyze if the peer effect results are stronger for friends in earlier grades than in later ones. For that, we split our sample between students who were in grades 7-9 and those who were in grades 10-12. We find that peer effects are stronger for the latter than for the former. In other words, peers later in high school seem to matter more than peers earlier in high school. This is primarily due to two factors: peers in the last years of high school are the interacting friends when the choice to go to college is made and/or they are peers carried on from early years at school with whom friendship ties are stronger.

To better understand these mechanisms, we investigate the role of *weak* and *strong ties* in educational decisions. We define a *strong-tie* relationship between two students if they have nominated each other in both waves (i.e. in Wave I in 1994-95 and in Wave II in 1995-96) and a *weak-tie* relationship if they have nominated each other in one wave only. We then try to disentangle the *decision proximity* effect and the *strength of interaction* effect by exploiting variation in peer nominations between Waves I and II. Is it really the timing of friendship formation that is the crucial factor for educational outcome in the future (decision proximity effect) or is it rather the frequency of social interactions underlying common choices (strength of interaction effect)? We find that the education decision of weak ties have no significant effect on individual education outcomes, regardless on whether peers are interacting in lower or higher grades. On the contrary, we find that strong ties' educational choice have a positive and significant effect on own educational outcome and that this is effect is stronger when the network is weighted by the order of nomination. Using the theoretical model above, we can explain why strong ties matter more than weak ties. In the model, in order to keep the same friendship relationship for two periods (strong ties) rather than one period, it has to be that the costs of having friends are relatively low and the characteristics of each individual

involved in the friendship relationship are relatively close.

A further explanation of why peers nominated in different time periods have a different long-run effect could be the fact that students value differently peer characteristics in friendship decision making over time. Do students select peers differently between the first and the second wave or is it really that distinct types of peers (weak versus strong ties) matter differently? We provide evidence showing that there are no differences between peers in Wave I and II in terms of observable characteristics and that link formation is not different between the different waves.

We also look at the long versus short-run effects of peers on education. While in the long run, only strong ties matter, we find that, in the short run, both weak and strong ties are important in determining a student's performance at school. The magnitude of these effects is comparable: for weighted networks, a standard deviation increase in aggregate GPA of peers translates respectively into a 3.3 (for peers only in wave I), a 4.4 (for peers only in wave II) and a 3.2 (for peers in both waves) percent increase of a standard deviation in the individual's GPA. This means that peer effects for GPAs are substantially homogenous among different type of ties. Taken our analysis as a whole, our results then seem to suggest that, in the short run, all peers matter for education (i.e. grades) while, in the long run, only strong ties matter for future educational choices.

In the last part of the paper, we test the robustness of our results with respect to mis-specification of network topology or relevant peers. For that, we provide a falsification check using simulated data by randomly deleting some friendship links and randomly creating new links, keeping total number of links the same in each network. We find that peer effects remain statistically significant up to a percentage of randomly replaced (interchangeable) links of about 20 percent. This implies that even if we do not observe or imprecisely observe a portion of each individual's social network, our results on the existence of peer effects hold.

The rest of the paper unfolds as follows. In the next section, we discuss the related literature and explain our contribution. The theoretical model is developed in Section 3. Our data are described in Section 4, while the estimation and identification strategy are discussed in Section 5. The empirical results of the impact of peer effects on educational outcomes are provided in Section 6. Section 7 investigates the economic mechanisms behind our peer-effect results. In Section 8, we perform some robustness checks by looking at what happens to the peer-effect results if links are mis-specified. Finally, Section 9 concludes and discusses some policy implications.

# 2 Related literature

**Theory of network formation and education outcomes** There are some early theoretical contributions of peer effect models in education (De Bartolome, 1990; Benabou, 1993) but very few papers are using a network approach. Calvó-Armengol et al. (2009) is the first model with an explicit network that looks at the impact of peers on education outcomes. Here we add a first stage by looking at both network formation and educational outcomes. It is well-known that this issue is very complicated (Jackson and Zenou, 2013) because of coordination problems and, in fact, very difficult to solve. While we can solve the second stage for any network, we can only provide examples solving the whole two-stage problem. They will, however, help us interpret our empirical results.

**Empirical aspects of education with social interactions** There is an important literature on peer effects in education (for a survey, see Sacerdote, 2011). Different empirical strategies have been used to identify these effects ranging from an instrumental variable approach (see e.g. Evans et al., 1992; Cutler and Glaeser, 1997; Card and Rothstein, 2007; Weinberg, 2004) to specific social experiments or quasi-experimental data (e.g. Katz et al., 2001; Sacerdote, 2001; Zimmerman 2003). Based on recent papers (Bramoullé et al., 2009; Calvó-Armengol et al., 2009; Lin 2010; Liu et al. 2012; Patacchini and Zenou, 2012), we consider the identification of peer effects in *social networks* where they can be separately identified from contextual effects using the variations in the reference groups across individuals (see also De Giorgi et al. 2010). In particular, Lin (2010) and Liu et al. (2012) present a network model specification and an empirical strategy that is closely related to the one presented in this paper. Using data from the first wave of the AddHealth data, these studies provide an assessment of peer effects in student academic performance (GPA) and in crime, respectively. Even though we adopt a relative similar empirical strategy,[1] we do not focus on the same issues since we look at long-term effects of peers on education and distinguish between weak and strong ties.

**Long-run effects on educational outcomes** There are very few studies looking at the long-run effects of education. There is a literature that looks at the long-run effects of education interventions on outcomes. For example, Angrist and Krueger (1991) examine the long-term effects of compulsory schooling laws on adult educational attainment. Others have shown evidence for the effect of childhood investments on post-secondary attainment

---

[1]See Section 5 to see the differences.

(Krueger and Whitmore, 2001; Dynarski et al., 2011). The main research question of this literature (whether the effects of certain educational interventions can persist beyond test scores and lead to long-term enhancements to human capital) is quite different to ours since we are interested in the impact of peers on long-run education outcomes.

Using the Wisconsin Longitudinal Study of Social and Psychological Factors in Aspiration and Attainment (WLS), Zax and Rees (2002) where the first to analyze the role of friendships in school on future earnings. Their paper is, however, quite different from ours since they do not have a theoretical model driving the empirical analysis and do not tackle the issue of endogenous sorting of individuals into groups. They also do not provide the mechanisms behind their result. Using the British National Child Development Study (NCDS), Patacchini and Zenou (2011) investigate the effects of neighborhood quality (in terms of education) when a child is thirteen on his/her educational outcomes when he/she is adult. Similarly, Gould et al. (2011), using Israeli data, estimate the effect of the early childhood environment on a large array of social and economic outcomes lasting almost 60 years. In both studies, peer effects are measured by the neighborhood where people live and not by friendship nominations. Finally, Bifulco et al. (2011), using the AddHealth data, study the effect of school composition (percentage of minorities and college educated mothers among the students in one's school cohort) on high-school graduation and post-secondary outcomes. Our research question is different as we mostly focus on the mechanisms behind our results.

**Weak and strong ties in education**   There is a large literature on the role of weak ties in the labor market. Granovetter (1973, 1974, 1983) put forward the idea that weak ties are superior to strong ties for providing support in getting a job.[2] This is because, in a close network where everyone knows each other, information is shared and so potential sources of information are quickly shaken down so that the network quickly becomes redundant in terms of access to new information. In contrast Granovetter stresses the *strength of weak ties* involving a secondary ring of acquaintances who have contacts with networks outside ego's network and therefore offer new sources of information on job opportunities. The existing empirical evidence lends some support to Granovetter's ideas. Yakubovich (2005) uses a large scale survey of hires made in 1998 in a major Russian metropolitan area and finds that a

---

[2]In his seminal papers, Granovetter (1973, 1974, 1983) defines *weak ties* in terms of lack of overlap in personal networks between any two agents, i.e. weak ties refer to a network of acquaintances who are less likely to be socially involved with one another. Formally, two agents A and B have a weak tie if there is little or no overlap between their respective personal networks. Vice versa, the tie is *strong* if most of A's contacts also appear in B's network.

worker is more likely to find a job through weak ties than through strong ones. These results come from a within-agent fixed effect analysis, so are independent of workers' individual characteristics. Using data from a survey of male workers from the Albany NY area in 1975, Lin et al. (1981) find similar results. Lai et al. (1998) and Marsden and Hurlbert (1988) also find that weak ties facilitate the reach to a contact person with higher occupational status, who in turn leads to better jobs, on average.[3] To the best of our knowledge, there is no study on the impact of weak or strong ties on educational outcomes. Our study is one of the first and, compared to the literature on the labor market, we find the opposite result. Indeed, we show that strong rather than weak ties matter in education outcomes (years of schooling).

**This paper's contributions**   To sum-up, compared to the literature, we have the following main contributions:

($i$) We provide a theoretical model where both friendship formation and educational effort and attainment are modeled.

($ii$) We look at long-run outcomes, in the effect of high-school friends on years of schooling. This has been considered in other settings (Bifulco et al., 2011) but not for endogenous peer effects.

($iii$) We explore more specialized networks than have been previously considered in the literature (i.e., that the first person listed may be a better friend than the second, etc.) by assigning different weights to reported friends.

($iv$) We examine the role of strong ties versus weak ties on long-run effects of education.

($v$) We look at both the short-term and long-term effects of peers on education and show the different roles of weak and strong ties.

($vi$) We provide a falsification test that examines whether the set of relevant peers of each student is mis-measured.

# 3   Theoretical framework

Consider a population of $n$ individuals. The timing is as follows. In the first stage (time $t$), friendship relationships between the $n$ agents are made and a network connecting them is formed. Then, in the second stage (time $t + 1$), educational choices are made.

---

[3]See Topa (2011) for an overview and Patacchini and Zenou (2008) who find evidence of the strength of weak ties in crime.

## 3.1 Second-stage: Educational choices

Let us solve the second stage of this model. As stated above, in the first stage, a network[4] is formed with $n$ agents. It is a network of peer effects, where the network reflects the collection of active bilateral influences.

**The network** $N = \{1, \ldots, n\}$ is a finite set of agents. We keep track of social connections by a matrix $\mathbf{G}_t = \{g_{ij,t}\}$, where $g_{ij,t} = 1$ if $i$ and $j$ are direct friends, and $g_{ij,t} = 0$, otherwise. Friendship are reciprocal so that $g_{ij,t} = g_{ji,t}$. We also set $g_{ii,t} = 0$. Observe that we index the network variables by $t$ since links are made in the first stage (time $t$). For the sake of the presentation, we focus on un-weighted and undirected networks so that $\mathbf{G}_t$ is a $(0,1)$ symmetric matrix. All our results hold for *non-symmetric* and *weighted* networks but, for the ease of the presentation, we focus on symmetric and unweighted networks in the theoretical model.

**Preferences** Individuals decide how much effort to exert in education (e.g. how many years to study). As stated above, this decision is made in the second stage and thus all variables corresponding to this choice will be indexed by $t+1$. We denote by $y_{i,t+1}$ the educational effort level of individual $i$ (i.e. years of schooling) and by $\mathbf{y}_{t+1} = (y_{1,t+1}, ..., y_{n,t+1})'$ the population effort profile. Each agent $i$ selects an effort $y_{i,t+1} \geq 0$, and obtains a payoff $u_i(\mathbf{y}_{t+1}, g_t)$ that depends on the effort profile $\mathbf{y}_{t+1}$ and on the underlying network $g_t$ (formed at time $t$), in the following way:

$$u_i(\mathbf{y}_{t+1}, g_t) = (a_{i,t,t+1} + \eta_t + \varepsilon_{i,t+1})\, y_{i,t+1} - \frac{1}{2}y_{i,t+1}^2 + \phi \sum_{j=1}^{n} g_{ij,t}\, y_{i,t+1} y_{j,t+1} \qquad (1)$$

where $\phi > 0$. Two key aspects characterize the utility function $u_i(\mathbf{y}_{t+1}, g_t)$ of individual $i$. There is the idiosyncratic exogenous part $(a_{i,t,t+1} + \eta_t + \varepsilon_{i,t+1})\, y_{i,t+1} - \frac{1}{2}y_{i,t+1}^2$ and the endogenous peer effect aspect $\phi \sum_{j=1}^{n} g_{ij,t,t}\, y_{i,t+1} y_{j,t+1}$. In (1), $\eta_t$ denotes the unobservable network characteristics and $\varepsilon_{i,t+1}$ is an error term, meaning that there is some uncertainty in the benefit part of the utility function. There is also an ex ante *idiosyncratic heterogeneity*, $a_{i,t,t+1}$, which is assumed to be deterministic, perfectly *observable* by all individuals in the network and corresponds to the observable characteristics of individual $i$ (like e.g. sex, race, parental education, etc.) and to the observable average characteristics of individual $i$'s best

---

[4]See Goyal (2007), Jackson (2008), Jackson and Zenou (2013) for an overview on the theory of social networks.

friends, i.e. average level of parental education of $i$'s friends, etc. (contextual effects).[5] To be more precise, $a_{i,t,t+1}$ can be written as:

$$a_{i,t,t+1} = \sum_{m=1}^{M} \beta_m x_{i,t,t+1}^m + \frac{1}{g_{i,t}} \sum_{m=1}^{M} \sum_{j=1}^{n_r} g_{ij,t} \, x_{j,t,t+1}^m \gamma_m \tag{2}$$

where $x_{i,t,t+1}^m$ is a set of $M$ variables accounting for observable differences in individual characteristics of individual $i$, $\beta_m, \gamma_m$ are parameters and $g_{i,t} = \sum_{j=1}^{n} g_{ij,t}$ is the total number of friends individual $i$ has. The benefits from the utility are given by $(a_{i,t,t+1} + \eta_t + \varepsilon_{i,t+1}) \, y_{i,t+1}$ and are increasing in own educational effort $y_{i,t+1}$. In this first part, there is also a cost of providing educational effort, $\frac{1}{2}y_{i,t+1}^2$, which is also increasing in effort $y_{i,t+1}$. The second part of the utility function is: $\phi \sum_{j=1}^{n} g_{ij,t} \, y_{i,t+1}y_{j,t+1}$, which reflects the influence of friends' behavior on own action. The peer effect component is also heterogeneous, and this *endogenous heterogeneity* reflects the different locations of individuals in the friendship network and the resulting effort levels.

To summarize, when individual $i$ exerts some effort in education, the benefits of the activity depends on individual characteristics $a_{i,t,t+1}$, some network characteristics $\eta_t$ and on some random element $\varepsilon_{i,t+1}$, which is specific to individual $i$. In other words, $a_{i,t,t+1}$ is the observable part (by the econometrician) of $i$'s characteristics while $\varepsilon_{i,t+1}$ captures the unobservable characteristics of individual $i$. Note that the utility (1) is concave in own decisions, and displays decreasing marginal returns in own effort levels. In sum,

$$u_{i,t+1}(\mathbf{y}_{t+1}, g_t) = \underbrace{(a_{i,t,t+1} + \eta_t + \varepsilon_{i,t+1}) \, y_{i,t+1}}_{\text{Benefits from own effort}} \underbrace{-\frac{1}{2}y_{i,t+1}^2}_{\text{Costs}} + \underbrace{\phi \sum_{j=1}^{n} g_{ij,t} \, y_{i,t+1}y_{j,t+1}}_{\text{Benefits from friends' effort}}$$

**Nash equilibrium** We now characterize the Nash equilibrium of the game where agents choose their effort level $y_{i,t+1} \geq 0$ simultaneously. At equilibrium, each agent maximizes her utility (1). The corresponding first-order conditions are:

$$\frac{\partial u_{i,t+1}(\mathbf{y}_{t+1}, g_t)}{\partial y_{i,t+1}} = a_{i,t,t+1} + \eta_t + \varepsilon_{i,t+1} - y_{i,t+1} + \phi \sum_{j=1}^{n} g_{ij,t} \, y_{j,t+1} = 0.$$

Therefore, we obtain the following best-reply function for each $i = 1, ..., n$:

$$y_{i,t+1} = \phi \sum_{j=1}^{n} g_{ij,t} \, y_{j,t+1} + a_{i,t,t+1} + \eta_t + \varepsilon_{i,t+1} \tag{3}$$

---

[5]Observe that $a_{i,t,t+1}$ is indexed by both $t$ and $t+1$ because some characteristics (like e.g. sex, race, parental education, etc.) do not change over time while others (like age) do.

where $a_{i,t,t+1}$ is given by (2). Denote by $\mu_1(\mathbf{G}_t)$ the spectral radius of $\mathbf{G}_t$. We have:[6]

**Proposition 1** *If $\phi\mu_1(\mathbf{G}_t) < 1$, the peer effect game with payoffs (1) has a unique Nash equilibrium in pure strategies given by (3)*

We can calculate the equilibrium utility of the Nash equilibrium. Denote by $\mathbf{I}$ the identity matrix.

**Definition 1** *Given a vector $\boldsymbol{\alpha} \in \mathbb{R}^n_+$, and $\phi \geq 0$ such that $\phi\mu_1(\mathbf{G}_r) < 1$, we define the Katz-Bonacich vector of $\boldsymbol{\alpha}$-weighted centrality of parameter $\phi$ in network $g$ as:*

$$\mathbf{b}_{\boldsymbol{\alpha}}(g, \phi) = (\mathbf{I} - \phi\mathbf{G})^{-1} \boldsymbol{\alpha} = \sum_{p=0}^{+\infty} \phi^p \mathbf{G}^p \boldsymbol{\alpha} \tag{4}$$

The $i$th row of vector $\mathbf{b}_{\boldsymbol{\alpha}}(g, \phi)$ is denoted by $b_{i,\boldsymbol{\alpha}}(g, \phi)$. Denote $\alpha_{i,t,t+1} = a_{i,t,t+1} + \eta_t + \varepsilon_{i,t+1}$. Using the results of Ballester et al. (2006) and Calvo-Armengol et al. (2009), it is straightforward to show that the equilibrium educational effort given by (3) is

$$\mathbf{y}^*_{t+1} = \mathbf{b}_{\boldsymbol{\alpha}_{t,t+1}}(g_t, \phi) \tag{5}$$

so that each individual $i$ provide equilibrium effort

$$y^*_{i,t+1} = b_{i,\boldsymbol{\alpha}_{t,t+1}}(g_t, \phi)$$

## 3.2 First-stage: Friendship choices

In the empirical analysis, using data from adolescent friendships in the US, we will mainly test the second stage of this game, especially equation (3), and thus determine how strong are peer effects in educational choices by estimating the magnitude of $\phi$. In other words, we will take $g_{ij,t}$ as given and estimate $\phi$. However, in order to better understand our results, we would like to develop the first stage of the model and to explain how students make friendship relationships.

To determine the friendship choices, we need to calculate the equilibrium utility. The optimal effort of each student $i$ is given by (3). If we plug in (3) into (1), we easily obtain:

$$
\begin{aligned}
u_i(\mathbf{y}^*_{t+1}, g_t) &= (a_{i,t,t+1} + \eta_t + \varepsilon_{i,t+1})\, y^*_{i,t+1} - \frac{1}{2}y^{*2}_{i,t+1} + \phi\sum_{j=1}^{n} g_{ij,t}\, y^*_{i,t+1}y^*_{j,t+1} \\
&= \frac{1}{2}\left(y^*_{i,t+1}\right)^2 = \frac{1}{2}\left[b_{i,\boldsymbol{\alpha}_{t,t+1}}(g_t, \phi)\right]^2
\end{aligned}
$$

---

[6] All proofs of the theory can be found in Appendix 1.

Each individual $i$ has to pay a cost $c_i$ when creating a link. We assume that when a link is formed both students agreed (i.e. mutual consent) and both pay a cost. For example, if a link is created between $i$ and $j$, then both $i$ and $j$ have to better off than not creating this link and $i$ will pay $c_i$ and $j$ will pay $c_j$. We follow here Jackson and Wolinsky (1996) as it is the standard way of creating links. As a result, in the first stage the utility is

$$u_i(g_t, \phi) = \frac{1}{2} \left[ b_{i, \boldsymbol{\alpha}_{t, t+1}}(g_t, \phi) \right]^2 - c_i \sum_{j=1}^{n} g_{ij,t} \tag{6}$$

**Network stability**  In games played on a network, individuals payoffs depend on the network structure. In our case, this dependency is established in expression (6), that encompasses both the benefits and costs attributed to an individual given his/her position in the network of relationships. Any equilibrium notion introduces some stability requirements. The notion of pairwise-stability, introduced by Jackson and Wolinsky (1996), provides a widely used solution concept in networked environments. Let us now define this concept.

**Definition 2** *A network $g$ is **pairwise stable** if and only if:*

*(i) for all $ij \in g_t$, $u_i(g_t, \phi) \geq u_i(g_t - ij, \phi)$ and $u_j(g_t, \phi) \geq u_j(g_t - ij, \phi)$*

*(ii) for all $ij \notin g_t$, if $u_i(g_t, \phi) < u_i(g_t + ij, \phi)$ then $u_j(g_t, \phi) > u_j(g_t + ij, \phi)$.*

In words, a network is pairwise-stable if $(i)$ no player gains by cutting an existing link, and $(ii)$ no two players not yet connected both gain by creating a direct link with each other. Pairwise-stability thus only checks for one-link deviations.[7] It requires that any mutually beneficial link be formed at equilibrium but does not allow for multi-link severance. We will use throughout this equilibrium concept. Thus, network $g_t$ is an equilibrium network whenever it is pairwise stable.

**Two-stage equilibrium**  As it is well-known (Jackson and Wolinsky, 1996), it is extremely difficult to characterize all the pairwise stable equilibria of a network formation game with $n$ players where all agents simultaneously decide their link creation and deletion choices. This is why for example some recent papers have used a dynamic network formation model where only one person at a time decide to create or delete a link (see, e.g. Ehrhardt et al., 2006; König et al., 2012). Because here we have a network of friendship

---

[7]This weak equilibrium concept is often interpreted as a necessary conditions for stronger stability concepts.

that we observe at only one (or two) points in time, we will allow all individuals to create and delete simultaneously as in Jackson and Wolinsky (1996). As stated above, we are not able to characterize all equilibria, so we would like to provide a simple example to show how the two-stage equilibrium is calculated. This will help us for the intuition of the empirical results.

Consider the following network at time $t$
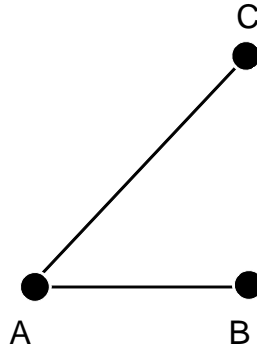


Figure 1: Pairwise stable star-shaped network

Its adjacency matrix $\mathbf{G}_t$ is given by (individual $A$ corresponds to the first row, $B$ to the second row, $C$ to the third row):

$$\mathbf{G}_t = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

The largest eigenvalue is $\sqrt{2}$, and thus the condition for a unique Nash-equilibrium is $\phi\mu_1(\mathbf{G}_r) < 1$, i.e.

$$\phi < \frac{1}{\sqrt{2}} = 0.707 \tag{7}$$

Ex ante heterogeneity is given by the vector $\boldsymbol{\alpha}_{t,t+1}^{\mathrm{T}} = \begin{pmatrix} \alpha_{A,t,t+1} & \alpha_{B,t,t+1} & \alpha_{C,t,t+1} \end{pmatrix}$. The vector of weighted Bonacich centrality is then equal to:[8]

$$\mathbf{b}_{\boldsymbol{\alpha}_{t,t+1}}\left(g_t, \phi\right) = \frac{1}{\left(1 - 2\phi^2\right)} \begin{pmatrix} \alpha_{A,t,t+1} + \phi\,\alpha_{B,t,t+1} + \phi\,\alpha_{C,t,t+1} \\ \phi\,\alpha_{A,t,t+1} + \left(1 - \phi^2\right)\alpha_{B,t,t+1} + \phi^2\,\alpha_{C,t,t+1} \\ \phi\,\alpha_{A,t,t+1} + \phi^2\,\alpha_{B,t,t+1} + \left(1 - \phi^2\right)\alpha_{C,t,t+1} \end{pmatrix}$$

---

[8] Observe that (7) guarantees that $1 - 2\phi^2 > 0$.

This gives the following equilibrium utilities

$$\mathbf{u}_t(g_t) = \frac{1}{2\left(1 - 2\phi^2\right)^2} \left( \begin{array}{c} \left(\alpha_{A,t,t+1} + \phi\,\alpha_{B,t,t+1} + \phi\,\alpha_{C,t,t+1}\right)^2 \\ \left[\phi\,\alpha_{A,t,t+1} + \left(1 - \phi^2\right)\alpha_{B,t,t+1} + \phi^2\,\alpha_{C,t,t+1}\right]^2 \\ \left[\phi\,\alpha_{A,t,t+1} + \phi^2\,\alpha_{B,t,t+1} + \left(1 - \phi^2\right)\alpha_{C,t,t+1}\right]^2 \end{array} \right) - \left( \begin{array}{c} 2c_A \\ c_B \\ c_C \end{array} \right)$$

Given utility (6), the following proposition gives under which condition the network described in Figure 1 is a pairwise stable equilibrium.

**Proposition 2** *Assume that $c_A = c_B = c_C = c$ and $\alpha_{A,t,t+1} = \alpha_{B,t,t+1} = \alpha_{C,t,t+1} = 1$. Then, if $\phi < 0.5$, the condition that guarantees that the star-shaped network displayed in Figure 1 is pairwise stable is:*

$$\frac{\phi\left(2 - 4\phi^2 - \phi\right)\left(1 + \phi\right)^2}{2\left(1 - \phi - 2\phi^2\right)^2\left(1 - 2\phi^2\right)^2} < c < \frac{\phi\left(2 + 5\phi - 4\phi^3\right)}{2\left(1 - 2\phi^2\right)^2} \tag{8}$$

Even in the simple star-shaped network described in Figure 1, it is quite cumbersome to determine the conditions under which this network is pairwise stable (see the proof of Proposition 2 in Appendix 1 where we also give the general conditions for any $c_A$, $c_B$ and $c_C$ and any $\alpha_{A,t,t+1}$, $\alpha_{B,t,t+1}$ and $\alpha_{C,t,t+1}$). It is because each individual, when deciding to forming a new link or deleting an existing one, has to anticipate the impact on the educational efforts of all individuals in the network (i.e. their weighted Bonacich centralities).

If we consider the general model with any $c_A$, $c_B$ and $c_C$ and any $\alpha_{A,t,t+1}$, $\alpha_{B,t,t+1}$ and $\alpha_{C,t,t+1}$, the conditions for this star-shaped network to be pairwise stable are given in the proof of Proposition 2 in Appendix 1 and are (18), (20) and (21).

# 4   Data description

As stated above, we will like to test the second stage of this game, i.e., Proposition 1, especially equation (3), and thus determine how strong are peer effects in educational choices by estimating the magnitude of $\phi$. In other words, we will take $g_{ij,t}$ as given and estimate $\phi$.

Our analysis is made possible by the use of a unique database on friendship networks from the National Longitudinal Survey of Adolescent Health (AddHealth).[9]

---

[9]This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies

The AddHealth survey 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 (wave I). Every pupil attending the sampled schools on the interview day is asked to compile a questionnaire (in-school data) containing questions on respondents' demographic and behavioral characteristics, education, family background and friendship. This sample contains information on roughly 90,000 students. A subset of adolescents selected from the rosters of the sampled schools, about 20,000 individuals, is then asked to compile a longer questionnaire containing more sensitive individual and household information (in-home and parental data). Those subjects are interviewed again in 1995–96 (wave II), in 2001–2 (wave III), and again in 2007-2008 (wave IV).[10]

From a network perspective, the most interesting aspect of the AddHealth data is the friendship information, which is based upon actual friends nominations. Indeed, pupils were asked to identify their best friends from a school roster (up to five males and five females).[11] This information is collected at wave I and one year after, i.e. at wave II. As a result, one can reconstruct the whole geometric structure of the friendship networks and their evolution at least in the short run. Such a detailed information on social interaction patterns allows us to measure the peer group more precisely than in previous studies by knowing exactly who nominates whom in a network (i.e. who interacts with whom in a social group) and by distinguishing between *strong ties* and *weak ties*. Indeed, we say that two students have a *strong-tie* relationship if they have nominated each other in both waves (i.e. in Wave I in 1994-95 and in Wave II in 1995-96). On the contrary, two students have a *weak-tie* relationship if they have nominated each other in one wave only).

We also exploit the *nomination order* to weight differently the influence of each peer within peer groups, i.e. we consider heterogenous peer effects. To the best of our knowledge this information has not been used before. More specifically, we exploit the directed nature of the nominations data and we denote a link from $i$ to $j$ as $g_{ij,r,t} = 1$ if $i$ has nominated $j$

as his/her friend in network $r$,[12] and $g_{ij,r,t} = 0$, otherwise.[13] We then weight each individual contact using a function, which is linearly decreasing with the corresponding order in the nomination list and also accounts for the total number of nominations made by the individual. Each non-zero entry $w_{ij,r,t}$ of the adjacency matrix $\mathbf{G}_{r,t}$ is:

$$w_{ij,r,t} = 1 - \frac{(\vartheta - 1)}{g_{i,r,t}} \tag{9}$$

where $\vartheta$ denotes the order of nomination given by individual $i$ to friend $j$ in his/her nomination list while $g_{i,r,t} = \sum_{j=1}^{n} g_{ij,r,t}$ is the total number of nominations made by individual $i$. By doing so, we allow for the fact that each individual can be affected differently by different peers within his/her peer group. For example, imagine that individual 1 has nominated three friends (i.e. $g_{i,r} = 3$), say first friend 2, then 4 and then 3. In that case, $\vartheta = 1$ for individual 2, $\vartheta = 2$ for individual 4 and $\vartheta = 3$ for individual 3. We will have the following weights: $w_{12,r,t} = 1$, $w_{14,r,t} = 2/3$ and $w_{13,r,t} = 1/3$ and thus, in the first row of the adjacency matrix $\mathbf{G}_{r,t}$ (for individual 1), there will be a 0 for individuals that 1 has not nominated and a 1, 2/3 and 1/3 for individuals 2, 4 and 3, respectively.

In order to verify that the order of nomination makes sense, in Figure 2, we display the percentage of deleted links from Wave I to Wave II by order of nomination. We see that, among all the first-nominated friends ("first nomination"), less than 40 percent of these friendship links are severed between Wave I and Wave II while, for individuals who have been nominated fifth ("fifth nomination"), this number is 77 percent. More generally, there is an increasing relationship between the percentage of deleted links and the order of nomination, indicating that the lower is the friend's ranking in terms of nomination, the weaker is the relationship between the two individuals and the more likely this relationship will not last more than one wave. This confirms that the order of nomination is not random and that weak-tie relationships are more likely to be severed between the two waves than strong ties.

<center>[<em>Insert Figure 2 here</em>]</center>

By matching the identification numbers of the friendship nominations to respondents' identification numbers, one can also obtain information on the characteristics of nominated friends. In addition, the longitudinal structure of the survey provides information on both respondents and friends during the adulthood. In particular, the questionnaire of wave

---

[12]In the data, we have different networks. This is why we need to index all variables by $r$.

[13]As highlighted by Wasserman and Faust (1994), centrality indices for directional relationships generally focus on choices made.

IV contains detailed information on the highest education qualification achieved. We measure education attainment in completed years of full time education.[14] Social contacts (i.e. friendship nominations) are, instead, collected in Wave I (and II).

Our final sample of in-home Wave I students (and friends) that are followed over time and have non-missing information on our target variables both in Waves I, II and IV consists of 1,749 individuals distributed over 183 networks. This large reduction in sample size with respect to the original sample is mainly due to the network construction procedure, roughly 20 percent of the students do not nominate any friends and another 20 percent cannot be correctly linked. In addition, we exclude networks composed by 2-3 individuals, those with more than 400 members and individuals who are not followed in Wave IV.[15] The mean and the standard deviation of network size are roughly 9.5 and 15, respectively.[16]

Table 1 gives the definition of the variables used in our study[17] as well as their descriptive statistics. Among the individuals selected in our sample, 53 percent are female and 19 percent are blacks. The average parental education is high-school graduate. Roughly 10 percent have parents working in a managerial occupation, another 10 percent in the office or sales sector, 20 percent in a professional/technical occupation, and roughly 30 percent have parents in manual occupations. More than 70 percent of our individuals come from household with two married parents and from an household of about four people on average. At Wave IV, 42 percent of our adolescents are now married and nearly half of them (43 percent) have at least a son or a daughter. The mean intensity in religion practice slightly decreases during the transition from adolescence to adulthood. On average, during their teenage years, our individuals felt that adults care about them and had a good a good relationship teachers. Roughly, 30 percent of our adolescents were highly performing individuals at school, i.e. had

---

[14]More precisely the Wave IV questionnaire asks the highest education qualification achieved (distinguishing between 8th grade or less, high school, vocational/technical training, bachelor's degree, graduate school, master's degree, graduate training beyond a master's degree, doctoral degree, post baccalaureate professional education). Those with high school and above qualification are also asked to report the exact year when the highest qualification was achieved. Such an information allows us to construct a reliable measure of each individual's completed years of education.

[15]We do not consider networks at the extremes of the network size distribution (i.e. composed by 2-3 individuals or by more than 400) because peer effects can show extreme values (too high or too low) in these edge networks.

[16]On average, these adolescents declare having 2.49 friends with a standard deviation of 1.78.

[17]Information at the school level, such as school quality and teacher/pupil ratio is also available but we don't use it since our sample of networks are within schools and we use fixed network effects in our estimation strategy.

the highest mark in mathematics.

[*Insert Table 1 here*]

Before we start our empirical analysis, we would like to look at simple correlations between the education attainment of an individual and the friends that s/he has nominated when s/he was adolescent at school in Wave I.[18] Figure 3 documents this correlation by differentiating between *direct* best friends ($k = 1$), friends of friends ($k = 2$), etc. In other words, the length of the shortest path distance between two individuals is $k$. We differentiate between unweighted networks (i.e. when the order of nomination is not taken into account) and weighted networks (i.e. when the order of nomination is taken into account). One clearly sees that the correlation curve is decreasing with the length of distance in friendship and is steeper when weights on the order of nomination of friends are included. This indicates that direct friends' educational outcomes have a higher correlation with own education outcome than indirect friends' educational outcome and that these relationships are stronger when the order of nominations is taken into account. For example, it can be seen from Figure 3 that the correlation in education between an individual and his/her direct friend is twice as high as the one between an individual and his/her indirect friend of length 8 (i.e. $k = 8$).

[*Insert Figure 3 here*]

# 5   Empirical analysis

## 5.1   Empirical model

Let $N_r = \{1, \ldots, n_r\}$ be a finite set of agents in network $r$ ($r = 1, ..., \bar{r}$), where $\bar{r}$ is the total number of networks in the sample ($\bar{r} = 183$ in our dataset), $n_r$ is the number of individuals in the $r$th network, and $n = \sum_{r=1}^{\bar{r}} n_r$ be the total number of individuals ($n = 1,749$ in our dataset). We keep track of social connections by a matrix $\mathbf{G}_r = \{g_{ij,r}\}$, where $g_{ij,r} = 1$ if $i$ and $j$ are direct friends, and $g_{ij,r} = 0$, otherwise. For $i = 1, \cdots, n_r$ and $r = 1, \cdots, \bar{r}$, our empirical model, which, using (2) and indexing by network $r$, corresponds exactly to (3) in the theoretical model, can be written as:

$$y_{i,r,t+1} = \phi \sum_{j=1}^{n_r} g_{ij,r,t} y_{j,r,t+1} + x'_{i,r,t,t+1}\delta + \frac{1}{g_{i,r,t}} \sum_{j=1}^{n_r} g_{ij,r,t} x'_{j,r,t,t+1}\gamma + \eta_{r,t} + \epsilon_{i,r,t+1}, \qquad (10)$$

---

[18]Data from Wave II show a similar graph.

where $y_{i,r,t+1}$ is the highest education level reached by individual $i$ at time $t+1$ who belonged to network $r$ at time $t$, where time $t+1$ refers to wave IV in 2007-2008 while time $t$ refers to wave I in 1994-95. Similarly, $y_{j,r,t+1}$ is the highest education level reached by individual $j$ at time $t+1$ who has been nominated as his/her friend by individual $i$ at time $t$ in network $r$. Furthermore, $x'_{i,r,t,t+1} = (x^1_{i,r,t,t+1}, \cdots, x^m_{i,r,t,t+1})'$ indicates the $M$ variables accounting for observable differences in individual characteristics of individual $i$ both at times $t$ (e.g. self esteem, mathematics score, quality of the neighborhood, etc.) and $t+1$ (marital status, age, children, etc.) of individual $i$. Some characteristics are clearly the same at times $t$ and $t+1$, such as race, parents' education, gender, etc. Also $g_{i,r,t} = \sum_{j=1}^{n} g_{ij,r,t}$ is the total number of friends individual $i$ has in network $r$ at time $t$. Finally, $\epsilon_{i,r}$'s are i.i.d. innovations with zero mean and variance $\sigma^2$ for all $i$ and $r$.

In the next two sections, to avoid too cumbersome notations, we omit the time index.

## 5.2   Identification strategy

The aim of our empirical analysis is twofold, $(i)$ to assess the presence of long-run peer effects in education and $(ii)$ to differentiate between the impact of weak ties and strong ties. The identification of peer effects ($\phi$ in model (10)) raises, however, different challenges.

In *linear-in-means* models, simultaneity in behavior of interacting agents introduces a perfect collinearity between the expected mean outcome of the group and its mean characteristics. Therefore, it is difficult to differentiate between the effect of peers' choice of effort and peers' characteristics that do impact on their effort choice (the so-called *reflection problem*; see Manski, 1993). Basically, the reflection problem arises because, in the standard approach, individuals interact in groups, that is individuals are affected by all individuals belonging to the same group and by nobody outside the group. In other words, groups completely overlap. In the case of social networks, instead, this is nearly never true since the reference group has individual-level variation. Formally, as shown by Bramoullé et al. (2009), social effects are identified (i.e. there is no reflection problem) if $\mathbf{I}$, $\mathbf{G}_r$ and $\mathbf{G}_r^2$ are linearly independent where $\mathbf{I}$ is the identity matrix and $\mathbf{G}_r^2$ keeps track of indirect connections of length 2 in network $r$. In other words, if $i$ and $j$ are friends and $j$ and $k$ are friends, it does not necessarily imply that $i$ and $k$ are also friends. Denote $\mathbf{X}_r = (x_{1,r}, \cdots, x_{n_r,r})'$ and $\mathbf{Y}_r = (y_{1,r}, \cdots, y_{n_r,r})'$. Then, because of these intransitivities, $\mathbf{G}_r^2 \mathbf{X}_r$, $\mathbf{G}_r^3 \mathbf{X}_r$, etc. are not collinear with $\mathbf{G}_r \mathbf{X}_r$ and they can therefore act as valid instruments. Take, for example, individuals $i$, $j$ and $k$ in network $r$ such that $g_{ij,r} = 1$ and $g_{jk,r} = 1$ but $g_{ik,r} = 0$. In that case, for individual $i$, the characteristics of peers of peers $\mathbf{G}_r^2 \mathbf{X}_r$ (i.e. $x_{k,r}$) is a valid instrument for peers' behavior $\mathbf{G}_r \mathbf{Y}_r$ (i.e. $y_{j,r}$) since $x_{k,r}$ affects $y_{i,r}$ only indirectly through its effect

on $y_{j,r}$ (distance 2). The architecture of social networks implies that these attributes will affect each individual outcome only through their effect on his/her friends' outcomes. Even in linear-in-means models, the Manski's (1993) reflection problem is thus eluded.[19] Peer effects in social networks are thus identified and can be estimated using 2SLS or maximum likelihood (Lee 2007; Calvó-Armengol et al., 2009; Lin, 2010).[20]

Although this setting allows us to solve the reflection problem, the estimation results might still be flawed because of the presence of *unobservable factors* affecting both individual and peer behaviors. It is indeed difficult to disentangle the endogenous peer effects from the correlated effects, i.e. effects arising from the fact that individuals in the same network tend to behave similarly because they face a common environment. If individuals are not randomly assigned into networks, this problem might originate from the possible sorting of agents. If the variables that drive this process of selection are not fully observable, potential correlations between (unobserved) network-specific factors and the target regressors are major sources of bias. A number of papers using network data have dealt with the estimation of peer effects with correlated effects (e.g., Clark and Loheac 2007; Lee 2007; Calvó-Armengol et al., 2009; Lin, 2010; Lee et al., 2010). This approach is based on the use of *network fixed effects* and extends Lee (2003) 2SLS methodology. Network fixed effects can be interpreted as originating from a two-step model of link formation where agents self-select into different networks in a first step and, then, in a second step, link formation takes place within networks based on observable individual characteristics only. An estimation procedure alike to a panel within group estimator is thus able to control for these correlated effects. One can get rid of the network fixed effects by subtracting the network average from the individual-level variables.[21] As detailed in the next section, this paper follows this approach.

Finally, one might question the presence of problematic unobservable factors that are

[19]These results are formally derived in Bramoullé et al. (2009) (see, in particular, their Proposition 3) and used in Calvó-Armengol et al. (2009) and Lin (2010). Cohen-Cole (2006) presents a similar argument, i.e. the use of out-group effects, to achieve the identification of the endogenous group effect in the linear-in-means model (see also Weinberg et al., 2004; Laschever, 2009; De Giorgi et al., 2010).

[20]More technical results can be found in Liu and Lee (2010). Liu et al. (2011) explicitly study the case of a non row-normalized adjacency matrix and provides the conditions on the parameters that guarantee the identification of peer effects (similarly to the conditions derived by Bramoullé et al., 2009, who derive them for the case of a row-normalized adjacency matrix).

[21]Bramoullé et al. (2009) also deal with this problem in the case of a row-normalized $\mathbf{G}_r$ matrix. In their Proposition 5, they show that if the matrices $\mathbf{I}$, $\mathbf{G}_r$, $\mathbf{G}_r^2$ and $\mathbf{G}_r^3$ are linearly independent, then by subtracting from the variables the network average (or the average over neighbors, i.e. direct friends), social effects are again identified and one can disentangle endogenous effects from correlated effects. In our dataset this condition of linear independence is always satisfied.

not network-specific, but rather individual-specific. In this respect, 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. Specifically, to control for differences in leadership propensity across adolescents, we include an indicator of *self-esteem* and an indicator of the *level of physical development* compared to peers, and we use *mathematics score* as an indicator of ability. Also, we attempt to capture differences in attitude towards education, parenting and more general social influences by including indicators of the student's school attachment, relationship with teachers, parental care and social inclusion.

In addition, we present an IV approach that uses as instruments only variables lagged in time to ensure that the instruments are not correlated with the contemporaneous error term. Observe that any unobserved source of heterogeneity that can be captured at the network level is already taken into account by the inclusion of network fixed effects.

## 5.3  Econometric methodology

Our econometric methodology follows closely Liu and Lee (2010). Let us expose this approach and highlight the modification that is implemented in this paper.

Let $\mathbf{Y}_r = (y_{1,r}, \cdots, y_{n_r,r})'$, $\mathbf{X}_r = (x_{1,r}, \cdots, x_{n_r,r})'$, and $\boldsymbol{\epsilon}_r = (\epsilon_{1,r}, \cdots, \epsilon_{n_r,r})'$. Denote the $n_r \times n_r$ sociomatrix by $\mathbf{G}_r = [g_{ij,r}]$, the row-normalized of $\mathbf{G}_r$ by $\mathbf{G}_r^*$, and the $n_r$-dimensional vector of ones by $\mathbf{l}_{n_r}$. Then model (10) can be written in matrix form as:

$$\mathbf{Y}_r = \phi \mathbf{G}_r \mathbf{Y}_r + \mathbf{X}_r^* \beta + \eta_r \mathbf{l}_{n_r} + \boldsymbol{\epsilon}_r, \tag{11}$$

where $\mathbf{X}_r^* = (\mathbf{X}_r, \mathbf{G}_r^* \mathbf{X}_r)$ and $\beta = (\delta', \gamma')'$.

For a sample with $\bar{r}$ networks, stack up the data by defining $\mathbf{Y} = (\mathbf{Y}_1', \cdots, \mathbf{Y}_{\bar{r}}')'$, $\mathbf{X}^* = (\mathbf{X}_1^{*\prime}, \cdots, \mathbf{X}_{\bar{r}}^{*\prime})'$, $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1', \cdots, \boldsymbol{\epsilon}_{\bar{r}}')'$, $\mathbf{G} = \mathrm{D}(\mathbf{G}_1, \cdots, \mathbf{G}_{\bar{r}})$, $\boldsymbol{\iota} = \mathrm{D}(\mathbf{l}_{n_1}, \cdots, \mathbf{l}_{n_{\bar{r}}})$ and $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_{\bar{r}})'$, where $\mathrm{D}(\mathbf{A}_1, \cdots, \mathbf{A}_K)$ is a block diagonal matrix in which the diagonal blocks are $m_k \times n_k$ matrices $\mathbf{A}_k$'s. For the entire sample, the model is

$$\mathbf{Y} = \mathbf{Z}\theta + \boldsymbol{\iota} \cdot \boldsymbol{\eta} + \boldsymbol{\epsilon}, \tag{12}$$

where $\mathbf{Z} = (\mathbf{GY}, \mathbf{X}^*)$ and $\theta = (\phi, \beta')'$.

We treat $\eta$ as a vector of unknown parameters. When the number of networks $\bar{r}$ is large, we have the incidental parameter problem. Let $\mathbf{J} = \mathrm{D}(\mathbf{J}_1, \cdots, \mathbf{J}_{\bar{r}})$, where $\mathbf{J}_r = \mathbf{I}_{n_r} - \frac{1}{n_r} \mathbf{l}_{n_r}' \mathbf{l}_{n_r}$. The network fixed effect can be eliminated by a transformation with $\mathbf{J}$ such that:

$$\mathbf{JY} = \mathbf{JZ}\theta + \mathbf{J}\boldsymbol{\epsilon}. \tag{13}$$

Let $\mathbf{M} = (\mathbf{I} - \phi\mathbf{G})^{-1}$. The equilibrium outcome vector $\mathbf{Y}$ in (12) is then given by the reduced form equation:

$$\mathbf{Y} = \mathbf{M}(\mathbf{X}^*\delta + \boldsymbol{\iota} \cdot \boldsymbol{\eta}) + \mathbf{M}\boldsymbol{\epsilon}. \tag{14}$$

It follows that $\mathbf{GY} = \mathbf{GMX}^*\beta + \mathbf{GM}\boldsymbol{\iota}\eta + \mathbf{GM}\boldsymbol{\epsilon}$. $\mathbf{GY}$ is correlated with $\boldsymbol{\epsilon}$ because $E[(\mathbf{GM}\boldsymbol{\epsilon})'\boldsymbol{\epsilon}] = \sigma^2 \mathrm{tr}(\mathbf{GM}) \neq 0$. Hence, in general, (13) cannot be consistently estimated by OLS.[22] If $\mathbf{G}$ is row-normalized such that $\mathbf{G} \cdot \mathbf{l}_n = \mathbf{l}_n$, where $\mathbf{l}_n$ is a $n$-dimensional vector of ones, the endogenous social interaction effect can be interpreted as an average effect. With a row-normalized $\mathbf{G}$, Lee et al. (2010) have proposed a partial-likelihood estimation approach for the estimation based on the transformed model (13). However, for this empirical study, we are interested in the *aggregate endogenous effect* instead of the *average effect*. Hence, row-normalization is not appropriate. Furthermore, we are also interested in the centrality of networks that are captured by the variation in row sums in the adjacency matrix $\mathbf{G}$. Row-normalization could eliminate such information. If $\mathbf{G}$ is not row-normalized as it is in this empirical study, the (partial) likelihood function for (13) could not be derived, and alternative estimation approaches need to be considered. Liu and Lee (2010) use an instrumental variable approach and propose different estimators based on different instrumental matrices, denoted here by $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\mathbf{Q}_3$. They first consider the 2SLS estimator based on the conventional instrumental matrix for the estimation of (13): $\mathbf{Q}_1 = \mathbf{J}(\mathbf{GX}^*, \mathbf{X}^*)$ (*finite-IVs 2SLS*). For the case that the adjacency matrix $\mathbf{G}$ is not row-normalized, Liu and Lee (2010) then propose to use additional instruments (IVs) $\mathbf{JG}\boldsymbol{\iota}$ and enlarge the instrumental matrix: $\mathbf{Q}_2 = (\mathbf{Q}_1, \mathbf{JG}\boldsymbol{\iota})$ (*many-IVs 2SLS*). The additional IVs of $\mathbf{JG}\boldsymbol{\iota}$ are based on the row sums of $\mathbf{G}$ (i.e. the outdegrees of a network) and thus use the information on centrality of a network. They show that those additional IVs could help model identification when the conventional IVs are weak and improve upon the estimation efficiency of the conventional 2SLS estimator based on $\mathbf{Q}_1$. However, the number of such instruments depends on the number of networks. If the number of networks grows with the sample size, so does the number of IVs. The 2SLS could be asymptotic biased when the number of IVs increases too fast relative to the sample size (see, e.g., Bekker, 1994; Bekker and van der Ploeg, 2005; Hansen et al., 2008). Liu and Lee (2010) have shown that the proposed many-IV 2SLS estimator has a properly-centered asymptotic normal distribution when the average group size needs to be large relative to the number of networks in the sample. As detailed in Section 4, in this empirical study, we have a number of small networks. Liu and Lee (2010) also propose a bias-correction

---

[22]Lee (2002) has shown that the OLS estimator can be consistent in the spatial scenario where each spatial unit is influenced by many neighbors whose influences are uniformly small. However, in the current data, the number of neighbors are limited, and hence that result does not apply.

procedure based on the estimated leading-order many-IV bias: $\mathbf{Q}_3$ (*bias-corrected 2SLS*). The bias-corrected many-IV 2SLS estimator is properly centered, asymptotically normally distributed, and efficient when the average group size is sufficiently large. It is thus the more appropriate estimator in our case study (see Liu and Lee, 2010, for a detailed derivation and an analysis of the asymptotic properties of the different estimators).[23]

In this paper, we use these estimators and also implement a modification of this approach, which takes advantage of the longitudinal structure of our data. The exact equivalent of (10) can be written in matrix form as:

$$\mathbf{Y}_{r,t+1} = \phi \mathbf{G}_{r,t} \mathbf{Y}_{r,t+1} + \mathbf{X}_{r,t}^* \beta_1 + \mathbf{X}_{r,t+1}^* \beta_2 + \boldsymbol{\eta}_{r,t} \mathbf{l}_{n_r} + \boldsymbol{\epsilon}_{r,t+1}.$$

Our modification of the IV approach proposed by Liu and Lee (2010) consists in including in the different instrumental matrices only values lagged in time (i.e. observed in wave I). So, for instance the first instrumental matrix for the finite-IVs 2SLS will thus be: $\mathbf{Q}_1' = \mathbf{J}(\mathbf{G}_t \mathbf{X}_t^*, \mathbf{X}_t^*)$. Such a strategy should ensure that the instruments are not correlated with the contemporaneous (wave IV) error term $\epsilon_{t+1}$, thus strengthening our identification strategy.

# 6    Estimation results

Table 2 (columns 2 and 3) collects the estimation results of model (10) when using the different estimators discussed in the previous section, without using the information on the nomination order, i.e. all nominated friends receive the same weight equals to 1. The first column shows the results when using the traditional set of instruments whereas, in the second column, the instrumental set contains only variables lagged in time.

As explained above, for the estimation of $\phi$, we pool all the networks together by constructing a block-diagonal network matrix with the adjacency matrices from each network on the diagonal block. Hence we implicitly assume that the $\phi$ in the empirical model is the same for all networks. The difference between networks is controlled for by network fixed effects. Indeed, the estimation of $\phi$ for each network might be difficult (in terms of precision) for the small networks. Furthermore, as stated above, it is a crucial empirical concern to control for unobserved network heterogeneity by using network fixed effects.

Proposition 1 requires that $\phi$ is in absolute value smaller than the inverse of the largest eigenvalue of the block-diagonal network matrix $\mathbf{G}_r$, i.e. $\phi < 1/\mu(\mathbf{G}_r)$. In our case, the largest eigenvalue of $\mathbf{G}_r$ is 3.70 . Furthermore our theoretical model postulates that $\phi \geq 0$.

---

[23]Liu and Lee (2010) also generalize this 2SLS approach to the GMM using additional quadratic moment conditions.

As a result, we can accept values within the range $[0, 0.280)$. Table 2 shows, in the second column, the results obtained when using our most extensive set of instruments and, in column 3, those produced when using as instruments only variables lagged in time. All our estimates of $\phi$ are within the acceptable parameter space $[0, 0.280)$ and are all significant. Looking across columns, it appears that the results are similar and only slightly higher in magnitude in the third column. This finding (incidentally) validates the empirical identification strategies used by Lin (2010) and Liu et al. (2012). Indeed, given the extensive set of controls available in the AddHealth, the inclusion of network fixed effects, and, most importantly, because friendship networks are quite small (see Section 4), the presence of uncaptured (troubling) individual unobserved within network characteristics is very unlikely. If these factors were at work, we should have found a substantial difference between the results in the second and third column in Table 2, since the latter controls for such influences.

$$[Insert\ Table\ 2\ here]$$

Looking now within each column, as explained above, in our case study with relatively small networks in the sample, the preferred estimator is the bias-corrected 2SLS one. Let us thus focus on the bias-corrected estimator. First, the effect of friends' education on own education is always significant and positive, i.e., there are *strong and persistent peer effects in education*. This shows that the "quality" of friends (in terms of future educational achievement) from high school has a positive and significant impact on own future education level, even though it might be that individuals who were close friends in 1994-1995 (Wave I) might not be friends anymore in 2007-2008 (Wave IV). In terms of magnitude, we find that a standard deviation increase in peers' aggregate years of education (roughly two more high-school graduate friends) translates into roughly a 10 percent increase of a standard deviation in the individual's education attainment (roughly 3.5 more months of education). This is a strong effect, especially given our long list of controls and the fact that friendship networks might have changed over time. The influence of peers at school seem to be carried over time.

When the information on the nomination order is exploited (Table 2, last two columns), thus allowing to weight differently best friends, the magnitude of the effects is higher. A standard deviation increase in peers' education attainment translates into roughly a 15 percent increase of a standard deviation in the individual's education attainment (roughly 6 more months of education).[24] This confirms the role of the order of nomination in friendships, as

---

[24]When $\mathbf{G}_r$ is weighted the largest eigenvalue is 2.59. We can thus accept values within the range $[0, 0.385)$. All our estimates of $\phi$ in Table 2 (last two columns) are within this parameter space.

already highlighted in Figure 2.

# 7 The mechanisms of peer effects in long-run educational outcomes

In the previous section, we have seen that being friend in high school with individuals who will obtain a college degree later in life has a positive impact on own education attainment. The mechanisms behind this result is, however, not very clear. In the present section, we would like to narrow down this mechanism by proposing different empirical tests.

## 7.1 Do peers later in high school matter more than peers earlier in high school?

The strength of peer effects may depend on when students become friends with other students. In this section, we would like to investigate if peers *later* in high school matter more than peers *earlier* in high school. Indeed, earlier peers in a student's career are important as they affect attitutes towards education (which are formed in the early years of high school) such as: how much students learn, how much human capital will be accumulated or how much students value achievement. Later peers in a student's career are important as they are the interacting friends when student decide to go to college. To test the relative magnitude of these effects, we split our sample between students who were in grades 7-9 and those who were in grades 10-12 in Wave I and estimate model (10) on these two sub-samples separately. Table 3 presents the empirical results for students in grades 7-9 in Wave I while Table 4 displays the results for students in grades 10-12 in Wave I. We consider again unweighted networks (first two columns) and weighted networks (last two columns). We find that the impact on educational outcomes is stronger between peers in higher grades (i.e. in the last years of high school) than in lower grades. In terms of magnitude, a standard deviation increase in peers' aggregate years of education for students in grade 10-12 translates into roughly a 13.5 percent increase of a standard deviation in the individual's education attainment (roughly 5.7 more months of education).

[*Insert Tables* 3 *and* 4 *here*]

This result helps us understand better how we might expect peers to matter in the long run. We have shown that peers later in high school might matter more than peers earlier

24

in high school. This is primarily due to two factors: peers in the last years of high school are the interacting friends when the choice to go to college is made and/or these are peers carried on from early years at school with whom friendship ties are stronger.

To better understand these mechanisms, we investigate the role of *weak* and *strong ties* in educational decisions and try to disentangle the *decision proximity* effect and the *strength of interaction* effect. Indeed, is it really the timing of friendship formation that is the crucial factor (decision proximity effect) or is it rather the frequency of interactions underlying common choices (strength of interaction effect)? On the one hand, peers in the last years of high school might be more relevant because they are the interacting friends when the choice to go to college is made. On the other hand, it might also be that peers carried on from early years at school with whom friendship ties are stronger that matter. And, how about the influence of the peers met in earlier years and lost during the school years? They might still be important if the choice to go to college derives from preferences and attitudes towards education developed early on in the school years as a results of friends interactions. This is what we want to investigate in the next section.

## 7.2   The role of strong ties

### 7.2.1   An empirical investigation

The structure of the Addhealth dataset allow us to shed some light on these issues by distinguishing different types of peers. For this purpose, we use the information on the nominations that is collected in Wave II. Indeed, one year after the first interview (Wave I), (the same) students are asked again to nominate their best friends, with a similarly structured questionnaire. We consider the students that were in grade 10-12 in Wave I and collect their nominations in Wave II. Of course, we lose those who were in grade 12 in Wave I, unless a grade is repeated. We define a *strong-tie* relationship between two students if they have nominated each other in both waves (i.e. in Wave I in 1994-95 and in Wave II in 1995-96) and a *weak-tie* relationship if they have nominated each other in one wave only.

We split friendship links among teenagers in three categories: friendship links observed only in Wave I (weak ties), friendship links observed only in Wave II (weak ties) and friendship links observed in both waves (strong ties). By doing so, we can disentangle between the *decision proximity* effect and the *strength of interaction* effect, mentioned above. Indeed, contrary to strong ties, all weak ties have a low strength of interaction. However, the weak ties who are only nominated in Wave II have a higher decision proximity than those only nominated in Wave I. Tables 5-7 show the estimation results of model (10) when different

types of peers are considered. We analyze weak ties where best friends have only been nominated in Wave I and not in Wave II (Table 5) for unweighted (first two columns) and weighted networks (last two columns), weak ties where best friends have only been nominated in Wave II and not in Wave I (Table 6) for unweighted (first two columns) and weighted networks (last two columns) and strong ties (Table 7) for unweighted (first two columns) and weighted networks (last two columns).

We find that weak ties' educational choices have no significant impact on individual education outcomes (years of schooling), regardless whether peers are interacting in lower or higher grades. On the contrary, we find that strong ties' educational choices have a positive and significant effect on own educational outcome and that this effect is stronger when the network is weighted by the order of nomination. In terms of magnitude, a standard deviation increase in aggregate years of education of peers nominated both in Waves I and II (strong ties) translates into roughly a 10 percent increase of a standard deviation in the individual's education attainment (roughly 4 more months of education).

This suggests that *strong ties* rather than *weak ties* matter for educational outcomes in the long run. If we look at Figure 1, this also indicates that peers who are nominated first or second are more likely to be friend in both waves and thus more likely to influence their friends' future choice of education. In other words, even though students tend to nominate more than one or two friends (on average, our adolescents declare having 2.49 friends on average with a standard deviation of 1.78; see Section 4), it is really the first or second nominated friend who seems to influence individual's educational choice.

[*Insert Tables* 5, 6 *and* 7 *here*]

To better understand this result, Figure 4 provides a simple example. If we look at the networks on the left-hand side of the figure, one can see that, between $t$ (Wave I) and $t + 1$ (Wave II), individual $A$ has kept his/her friendship link with $B$, severed his/her friendship link with $D$, and created a new friendship link with $C$. According to our definitions, individual $B$ holds a strong tie link with $A$ while individuals $D$ and $C$ hold a weak tie link with $A$. The table on the right-hand side of the figure summarizes this by also highlighting the role of the *decision proximity* versus the *strength of interaction* effect. This allows us to differentiate between the two weak ties $D$ and $C$ who both have a low strength of interaction with $A$ (they only interact with $A$ in one wave) but individual $D$ (first column) has a low decision proximity (since he/she is not anymore friend with $A$ in Wave II) while individual $C$ (last column) has a high decision proximity (since he/she is friend with $A$ in wave II). The empirical results shown in Tables 5-7 show that what matters most for peer effects is the

strength of interaction rather than decision proximity effect (see also the last row of Figure 4). Being a strong tie (which includes both effects) is crucial to understand the role of peer effects in future educational attainment. A weak tie has no significant influence on his/her peers even if she/he has been nominated as best friend in Wave II.

[*Insert Figure* 4 *here*]

We would like to better understand what does this mean from an economic viewpoint. If the *decision proximity* effect was at work, then it would mean that peers matter more later in the student's career in shaping whether the student decides to go to college is worthwhile in the end. This is not what we found. We found that it is the *strength of interaction* that matters, indicating that peers matter in a student's career because they affect how much they learn, the human capital accumulation, how much they value achievement. It also shapes social norms that accumulate over time, which affect years of schooling both directly and indirectly. To summarize, we have found that education outcomes of students who have been nominated by a given person in both Wave I (1994-1995) and Wave II (1995-1996) have a positive impact on his/her education outcomes (years of schooling) made later in life as reported in Wave IV (2007-2008). We do not know, however, if these students are still friends with this given person in Wave IV since we do not have this information. We also found that the education outcome of friends who have been nominated in one wave only (it does not matter which one) have no significant impact on years of schooling of this individual. One natural question is: what happens to our results if the peers groups are not measured correctly? In Section 8, we perform a robustness check by manipulating the friendship links in each network, showing that our results holds true even if we do not observe or imprecisely observe a (non-negligible) portion of each individual's social network.

### 7.2.2   A theoretical investigation

To better understand the importance of weak and strong ties, let us now propose a model based on that of Section 3. In this section, for the specific networks described in Figure 4, we would like to see under which condition a student keeps a friend between Wave I and Wave II and when he/she doesn't. The second stage (at time $t+1$) for the educational outcome is exactly as in Section 3.1. The first stage is divided in two sub-stages: Wave I (at time $t-1$) and Wave II (at time $t$). We illustrate our model using the networks described in Figure 5, which are the same networks displayed in Figure 4. We only focus on student $A$ (red node). In Wave I, this student $A$ has two friends $D$ and $B$. In Wave II, we observe that this student $A$ has two friends $C$ and $B$.
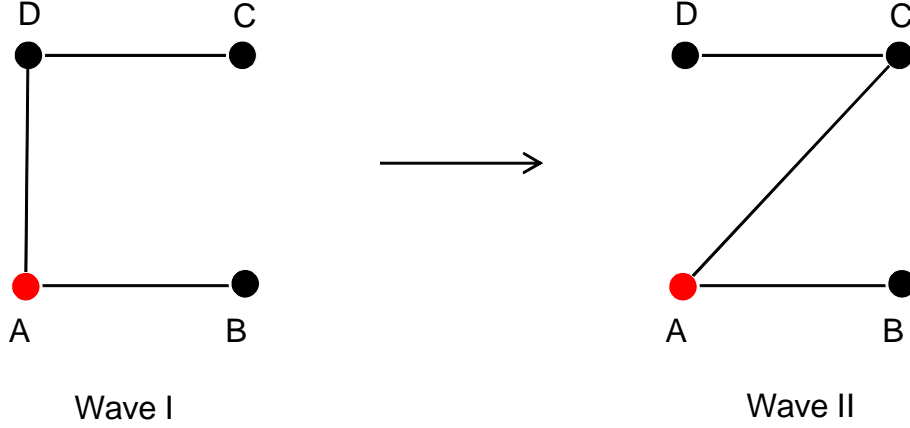
Figure 5: Strong and weak ties in link formation

Using our model, we would like to understand why, in Wave II, student $A$ first severs his/her link with $D$, then creates a link with $C$ and keeps his/her friendship link with $B$. Observe that we do not study under which conditions the networks described in Figure 5 are pairwise stable (as in Section 3) but we only focus on the behavior of student $A$. We have the following result.

**Proposition 3** *Assume* $\phi < 0.618$. *In Wave II, student $A$ will first severe the link with $D$, then creates a link with $C$ and keeps his/her friendship link with $B$ (Figure 5) if and only if*

$$\Gamma_1(\phi, \boldsymbol{\alpha}) < c_A < \min\{\Gamma_2(\phi, \boldsymbol{\alpha}), \Gamma_3(\phi, \boldsymbol{\alpha})\} \tag{15}$$

*where*

$$\Gamma_1(\phi, \boldsymbol{\alpha}) \equiv \frac{\left[(1-\phi)\alpha_A + \phi(1-\phi)\alpha_B + \phi\alpha_C + \phi\alpha_D\right]^2}{2\left(1 - \phi - 3\phi^2 + \phi^3\right)^2} - \Theta$$

$$\Gamma_2(\phi, \boldsymbol{\alpha}) \equiv \Theta - \frac{(\alpha_A + \phi\alpha_B)^2}{2\left(1 - \phi^2\right)^2}$$

$$\Gamma_3(\phi, \boldsymbol{\alpha}) \equiv \Theta - \frac{\left[(1-\phi^2)\alpha_A + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1 - 2\phi^2\right)^2}$$

*with*

$$\Theta \equiv \frac{\left[(1-\phi^2)\alpha_A + \phi(1-\phi^2)\alpha_B + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1 - 3\phi^2 + \phi^4\right)^2}$$

28

The results of this proposition are a little bit cumbersome and due to the fact that agents anticipate educational outcomes in the future (at time $t+1$). For this proposition to be true, it has to be that the cost of forming a link for $A$ is not too low so that he/she can sever the link $AD$ but also not too high so that he/she can creates the link $AC$ and keeps the one with $B$. These conditions also depend on $\alpha_A, \alpha_B, \alpha_C, \alpha_D$, which are the characteristics of all the individuals and $\phi$, the degree of interactions.

The key factors that affect these relationships are the costs of having friends, the strength of interactions in the network, and the characteristics of each individual involved in the friendship relationship. Individuals are more likely to severe a link with a friend (weak tie) if the cost of maintaining this link is too high and/or the $\alpha$'s are too different. On the contrary, students are more likely to keep friendship relationships (strong ties) if the reverse is true.

## 7.3 Weak versus strong ties

Our results in the previous sections suggest that the distinction between strong ties and weak ties is important for understanding long-run peer effects in education. In our analysis, we identified strong and weak ties using variations in network topology between two points in time, i.e. between Wave I and Wave II. We would now like to check if students select peers differently between the first and the second wave or if it is really that distinct types of peers (weak versus strong ties) matter differently. Table 8 investigates this issue but comparing the observable characteristics of peers who only appear in Wave I, those who only appear in Wave II and those who appear in both Waves. One can see that, in fact, there are no differences between these peers in terms of observable characteristics.

<center><em>[Insert Table 8 here]</em></center>

To further investigate this issue, we would like to test the fact that link formation is not different between different waves. For that, we pool the data for Wave I ($t = 1$) and Wave II ($t = 2$) and perform a *Chow test* based on the following regression:

$$g_{ij,r,t} = \alpha + \sum_{m=1}^{M} \beta_m |x_{i,r,t}^m - x_{j,r,t}^m| + \sum_{m=1}^{M} \gamma_m |x_{i,r,t}^m - x_{j,r,t}^m| \times \delta_{ij,r,t} + \epsilon_{ij,r,t}, \quad t = 1, 2, \quad (16)$$

where, as in (10), $g_{ij,r,t} = 1$ if there is a link between $i$ and $j$ belonging to network $r$ at time $t$, $x_{i,r,t}^m$ indicates the individual characteristic $m$ of individual $i$ in network $r$ at time $t$ and $\delta_{ij,r,t} = 1$ if a directed link $g_{ij,t}$ exists in Wave II, and zero otherwise. Here the variables that explain $g_{ij,r,t}$ are distances in terms of characteristics between students $i$ and $j$. For instance

<center>29</center>

two students might be friends because they both like to play football and watch TV. We want to test whether these effects are stable between Wave I and Wave II. From an economic viewpoint, not rejecting the null hypothesis of Chow test means that these effects, considered as a whole, does not change significantly after one year. Table 9 (Chow test p value) shows that, controlling for network fixed effects, we cannot reject the null hypothesis of $\gamma_m = 0$, $\forall m = 1, \ldots, M$. We have also tested separately the stability of each effect. Table 9 (second column) shows that most coefficients are not significant, meaning that the null hypothesis $\gamma_m = 0$ is not rejected so that there are no observable differences in link formation process between Waves I and II.

[*Insert Table* 9 *here*]

Observe that the aim of estimating (16) is not to test the model of Section 3.2 since its prediction is that not only the characteristics of the two persons involved in the friendship relationship matter but also the characteristics of all individuals of the network. Also the costs of forming links are an important determination of link formation but we do not have information about them. The aim of testing (16) is to show that the formation of a friendship link is not significantly different between Wave I and Wave II.

## 7.4   Short-run versus long run effects

So far, we have found that students nominate other students as their best friends but only their strong ties, i.e. students who are friends in both waves and who are more likely to be nominated first or second, influence them in their educational choice. It is really the intensity of the relationship that matters and only very close friends do impact on their friends' educational choices. Without looking at the difference between Waves I and II, Calvó-Armengol et al. (2009), using Addhealth data for Wave I only, have studied the *current* effect of peers on education. They show that peers do affect the current education activity (i.e. grades) of students. This could indicate that, in the short run, all peers affect current education (grades) of students while, in the long-run, only strong ties affect educational attainment of students.

To investigate further this issue, we would like now to oppose the long-run effects with short-run effects of peers on education by differentiating between strong ties and weak ties effects on school performance. To the best of our knowledge, nobody has looked at the effect of weak (Wave I *or* Wave II) and strong ties (Wave I *and* Wave II) on current educational outcomes (grades in Wave II). Is it true that also in the short run, only strong ties have an impact on current grades? For that, we estimate the short-run counterpart of equation (10):

$$y_{i,r,t} = \phi \sum_{j=1}^{n_r} g_{ij,r,t} y_{j,r,t} + x'_{i,r,t}\delta + \frac{1}{g_{i,r,t}} \sum_{j=1}^{n_r} g_{ij,r,t} x'_{j,r,t}\gamma + \eta_{r,t} + \epsilon_{i,r,t}, \qquad (17)$$

where $y_{i,r,t}$ is now the grade of student $i$ who belongs to network $r$ at time $t$, and $t$ refers to wave II instead of $y_{i,r,t+1}$, the highest education level reached by individual $i$ at time $t+1$ (i.e. in Wave IV in 2007-2008) who belonged to network $r$ at time $t$ (i.e. in Wave I in 1994-1995). The notation remain unchanged, which implies that we now deal with a traditional peer effects model where all the individual and peer group characteristics are contemporaneous (i.e. in Wave II in 1994-1995). As we did in our investigation on long run effects, we exploit variations in link formation in Waves I and II to differentiate between strong ties and weak ties. We will first estimate equation (17) for peers who are only friends in Wave I, i.e. $g_{ij,r,t} = 1$ in Wave I and $g_{ij,r,t} = 0$ in Wave II, then for peers who are only friends in Wave II, i.e. $g_{ij,r,t} = 0$ in Wave I and $g_{ij,r,t} = 1$ in Wave II, and finally for peers who are friends in both Waves I and II, i.e. $g_{ij,r,t} = 1$ in Wave I and $g_{ij,r,t} = 1$ in Wave II. The identification strategy remain unchanged with the difference that we cannot now use IV variables lagged in time. The school performance is measured using the respondent's scores received in Wave II in several subjects, namely English or language arts, history or social science, mathematics, and science. The scores are coded as 1=D or lower, 2=C, 3=B, 4=A. For each individual, we calculate an index of school performance using a standard principal component analysis. The final composite index (labeled as GPA index or grade point average index) is the first principal component score. It captures a general performance at school since it is positively and highly correlated to the scores in all subjects. Appendix 2 contains further details on this procedure.

We estimated equation (17) considering friends only in Wave I, friends only in Wave II and friends in both waves. The results are contained in Tables 10, 11 and 12, first column for unweighted networks and second column for weighted (by the nomination order) networks.

[*Insert Tables* 10, 11 and 12 *here*]

As for our results for the long run effects, we find stronger effects for weighted networks, suggesting that also, in the short run, friends nominated first or second seem to be more influential. However, *while in the long run, only strong ties matter, it appears that, in the short run, both weak and strong ties are important in determining a student's performance at school.* The magnitude of these effects is comparable: for weighted networks, a standard deviation increase in aggregate GPA of peers translates respectively into a 3.3 (for peers only in Wave I), a 4.4 (for peers only in Wave II) and a 3.2 (for peers in both waves) percent increase of a standard deviation in the individual's GPA. It means that peer effect for GPA is

31

substantially homogenous among different type of ties (see Tables 10, 11 and 12). Taken our analysis as a whole, our results suggest that, in the short run, all peers matter for education (i.e. grades) while, in the long run, only strong ties matter for future educational choices (i.e. years of schooling). In both cases, taking into account an heterogeneous effect of peers according to their nomination order might be important in shaping the magnitude of the peer effects.

# 8    Measurement errors in peer groups

We have shown that high-school friends have a long term impact on years of schooling if they are strong ties (i.e. friends in both Waves I and II). Observe that we have a measure of networks at two points in time (Waves I and II) and then an outcome measure in Wave IV (10 years after). It is possible that those students form other friendships before completing their schooling and that those extra peers are left out. This could explain why friends in the last grades of high school matter more than those made earlier. In any case, it may be that the set of relevant peers is mis-measured. Observe also that our identification and estimation strategies depend on the correct specification of network links. In particular, our identification strategy hinges upon non linearities in group membership, i.e. on the presence of *intransitive triads* (see Section 5.2).

In this section, we test the robustness of our results with respect to misspecification of network topology or relevant peers. So far, we have measured peer groups as precisely as possibly by exploiting the direction of the nomination data. However, there can be, for example, some "unobserved" network link that, if considered, would change the network topology and break some intransitivities in network links. In this section, a *falsification check* using simulated data will be performed. The general idea is as follows. Within the same network, let us randomly delete some friendship links and randomly create new links so that the total number of links stays the same. By doing so, we would like to see if our peer effect results stay the same or are affected by these changes. This will allow us to answer the following questions: Do our results change if some links are misspecified? To what extent? How many links need to be misspecified before our results become insignificant?

**A Numerical experiment**

We use a simulation approach to randomly change a certain percentage $p_r$ of links in each network $r$, one hundred times for each value of $p_r$ starting from 0 to 1 with a pace of 0.005. We thus draw one hundred network structures (samples) of size equal to the real one ($n = 1,749$) for each value of $p_r$, twenty thousand network structures in total. The desired

replacement rate is assumed to be the same for all networks, i.e. $p_r = p$.

The first empirical issue that we face in our procedure relates to the relationship between the strength of peer effects and network density. Because peer effects varies with network density (see, e.g. Calvó-Armengol et al., 2009), our numerical exercise needs to generate a constant number of links after deletion.

Let $L_r$, with cardinality $l_r$, be the set of existing links in network $r$ and $O_r$, with cardinality $o_r$, the set of non-existing links in the same network. The number of "possible changes" in network $r$ consistent with our constraint is $c_r = \min(l_r, o_r)$. In words, for each network $r$, we can exchange only a fraction of existing links with non existing links (and vice versa) if we want to maintain constant the total number of links in our network $r$ of a given size (network density). The percentage of randomly replaced links $p_r$ is thus calculated over the possibly interchangeable links (excluding overlapping), rather than over the total number of network links. Let $q_r$ be the theoretical number of links to replace in network $r$ when we want to remove a percentage $p_r$ of $c_r$. It is: $q_r = p_r \times c_r$.

The second empirical issue here is that this theoretical portion of links that we want to change may not correspond to a discrete number of links. For example, a replacement rate of 20 percent in a network with 7 possibly interchangeable links would imply that 1.4 links need to be changed. Do we swap one link (i.e. one existing into non existing and one non existing into existing at random) or two links (i.e. two couples)?

We rigorously implement this decision rule as follows.

Let $p_r \in (0,1)$ be our desired replacement rate in network $r$. In order to obtain a number of changes as close as possible to the desired one, the actual number of changes $s_r$ is:

$$s_r = \begin{cases} \lfloor q_r \rfloor & \text{if} \quad u > a \\ \lfloor q_r \rfloor + 1 & \text{if} \quad u < a \end{cases}$$

where $\lfloor q_r \rfloor$ denotes the largest integer smaller or equal than $q_r$ (the floor of $q_r$), $a = q_r - \lfloor q_r \rfloor$, and $u$ is a random extraction from a variable uniformly distributed on $(0,1)$.

To illustrate this point, let us consider a simple example where the network is given in Figure 6.
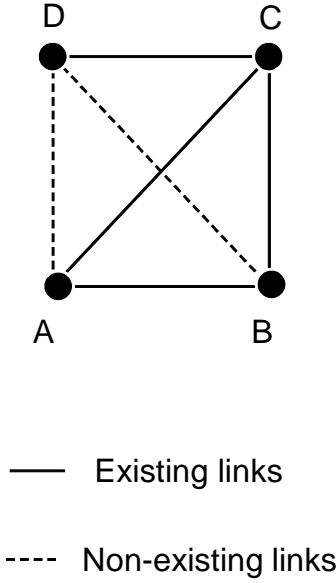
Figure 6: A simple example illustrating the simulation exercises

In Figure 6, the undirected network composed of 4 nodes, $A, B, C, D$, and 4 links $\{AB, AC, BC, CD\}$. In this situation, $l_r$ (the number of existing links in the network) is equal to 4, $o_r$ (the number of non-existing links. i.e. the non-existing links $AD$ and $BD$) is equal to 2 and therefore $c_r = 2$. Within the set of "possible changes"

$$\{(ABAD), (ABBD), (ACAD), (ACBD), (BCAD), (BCBD), (CDAD), (CDBD)\},$$

we can make at most two changes. This means that we can extract randomly just two couples out of eight. Now suppose that our desired replacement rate $p_r$ is 0.3 (30 percent), yielding to an actual replacement rate $q_r = 0.6$ ($q_r = p_r c_r = 0.3 \times 2$). In that case, $a = 0.6 - \lfloor 0.6 \rfloor = 0.6 - 0 = 0.6$. We have:

$$s_r = \begin{cases} \lfloor 0.6 \rfloor = 0 & \text{if} \quad u > 0.6 \\ \lfloor 0.6 \rfloor + 1 = 1 & \text{if} \quad u < 0.6 \end{cases}$$

At this stage, our algorithm draws $u$. If $u < 0.6$, then $s_r = 1$ and we will replace one link (i.e. we extract at random one couple), otherwise nothing will happen. Clearly, given that 0.6 is closer to 1 than to 0, the probability to extract $u < 0.6$ is higher than the probability to extract $u > 0.6$, as it is desired.[25]

---

[25]This algorithm has been written in Matlab. The code is available upon request.

**Simulated evidence**

Our link replacement procedure enables us to simulate different network structures ($\mathbf{G}$ matrices in model (10)) that differ from the real one by a given (increasing) number of misspecified links. As mentioned before, for each percentage of randomly replaced links, we draw 100 network structures (samples) of size and network density equal to the real one. We then estimate model (10) replacing the real $\mathbf{G}$ matrix with the simulated ones in turn, so that in total we estimate model (10) twenty thousand times for each type of estimator (see Section 5.3).

Figure 7 shows the results of our simulation experiment. The upper panel depicts the estimates of peer effects (model (10)), whereas the lower panel shows the $t-$statistics with 90 percent confidence bands. It appears that the higher the percentage of misspecified links, the wider is the range of the peer-effect estimates and the more often the $t-$statistics fail to reject the hypothesis of no effects. This means that the more the peer effects are mis-specified (in terms of links), the less effects they have on educational outcomes of their friends. This gives us confidence that students do declare their best friends and do not make mistakes or errors when nominating them. The crucial question now is what is the percentage of network structure misspecification over which peer effects become insignificant.

[*Insert Figure 7 here*]

Figure 8 plots the averages of the estimates of peer effects for each replacement rate with 90 percent confidence bands. Standard errors have been calculated assuming drawing independence and taking into account the variation between estimates for each replacement rate.[26] We find that peer effects remain statistically significant up to a percentage of randomly replaced (interchangeable) links of about 20 percent. This implies that even if we do not observe or imprecisely observe a portion of each individual's social network, our results on the existence of peer effects hold. The portion of network topology that can be misspecified is not extremely small.

[*Insert Figure 8 here*]

---

[26]Specifically, the standard error at each replacement rate, say $i$, is computed as follows:

$$\sigma_i = \sqrt{W_i + B_i}$$

where $W_i = \frac{1}{n}\sum_{j=1}^{n}\sigma_{ij}^2$ , $B_i = \frac{1}{n}\sum_{j=1}^{n}(\phi_{ij} - \bar{\phi}_i)^2$ , $\sigma_{ij}^2$ is the estimated variance of the *jth* estimator at the *ith* replacement rate, $\phi_{ij}$ is the *jth* estimate at the *ith* replacement rate and $\bar{\phi}_i$ is the mean across the $n$ estimates. In this experiment $n = 100$.

# 9 Concluding remarks and policy implications

This paper is about peer effects in education and we have tried to understand how peers met at school impact on educational choices (years of schooling) made later on in life. We have found that students who have a strong-tie relationship (i.e. who have been nominated more than once) with a given person have a positive impact on his/her education outcomes while those who have a weak-tie relationship (i.e. who have been nominated only once) do not. This means that the strength of interactions between students matters in a student's career because it affects how much they learn, how much they value achievement, the human capital accumulation, and how social norms are formed. We have also found that, in the short run, any relationship (whether it is a weak or a strong one) matters in influencing current grades.

The presence of peer effects provides opportunities for policies aiming at improving social welfare and increasing educational outcomes (Hoxby, 2000). If one wants to implement an effective education policy, it needs to internalize peer effects. For instance, education vouchers could lead to a more efficient human capital investment profile (see e.g., Epple and Romano 1998; Nechyba 2000). Our results suggest that one should give vouchers to students who generate most positive education spillovers. Using, for example, the AddHealth data, where information on students' grades and students' strong-tie relationships (a good proxy for positive spillovers) is provided, one could give vouchers to the students who generate the highest positive externalities in education. Furthermore, policies such as school desegregation, busing, magnet schools, Moving to Opportunity programs[27] could also be effective if the government understands the magnitude and nature of peer effects in student outcomes. In the present paper, we have shown that the strength of relationships matters in long-term educational outcomes. This may be an indication of good social norms where students influence each other over a long-term period. Since adolescent friendships are often made in school and, more precisely, in the classroom, one natural policy implication could be not to change the composition of the classroom over time, as it is often the case in the United States. Indeed, we believe that, in the last years of high school, it would not be efficient to move students to another class because, in that case, his/her strong ties will also be removed and it is unlikely that his/she is going to have strong-tie friends in the new environment. On the contrary, in earlier grades, moving a student to another class could have positive effects since this student has not yet formed strong-tie relationhsips and it is still possible for him/her to create such relationship in the new environment. This, of course, would be

---

[27]See Lang (2007) for an overview of these policies in the U.S.

efficient if peers have positive effects on each other. In the case of disruptive kids and negative peer effects, then the opposite should be done. This paper points toward the fact that friends made at school over a long period of time have an important impact of the decision to whether or not to go to college. This could be for the best but also for the worse.

# References

[1] Angrist, J. D. and A.B. Krueger (1991), "Does compulsory school attendance affect schooling and earnings?," *Quarterly Journal of Economics* 106, 979-1014.

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

[3] Bekker, P. (1994), "Alternative approximations to the distributions of instrumental variable estimators," *Econometrica* 62, 657-681.

[4] Bekker, P. and J. van der Ploeg (2005), "Instrumental variable estimation based on grouped data," *Statistica Neerlandica* 59, 239-267.

[5] Benabou, R. (1993), "Workings of a city: location, education, and production", *Quarterly Journal of Economics* 108, 619-52.

[6] Bifulco, R., Fletcher, J.M. and S.L. Ross (2011), "The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health," *American Economic Journal: Economic Policy* 3, 25-53.

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

[8] Brooks-Gunn, J., Duncan, G.J., Kato Klebanov, P. and N. Sealand (1993), "Do neighborhoods influence child and adolescent development," *American Journal of Sociology* 99, 353-395.

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

[10] Card, D. and J. Rothstein (2007), "Racial segregation and the black-white test score gap," *Journal of Public Economics* 91, 2158-2184.

[11] Clark, A.E. and Y. Loheac (2007), "It wasn't me, it was them! Social influence in risky behavior by adolescents," *Journal of Health Economics* 26, 763-784.

[12] Cohen-Cole, E. (2006), "Multiple groups identification in the linear-in-means model," *Economics Letters* 92, 157-162.

[13] Corcoran, M., Gordon, R., Laren, D. and G. Solon (1992), "The association between men's economic status and their family and community origins" *Journal of Human Resources* 27, 575-601.

[14] Cutler, D. M. and E. L. Glaeser (1997), "Are ghettos good or bad?" *Quarterly Journal of Economics* 112, 827-872.

[15] Datcher, L. (1982), "Effects of communuty and family background on achievement," *Review of Economics and Statistics* 64, 32-41.

[16] De Bartolome, C.A.M. (1990), "Equilibrium and inefficiency in a community model with peer group effects," *Journal of Political Economy* 98, 110-133.

[17] De Giorgi, G., Pellizzari, M. and S. Redaelli (2010), "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics* 2, 241-275.

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

[19] Dynarski, S., J. Hyman, and D. Whitmore Schanzenbach (2011), "Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion," NBER Working Paper No. 17533.

[20] Epple, D. and R.E. Romano (1998), "Competition between private and public schools: Vouchers and peer group effects," *American Economic Review* 88, 33-62.

[21] Ehrhardt, G., Marsili, M. and F. Vega-Redondo (2006), "Diffusion and growth in an evolving network," *International Journal of Game Theory* 34.

[22] 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.

[23] Gould, E.D., Lavy, V. and D. Paserman (2011), "Sixty years after the magic carpet ride: The long-run effect of the early childhood environment on social and economic outcomes," *Review of Ecoconomic Studies* 78, 938-973.

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

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

[26] Granovetter, M.S. (1973), "The strength of weak ties," *American Journal of Sociology* 78, 1360-1380.

[27] Granovetter, M.S. (1974), *Getting a Job: A Study of Contacts and Careers*, Cambridge, MA: Harvard University Press.

[28] Granovetter, M.S. (1983), "The strength of weak ties: A network theory revisited," *Sociological Theory* 1, 201-233.

[29] Hoxby, C. (2000), "Peer effects in the classroom, learning from gender and race variation," NBER Working Paper no. 7867.

[30] Ioannides, Y.M. (2011), "Neighborhood effects and housing," In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics, Vol. 1B*, Amsterdam: Elsevier Science, pp. 1281-1340.

[31] Ioannides, Y.M. and G. Topa (2010), "Neighborhood effects: Accomplishments and looking beyond them," *Journal of Regional Science* 50, 343-362.

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

[33] Jackson, M.O and A. Wolinsky (1996), "A strategic model of social and economic networks," *Journal of Economic Theory* 71, 44-74.

[34] Jackson, M.O. and Y. Zenou (2013), "Games on networks," In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier, forthcoming.

[35] Katz, L.F., Kling, J.R. and J.B. Liebman (2001), "Moving to opportunity in Boston: Early results of a randomized mobility experiment," *Quarterly Journal of Economics* 116, 607-654.

[36] König, M., Tessone, C. and Y. Zenou (2012), "Nestedness in networks: A theoretical model and some applications," Unpusblished manuscript, Stockholm University.

[37] Krueger, A.B., and D.M. Whitmore (2001), "The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project STAR," *Economic Journal* 111, 1-28.

[38] Lai, G., N. Lin and S.-Y. Leung (1998), "Network resources, contact resources, and status attainment," *Social Networks* 20, 159-178.

[39] Lang, K. (2007), *Poverty and Discrimination*, Princeton: Princeton University Press.

[40] Laschever, R. (2009), "The doughboys networks: Social interactions and labor market outcomes of World War I veterans," Unpusblished manuscript, University of Illinois at Urbana-Champaign.

[41] Lee, L-F. (2002), "Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models," *Econometric Theory* 18, 252-277.

[42] Lee, L-F. (2003), "Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances," *Econometric Reviews* 22, 307-335.

[43] Lee, L-F. (2007), "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," *Journal of Econometrics* 140, 333-374.

[44] Lee, L-F., Liu, X. and X. Lin (2010), "Specification and estimation of social interaction models with network structures," *Econometrics Journal* 13, 145-176.

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

[46] Lin, N., W.M. Ensel and J.C. Vaughn (1981), "Social resources and strength of ties: Structural factors in occupational status attainment," *American Sociological Review* 46, 393-405.

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

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

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

[50] Marsden, P.V. and J.S. Hurlbert (1988), "Social resources and mobility outcomes: A replication and extension," *Social Forces* 66, 1038-1059.

[51] Nechyba, T.J. (2000), "Mobility, targeting, and private-school vouchers," *American Economic Review* 90, 130-46.

[52] Patacchini, E. and Y. Zenou (2008), "The strength of weak ties in crime," *European Economic Review* 52, 209-236.

[53] Patacchini, E. and Y. Zenou (2011), "Intergenerational education transmission: Neighborhood quality and/or parents' involvement?", *Journal of Regional Science* 51, 987-1013.

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

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

[56] Sacerdote, B. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?", In: E.A. Hanushek, S. Machin and L. Woessmann (Eds.), *Handbook of Economics of Education, Vol. 3*, Amsterdam: Elevier Science, pp. 249-277.

[57] Topa, G. (2011), "Labor markets and referrals," In: J. Benhabib, A. Bisin and M.O. Jackson (Eds.), *Handbook of Social Economics*, Amsterdam: Elsevier Science, pp. 1193-1221.

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

[59] Weinberg, B.A. (2004), "Testing the spatial mismatch hypothesis using inter-city variations in industrial composition," *Regional Science and Urban Economics* 34, 505-532.

[60] Weinberg, B., P.B. Reagan, and J.J. Yankow (2004), "Do neighborhoods affect work behavior? Evidence from the NLSY 79," *Journal of Labor Economics* 22, 891-924.

[61] Yakubovich, V. (2005), "Weak ties, information, and influence: How workers find jobs in a local Russian labor market," *American Sociological Review* 70, 3, 408-421.

[62] Zax, J.S. and D.I. Rees (2002), "IQ, academic performance, environment, and earnings," *Review of Economics and Statistics* 84, 600-616.

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

# APPENDIX 1: Proofs for the theory

**Proof of Proposition 1.** Apply Theorem 1, part b, in Calvó-Armengol et al. (2009) to our problem. ∎

**Proof of Proposition 2.** Let us first consider the possible link creation. Because it is a star-shaped network, it should be clear that there is only one link creation to check:

$(i)$ Condition for which $B$ does not want to form a link with $C$:

To determine the benefit of forming the link $BC$, we need to calculate the new Nash equilibrium and the new Bonacich centralities of all individuals. First, if the link $BC$ is created, the adjacency matrix now becomes

$$\mathbf{G}_t + BC = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

The largest eigenvalue is 2 and the condition is $\phi < 0.5$. The equilibrium utilities are now given by

$$\mathbf{u}_t\left(g_t + BC\right) = \frac{1}{2\left(1 - \phi - 2\phi^2\right)^2} \begin{pmatrix} \left[(1 - \phi)\,\alpha_{A,t,t+1} + \phi\alpha_{B,t,t+1} + \phi\alpha_{C,t,t+1}\right]^2 \\ \left[\phi\alpha_{A,t,t+1} + (1 - \phi)\,\alpha_{B,t,t+1} + \phi\alpha_{C,t,t+1}\right]^2 \\ \left[\phi\alpha_{A,t,t+1} + \phi\alpha_{B,t,t+1} + (1 - \phi)\,\alpha_{C,t,t+1}\right]^2 \end{pmatrix} - \begin{pmatrix} 2c_A \\ 2c_B \\ 2c_C \end{pmatrix}$$

As a result, $B$ will not want to form a link with $C$ if: $u_{B,t}\left(g_t + BC\right) < u_{B,t}\left(g_t\right)$, which is equivalent to:

$$c_B > \frac{\left[\phi\alpha_{A,t,t+1} + (1 - \phi)\,\alpha_{B,t,t+1} + \phi\alpha_{C,t,t+1}\right]^2}{2\left(1 - \phi - 2\phi^2\right)^2} - \frac{\left[\phi\alpha_{A,t,t+1} + \left(1 - \phi^2\right)\alpha_{B,t,t+1} + \phi^2\alpha_{C,t,t+1}\right]^2}{2\left(1 - 2\phi^2\right)^2}$$

If we assume that $\alpha_{A,t,t+1} = \alpha_{B,t,t+1} = \alpha_{C,t,t+1} = 1$, then

$$c_B > \frac{\phi\left(2 - 4\phi^2 - \phi\right)\left(1 + \phi\right)^2}{2\left(1 - \phi - 2\phi^2\right)^2\left(1 - 2\phi^2\right)^2} \tag{18}$$

Because of mutual consent if $B$ does not want to form a link with $C$ then the link $BC$ will not be created even if $C$ wants to form this link. Thus, we don't need to check the condition for which $C$ does not want to form a link with $B$.

In terms of link deletions, because it is a star-shaped network, only two deviations need to be checked:

43

($ii$) Condition for which $A$ does not want to severe a link with $B$ (or $C$). If this link is severed, the adjacency matrix now becomes

$$\mathbf{G}_t - AB = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

The largest eigenvalue is 1 and the condition is $\phi < 1$. The equilibrium utilities are given by:

$$\mathbf{u}_t \left( g_t - AB \right) = \frac{1}{2} \begin{pmatrix} \left( \alpha_{A,t,t+1} + \phi \alpha_{C,t,t+1} \right)^2 / \left( 1 - \phi^2 \right)^2 \\ \alpha_{B,t,t+1}^2 \\ \left( \phi \alpha_{A,t,t+1} + \alpha_{C,t,t+1} \right)^2 / \left( 1 - \phi^2 \right)^2 \end{pmatrix} - \begin{pmatrix} c_A \\ 0 \\ c_C \end{pmatrix} \tag{19}$$

As a result, $A$ does not want to severe a link with $B$ or $C$ if: $u_{A,t} \left( g_t - AB \right) < u_{A,r,t} \left( g_t \right)$, which is equivalent

$$c_A < \frac{\left( \alpha_{A,t,t+1} + \phi \alpha_{B,t,t+1} + \phi \alpha_{C,t,t+1} \right)^2}{2 \left( 1 - 2\phi^2 \right)^2} - \frac{\left( \alpha_{A,t,t+1} + \phi \alpha_{C,t,t+1} \right)^2}{2 \left( 1 - \phi^2 \right)^2}$$

If we assume that $\alpha_{A,t,t+1} = \alpha_{B,t,t+1} = \alpha_{C,t,t+1} = 1$, then

$$c_A < \frac{\phi \left( 1 + \phi \right)^2 \left( 2 + \phi - 4\phi^2 \right)}{2 \left( 1 - 2\phi^2 \right)^2 \left( 1 - \phi^2 \right)^2} \tag{20}$$

($iii$) Condition for which $B$ (or $C$) does not want to severe a link with $A$. The adjacency matrix is as before and equal to $\mathbf{G}_t - AB$. The largest eigenvalue is 1 and the condition is $\phi < 1$. The equilibrium utilities are still given by (19). As a result, $B$ does not want to severe a link with $A$ if: $u_{B,t} \left( g_t - AB \right) < u_{B,t} \left( g_t \right)$, which is equivalent to:

$$c_B < \frac{\left[ \phi \alpha_{A,t,t+1} + \left( 1 - \phi^2 \right) \alpha_{B,t,t+1} + \phi^2 \alpha_{C,t,t+1} \right]^2}{2 \left( 1 - 2\phi^2 \right)^2} - \frac{\alpha_{B,t}^2}{2}$$

If we assume that $\alpha_{A,t,t+1} = \alpha_{B,t,t+1} = \alpha_{C,t,t+1} = 1$, then

$$c_B < \frac{\phi \left( 2 + 5\phi - 4\phi^3 \right)}{2 \left( 1 - 2\phi^2 \right)^2} \tag{21}$$

Observe that, for $\phi < 0.5$,

$$\frac{\phi \left( 2 + 5\phi - 4\phi^3 \right)}{2 \left( 1 - 2\phi^2 \right)^2} < \frac{\phi \left( 1 + \phi \right)^2 \left( 2 + \phi - 4\phi^2 \right)}{2 \left( 1 - 2\phi^2 \right)^2 \left( 1 - \phi^2 \right)^2}$$

Therefore, if $c_A = c_B = c_C = c$, then, if $\phi < 0.5$, the condition that guarantees that this network is pairwise stable is (which put together conditions (18), (20) and (21)):

$$\frac{\phi \left(2 - 4\phi^2 - \phi\right)\left(1 + \phi\right)^2}{2\left(1 - \phi - 2\phi^2\right)^2 \left(1 - 2\phi^2\right)^2} < c < \frac{\phi \left(2 + 5\phi - 4\phi^3\right)}{2\left(1 - 2\phi^2\right)^2}$$

which is condition (8) in the proposition. ∎

**Proof of Proposition 3**

In Wave II, the adjacency matrix of the network displayed on the right-hand side of Figure 5

$$\mathbf{G}_t = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

whose largest eigenvalues is: $\frac{1+\sqrt{5}}{2} = 1.618$. Let us calculate the Nash equilibrium utilities of all agents for this network in Wave II. For $\phi < 0.618$, we easily obtain the Bonacich centralities of all players:

$$\mathbf{b}_\alpha\left(g_t, \phi\right) = \frac{1}{\left(1 - 3\phi^2 + \phi^4\right)} \begin{pmatrix} \left(1 - \phi^2\right)\alpha_A + \phi\left(1 - \phi^2\right)\alpha_B + \phi\alpha_C + \phi^2\alpha_D \\ \phi\left(1 - \phi^2\right)\alpha_A + \left(1 - 2\phi^2\right)\alpha_B + \phi^2\alpha_C + \phi^3\alpha_D \\ \phi\alpha_A + \phi^2\alpha_B + \left(1 - \phi^2\right)\alpha_C + \phi\left(1 - \phi^2\right)\alpha_D \\ \phi^2\alpha_A + \phi^3\alpha_B + \phi\left(1 - \phi^2\right)\alpha_C + \left(1 - 2\phi^2\right)\alpha_D \end{pmatrix}$$

The equilibrium utilities are thus given by

$$\mathbf{u}\left(g_t\right) = \frac{1}{2\left(1 - 3\phi^2 + \phi^4\right)^2} \begin{pmatrix} \left[\left(1 - \phi^2\right)\alpha_A + \phi\left(1 - \phi^2\right)\alpha_B + \phi\alpha_C + \phi^2\alpha_D\right]^2 \\ \left[\phi\left(1 - \phi^2\right)\alpha_A + \left(1 - 2\phi^2\right)\alpha_B + \phi^2\alpha_C + \phi^3\alpha_D\right]^2 \\ \left[\phi\alpha_A + \phi^2\alpha_B + \left(1 - \phi^2\right)\alpha_C + \phi\left(1 - \phi^2\right)\alpha_D\right]^2 \\ \left[\phi^2\alpha_A + \phi^3\alpha_B + \phi\left(1 - \phi^2\right)\alpha_C + \left(1 - 2\phi^2\right)\alpha_D\right]^2 \end{pmatrix} - \begin{pmatrix} 2c_A \\ c_B \\ 2c_C \\ c_D \end{pmatrix}$$

$$(22)$$

($i$) Let us calculate under which condition student $A$ will *not* keep the link with $D$ in Wave II.

If player $A$ keeps the link with $D$, the adjacency matrix becomes:

$$\mathbf{G}_t + AD = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

and the equilibrium utilities:

$$\mathbf{u}\left(g_t + AD\right) = \frac{1}{2}\begin{pmatrix} \dfrac{[(1-\phi)\alpha_A+\phi(1-\phi)\alpha_B+\phi\alpha_C+\phi\alpha_D]^2}{\left(1-\phi-3\phi^2+\phi^3\right)^2} \\[4mm] \dfrac{\left[\phi(1-\phi)\alpha_A+\left(1-\phi-2\phi^2\right)\alpha_B+\phi^2\alpha_C+\phi^2\alpha_D\right]^2}{\left(1-\phi-3\phi^2+\phi^3\right)^2} \\[4mm] \dfrac{\left[\phi(1+\phi)\alpha_A+(1+\phi)\phi^2\alpha_B+\left(1-2\phi^2\right)\alpha_C+\phi\left(1+\phi-\phi^2\right)\alpha_D\right]^2}{\left(1-4\phi^2-2\phi^3+\phi^4\right)^2} \\[4mm] \dfrac{\left[\phi(1+\phi)\alpha_A+\phi^2(1+\phi)\alpha_B+\phi\left(1+\phi-\phi^2\right)\alpha_C+\left(1-2\phi^2\right)\alpha_D\right]^2}{\left(1-4\phi^2-2\phi^3+\phi^4\right)^2} \end{pmatrix} - \begin{pmatrix} 3c_A \\ c_B \\ 2c_C \\ 2c_D \end{pmatrix} \qquad (23)$$

As a result, $A$ will *not* keep the link with $D$ iff: $u_A\left(g_t + AD\right) < u_A\left(g_t\right)$, i.e.,

$$\frac{\left[\alpha_A\left(1-\phi\right)\alpha_A + \phi\left(1-\phi\right)\alpha_B + \phi\alpha_C + \phi\alpha_D\right]^2}{2\left(1-\phi-3\phi^2+\phi^3\right)^2} - 3c_A$$

$$< \frac{\left[\left(1-\phi^2\right)\alpha_A + \phi\left(1-\phi^2\right)\alpha_B + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1-3\phi^2+\phi^4\right)^2} - 2c_A$$

which is equivalent to:

$$c_A > \frac{\left[\left(1-\phi\right)\alpha_A + \phi\left(1-\phi\right)\alpha_B + \phi\alpha_C + \phi\alpha_D\right]^2}{2\left(1-\phi-3\phi^2+\phi^3\right)^2} - \frac{\left[\left(1-\phi^2\right)\alpha_A + \phi\left(1-\phi^2\right)\alpha_B + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1-3\phi^2+\phi^4\right)^2}$$
$$(24)$$

$(ii)$ Let us calculate under which condition student $A$ will create a link with $C$ in Wave II.

If student $A$ does not create a link with $C$ in Wave II, the adjacency matrix would be given by:

$$\mathbf{G}_t - AC = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

whose largest eigenvalue is 1. Thus, given that $\phi < 1$, the equilibrium utilities are:

$$\mathbf{u}\left(g_t - AC\right) = \frac{1}{2\left(1-\phi^2\right)^2}\begin{pmatrix} \left(\alpha_A + \phi\alpha_B\right)^2 \\ \left(\phi\alpha_A + \alpha_B\right)^2 \\ \left(\alpha_C + \phi\alpha_D\right)^2 \\ \left(\alpha_D + \phi\alpha_C\right)^2 \end{pmatrix} - \begin{pmatrix} c_A \\ c_B \\ c_C \\ c_D \end{pmatrix}$$

As a result, $A$ will create the link with $C$ in Wave II iff: $u_A\left(g_t\right) > u_A\left(g_t - AC\right)$, i.e.

$$\frac{\left[\left(1-\phi^2\right)\alpha_A + \phi\left(1-\phi^2\right)\alpha_B + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1-3\phi^2+\phi^4\right)^2} - 2c_A > \frac{\left(\alpha_A + \phi\alpha_B\right)^2}{2\left(1-\phi^2\right)^2} - c_A$$

46

which is equivalent to:

$$c_A < \frac{\left[\left(1 - \phi^2\right)\alpha_A + \phi\left(1 - \phi^2\right)\alpha_B + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1 - 3\phi^2 + \phi^4\right)^2} - \frac{(\alpha_A + \phi\alpha_B)^2}{2\left(1 - \phi^2\right)^2} \qquad (25)$$

($iii$) Let us finally calculate under which condition student $A$ will keep the link with $B$ in Wave II.

If student $A$ deletes the link with $D$, the adjacency matrix becomes:

$$\mathbf{G}_t - AB = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

whose largest eigenvalue is $\sqrt{2} = 1.414$. Thus, given that $\phi < 1.414$, the equilibrium utilities are given by:

$$\left[\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \phi \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\right]^{-1} \begin{pmatrix} \alpha_A \\ \alpha_B \\ \alpha_C \\ \alpha_D \end{pmatrix}$$

$$\mathbf{u}_t\left(g_t - AB\right) = \frac{1}{2\left(1 - 2\phi^2\right)^2}\begin{pmatrix} \left[\left(1 - \phi^2\right)\alpha_A + \phi\alpha_C + \phi^2\alpha_D\right]^2 \\ \left(1 - 2\phi^2\right)^2\alpha_B^2 \\ \left(\phi\alpha_A + \alpha_C + \phi\alpha_D\right)^2 \\ \left[\phi^2\alpha_A + \phi\alpha_C + \left(1 - \phi^2\right)\alpha_D\right]^2 \end{pmatrix} - \begin{pmatrix} c_A \\ 0 \\ 2c_C \\ c_D \end{pmatrix}$$

As a result, student $A$ will not delete a link with $B$ if $u_A\left(g_t - AB\right) < u_A\left(g_t\right)$, i.e.

$$\frac{\left[\left(1 - \phi^2\right)\alpha_A + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1 - 2\phi^2\right)^2} - c_A < \frac{\left[\left(1 - \phi^2\right)\alpha_A + \phi\left(1 - \phi^2\right)\alpha_B + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1 - 3\phi^2 + \phi^4\right)^2} - 2c_A$$

which is equivalent to:

$$c_A < \frac{\left[\left(1 - \phi^2\right)\alpha_A + \phi\left(1 - \phi^2\right)\alpha_B + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1 - 3\phi^2 + \phi^4\right)^2} - \frac{\left[\left(1 - \phi^2\right)\alpha_A + \phi\alpha_C + \phi^2\alpha_D\right]^2}{2\left(1 - 2\phi^2\right)^2} \qquad (26)$$

Therefore, conditions (24), (25) and (26) guarantee that, in Wave II, student $A$ will severe the link with $D$, will create the link with $C$ and will not delete the link with $B$ (Figure 5). These three conditions correspond to condition (15) in the proposition. ∎

# APPENDIX 2: GPA and Principal Component Analysis

The school performance is measured using the respondent's scores received in wave II in several subjects: English or language arts, history or social science, mathematics, and science. The scores are coded as 1=D or lower, 2=C, 3=B, 4=A. A student can have higher grades in a subset of subjects depending on his/her preferences and abilities. Given that we want to measure a general school performance, not a subject-specific one, we used the first principal component score as final composite index (labeled as GPA index or grade point average index). We apply a principal component analysis (PCA) to a matrix $S$ in which scores in each subject (columns) are registered for students (rows). PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of observations (student's scores) of possibly correlated variables (subjects) into a set of values of linearly uncorrelated variables called principal components. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it must be orthogonal to (i.e., uncorrelated with) the preceding components.

Let us explain more formally how these components are computed. So far, we have defined $S$ as a score matrix in which units are students (rows, $l = 1, .., n$) and variables are subjects (columns, $k = 1, ..., s$). The first elements to be defined are the covariance matrix $V = S'S$ and the metric matrix $M$, which is a $s \times s$ diagonal matrix whose generic element is the inverse of subject $s$' variance. Let us define the *Inertia* of a cloud of points (roughly defining the *variability in the data* ) as

$$I_g = \sum_{l=1}^{n} ||s_l||_M^2$$

where $s_l$ is the $l$th row of $S$, a $1 \times s$ row vector containing student $l$ scores. Essentially, we want to find a series of spaces that sequentially maximize the percentage of explained *Inertia* under the constraint of being orthogonal to previous ones, sequentially solving the following problem

$$\max_{W^k} I_{W^k} \quad s.t. \quad W^k \perp W_{k-1} \tag{28}$$

where $I_{W^k}$ is the explained inertia at the $k$th step by the space $W^k$ and $W_{k-1}$ is its ortogonal space, with

$$I_g \geq I_{W^1} \geq \cdot\cdot \geq \cdot\cdot \geq I_{W^k} \geq \cdot\cdot \geq \cdot\cdot \geq I_{W^s} \tag{29}$$

In our case, the first principal component explains 56 percent of total variance (see Figure A1). It is the only component which is positively and highly correlated to all subjects.
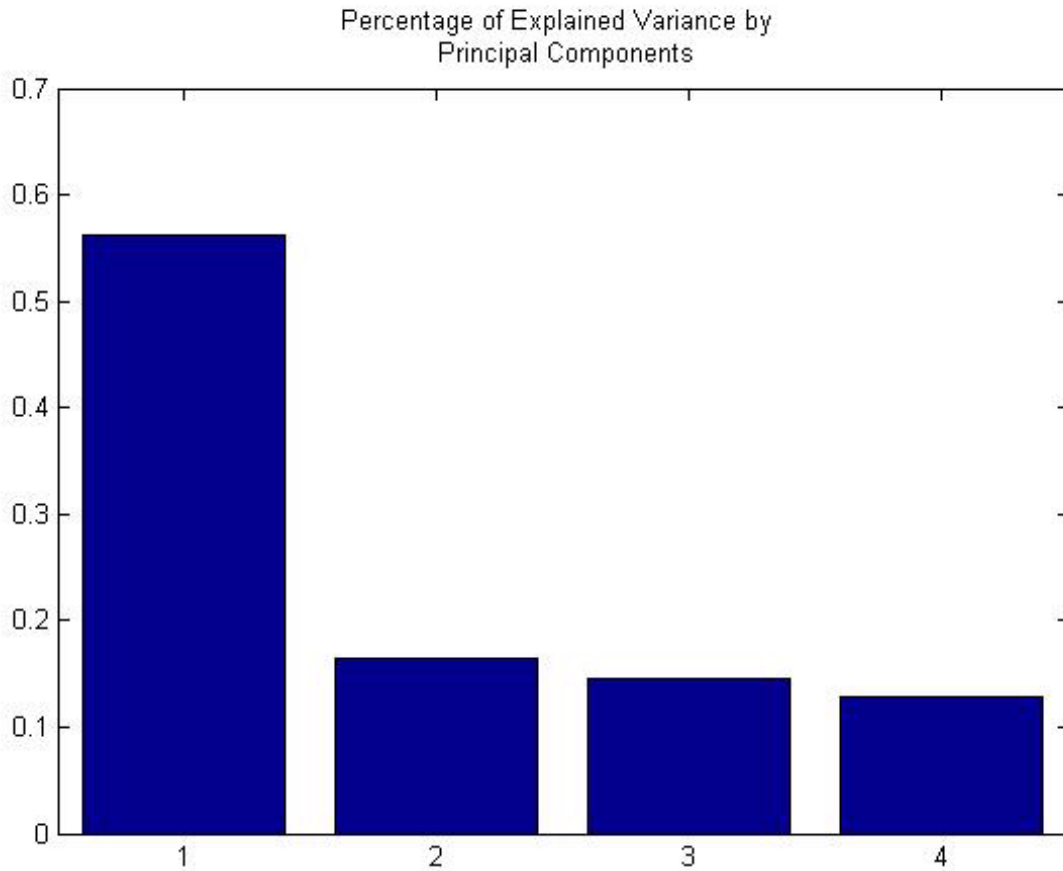


Figure A1

This first component captures a general high performance at school while other components capture the variability of scores among subjects (given the level of general performance). Correlations between components and subjects are represented in Figure A2. For example, in the North-West panel of Figure A2, subjects are points whose $x-$coordinate represents correlation between subject and first component while $y-$coordinate represents correlation between subject and second component. All subjects have roughly the same

49

correlation (about 0.5) with the first component while only mathematics has a strong and positive correlation with the second component, which discriminates between students who are relatively more successful in this subject relatively to all other subjects (expecially history) given their general performance's level.
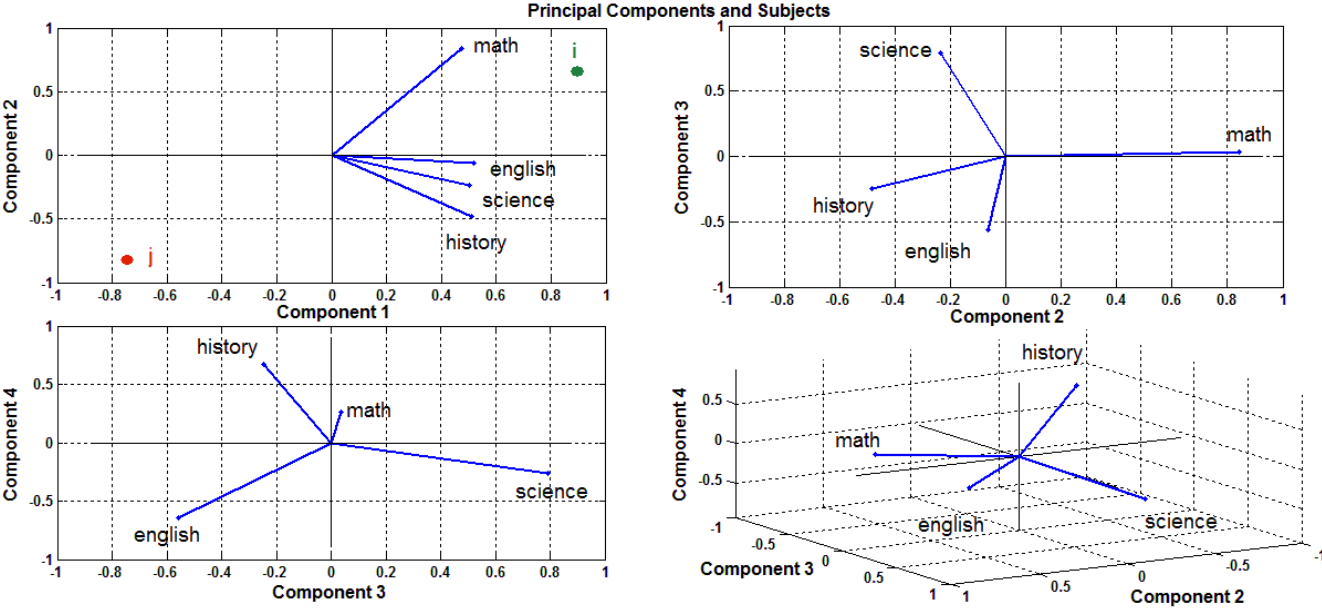


Figure A2: Principal Component Analysis

Let us illustrate this last point with an example (see the North-West panel of Figure A2). Student $i$ (green point) has a generalized above-the-average profile in every subject and has a high score for first principal component. This student is really excellent in mathematics and has thus a high score also in the second principal component (which distinguishes a propensity for mathematics). Student $i$ is therefore in the first quadrant of the coordinate plane formed by the first two principal components. Student $j$ (red point) is a below-average student and is particularly weak in mathematics but relatively good, for example, in history. As a result, this student will lie on the third quadrant of the coordinate plane. By considering the first component, we are able to isolate a global performance index at school and peer effects on this outcome might thus be considered as a spillover in general effort at school, excluding preferences or abilities in specific subjects.

Figure 2: Percentage of deleted links from W1 to W2 by order of nomination
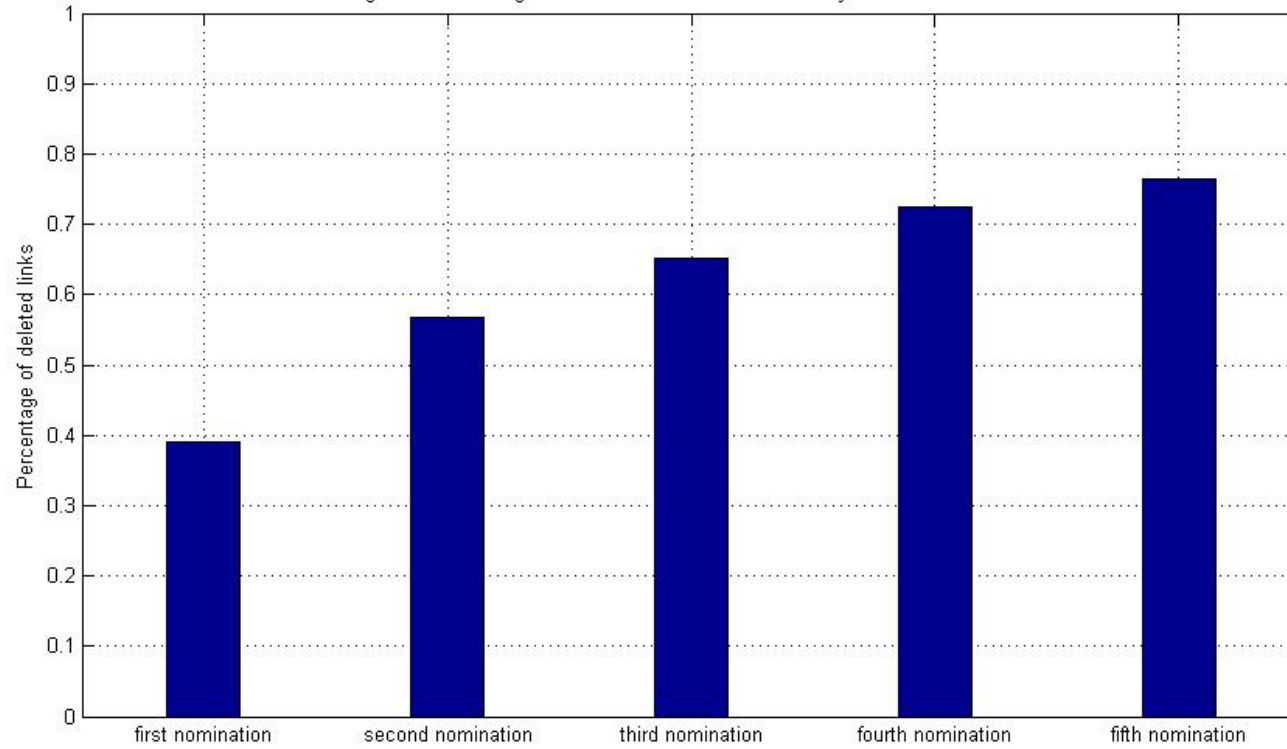
Figure 3: Correlations between own education and peers' education
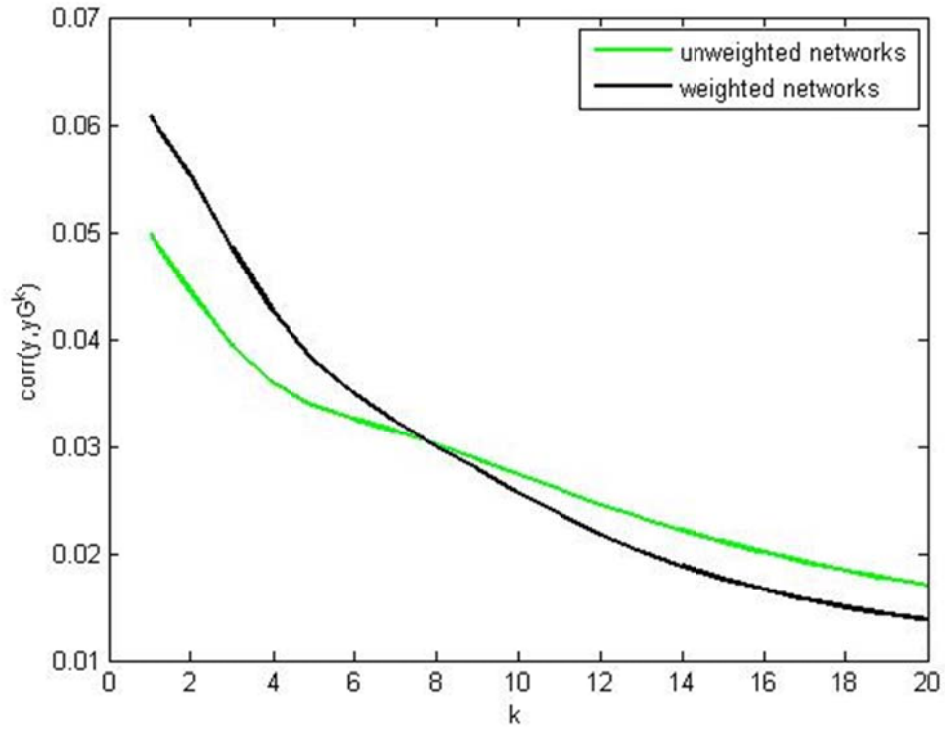


Notes: Network links are defined using the choices made (out-degree). The plotted correlations are statistically significant at the 1% level.

Figure 4: Strength of interactions or college decision proximity?
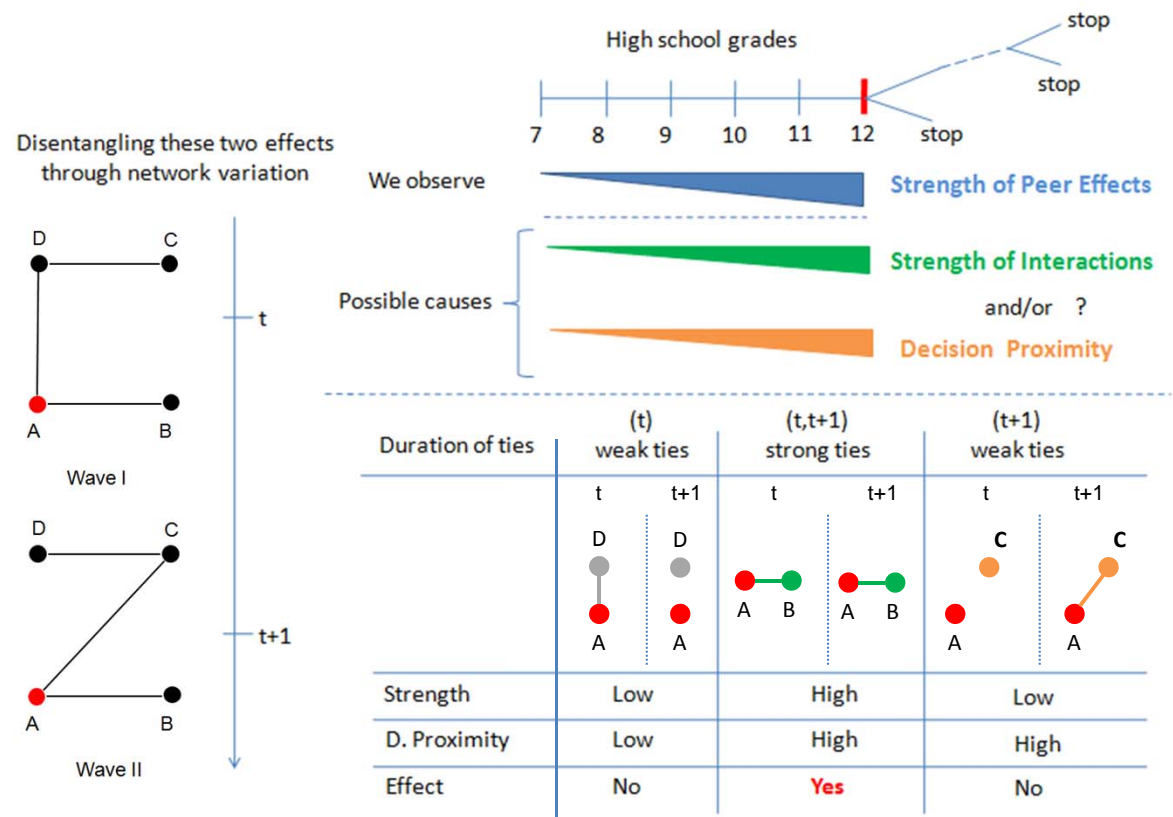
# Figure 7: Misspecification of network topology
## Numerical simulation

Estimates of peereffects



finite-IVs 2SLS     many-IVs 2SLS     bias-corrected 2SLS     lagged finite-IVs 2SLS     lagged many-IVs 2SLS     lagged bias-corrected 2SLS

T statistics



finite-IVs 2SLS     many-IVs 2SLS     bias-corrected 2SLS     lagged finite-IVs 2SLS     lagged many-IVs 2SLS     lagged bias-corrected 2SLS
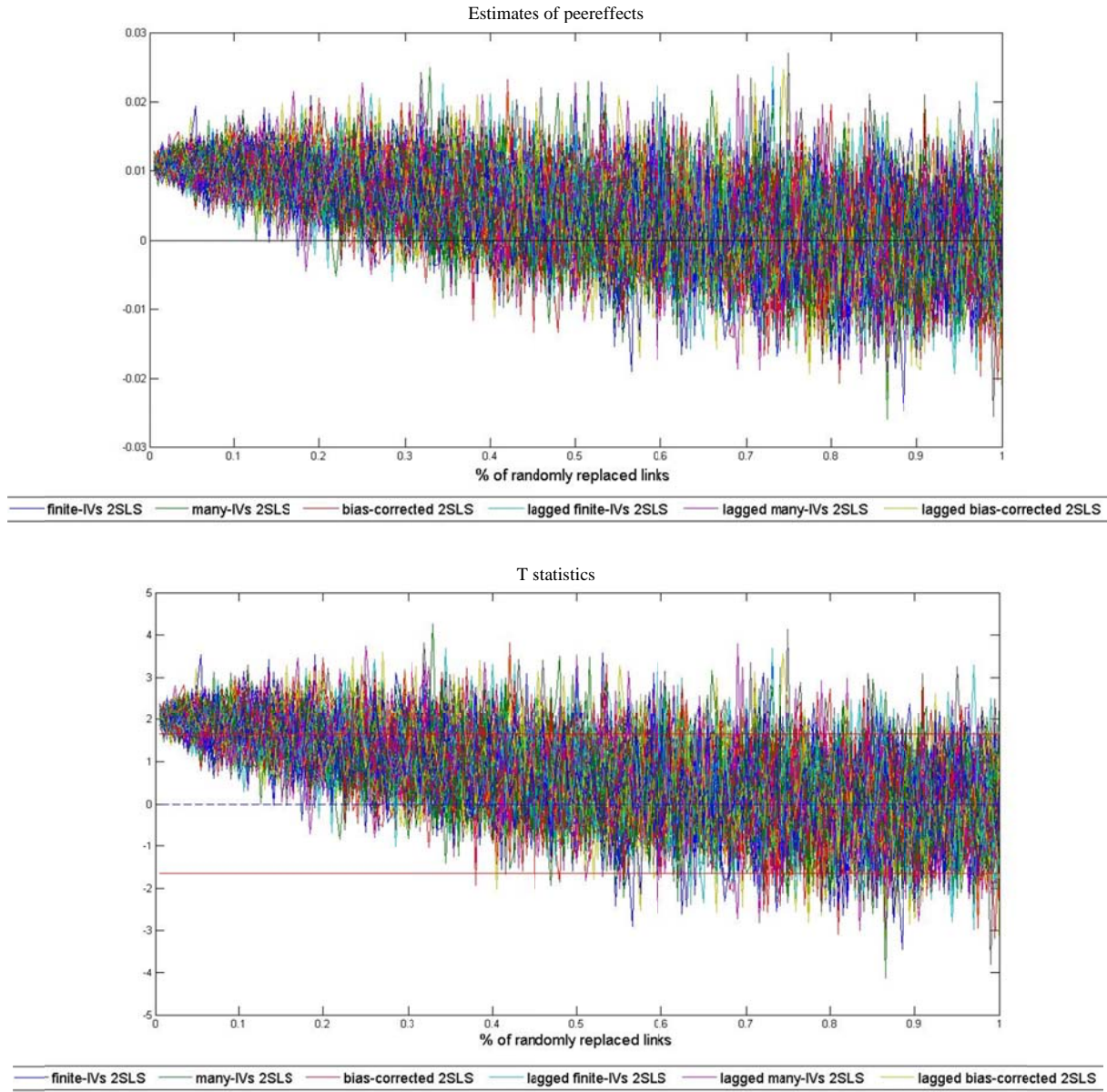
Notes: For each percentage of randomly replaced links, we draw 100 samples of size and network density equal to the real one and show the estimated peer effects and t-statistics (model specification (5))

Figure 8: Simulation experiment
Summarizing the evidence

Lagged bias-corrected 2SLS estimates of peer effects
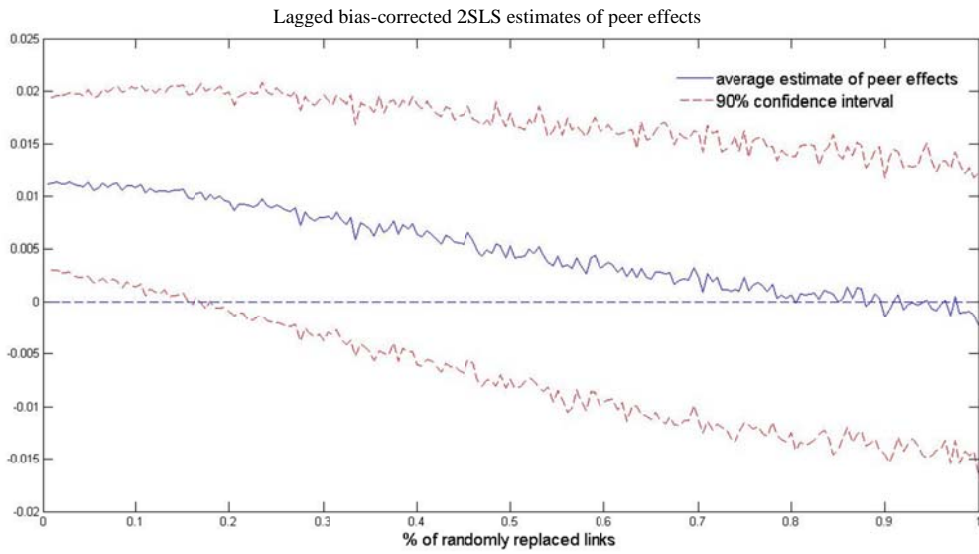


Notes: For each percentage of randomly replaced links, we average the estimates of peer effects across the drawn samples. The confidence bands are based on the derived standard errors, accounting for within

Table 1: Variables definition and summary statistics

| Variable | Description | Average (Std.Dev.) | Min - Max |
|---|---|---|---|
| **Wave II (grade 7-12 )** | | | |
| *Individual socio-demographic variables* | | | |
| Female | Dummy variable taking value one if the respondent is female. | 0.53 (0.50) | 0 - 1 |
| Black or African American | Race dummies. "White" is the reference group | 0.19 (0.39) | 0 - 1 |
| Other races | // | 0.10 (0.30) | 0 - 1 |
| Student grade | Grade of student in the current year. | 9.07 (1.65) | 7 - 12 |
| Religion practice | Response to the question: "In the past 12 months, how often did you attend religious services?", coded as 2= never, 3= less than once a month, 4= once a month or more, but less than once a week, 5= once a week or more. Coded as 1 if the previous is skipped because of response "none" to the question: "What is your religion?" | 3.79 (1.83) | 1 - 5 |
| Mathematics score A | Mathematics score dummies. Score in mathematics at the most recent grading period. D is the reference category, coded (A, B, C, D missing). | 0.29 (0.45) | 0 - 1 |
| Mathematics score B | // | 0.34 (0.48) | 0 - 1 |
| Mathematics score C | // | 0.21 (0.41) | 0 - 1 |
| Mathematics score Missing | // | 0.05 (0.21) | 0 - 1 |
| 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. | 4.00 (1.09) | 1 - 6 |
| GPA | In the text | 2.29 (1.49) | 0 - 6.09 |
| Physical development | Response to the question: "How advanced is your physical development compared to other boys/girls your age", coded as 1= I look younger than most, 2= I look younger than some, 3= I look about average, 4= I look older than some, 5= I look older than most. | 3.31 (1.11) | 1 - 5 |
| *Family background variables* | | | |
| Household size | Number of people living in the household | 3.40 (1.34) | 1 - 11 |
| 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. | 0.73 (0.44) | 0 - 1 |
| Parent education | Schooling level of the (biological or nonbiological) parent who is living with the child, distinguishing between "never went to school", "not graduate from high school","high school graduate","graduated from college or a university", "professional training beyond a four-year college", coded as 1 to 5. We consider only the education of the father if both parents are in the household. | 3.25 (0.97) | 1 - 5 |
| Parent occupation manager | Parent occupation dummies. Closest description of the job of (biological or nonbiological) parent that is living with the child is manager. If both parents are in the household, the occupation of the father is considered. "none" is the reference group | 0.11 (0.31) | 0 - 1 |
| Parent occupation professional/technical | // | 0.21 (0.41) | 0 - 1 |
| Parent occupation office or sales worker | // | 0.10 (0.33) | 0 - 1 |
| Parent occupation manual | // | 0.30 (0.46) | 0 - 1 |

| Parent occupation other | // | 0.14 (0.35) | 0 - 1 |
|---|---|---|---|
| *Protective factors* | | | |
| School attachment | Response to the question: "You feel like you are part of your school coded as 1= strongly agree, 2= agree, 3=neither agree nor disagree, 4= disagree, 5= strongly disagree. | 1.90 (0.90) | 1 - 5 |
| Relationship with teachers | Response to the question: "How often have you had trouble getting along with your teachers?" coded as 0= never, 1= just a few times, 2= about once a week, 3= almost everyday, 4=everyday | 0.91 (0.94) | 0 - 4 |
| Social inclusion | Response to the question: "How much do you feel that adults care about you, coded as 5= very much, 4= quite a bit, 3= somewhat, 2= very little, 1= not at all | 4.47 (0.73) | 1 - 5 |
| *Residential neighborhood variables* | | | |
| Residential building quality | Interviewer response to the question "How well kept is the building in which the respondent lives", coded as 4= very poorly kept (needs major repairs), 3= poorly kept (needs minor repairs), 2= fairly well kept (needs cosmetic work), 1= very well kept. | 1.52 (0.80) | 1 - 4 |
| **Wave IV (aged 25 - 31)** | | | |
| Years of education | Years of education attained by the individual. | 16.42 (3.21) | 9 - 26 |
| Years of education of peers | Sum of years of education attained by respondent peers. | 40.93 (29.48) | 9 - 207 |
| Children | Dummy variable taking value one if the respondent has a child. | 0.43 (0.50) | 0 - 1 |
| Married | Variable taking value one if the respondent is married | 0.42 (0.49) | 0 - 1 |
| Religion practice | Response to the question: 'How often have you attended religioius services in the past 12 months?", coded as 0= never, 1= a few times , 2= several times, 3= once a month, 4=2 or 3 times a month, 5=once a week, 6=more than once a week. | 1.75 (1.64) | 0 - 5 |

Table 2: Long-run peer effects in education

| Dep.Var. Years of Education | Unweighted Networks | | Weighted Networks | |
|---|---|---|---|---|
| | Total IV | Lagged IV | Total IV | Lagged IV |
| 2SLS finite IVs | 0.0076* | 0.0107* | 0.0127* | 0.0186** |
| | (0.0045) | (0.0062) | (0.0071) | (0.0095) |
| 2SLS many Ivs | 0.0067* | 0.008** | 0.0097* | 0.0122** |
| | (0.0039) | (0.0041) | (0.0058) | (0.0062) |
| Bias-corrected Ivs | 0.0077* | 0.0092** | 0.0113** | 0.0142** |
| | (0.0039) | (0.0041) | (0.0058) | (0.0062) |
| Individual socio-demographic variables | yes | yes | yes | yes |
| Family Background Variables | yes | yes | yes | yes |
| Protective Factors | yes | yes | yes | yes |
| Contextual Effects | yes | yes | yes | yes |
| Network Fixed Effects | yes | yes | yes | yes |
| Observations | 1749 | 1749 | 1749 | 1749 |

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: Long-run peer effects in education, grades 7-9

| Dep.Var. Years of Education | | | | |
|---|---|---|---|---|
| | Unweighted Networks | | Weighted Networks | |
| | Total IV | Lagged IV | Total IV | Lagged IV |
| 2SLS finite IVs | 0.01** | 0.0175** | 0.0121* | 0.0183** |
| | (0.0045) | (0.0065 ) | (0.007) | (0.0094) |
| 2SLS many Ivs | 0.0104** | 0.0123*** | 0.013** | 0.0152** |
| | (0.004) | (0.0043) | (0.006) | (0.0064) |
| Bias-corrected Ivs | 0.0109** | 0.013*** | 0.0137** | 0.016*** |
| | (0.004) | (0.0043) | (0.006) | (0.0064) |
| Individual socio-demographic variables | yes | yes | yes | yes |
| Family Background Variables | yes | yes | yes | yes |
| Protective Factors | yes | yes | yes | yes |
| Contextual Effects | yes | yes | yes | yes |
| Network Fixed Effects | yes | yes | yes | yes |
| Observations | 969 | 969 | 969 | 969 |

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Long-run peer effects in education, grades 10-12

| Dep.Var. Years of Education | | | | |
|---|---|---|---|---|
| | Unweighted Networks | | Weighted Networks | |
| | Total IV | Lagged IV | Total IV | Lagged IV |
| 2SLS finite IVs | 0.0179* | 0.0171 | 0.0296** | 0.031* |
| | (0.0096) | (0.0118) | (0.0144) | (0.0175) |
| 2SLS many Ivs | 0.0191* | 0.0196** | 0.031*** | 0.0326*** |
| | (0.0082) | (0.0088) | (0.0119) | (0.0128) |
| Bias-corrected Ivs | 0.0211*** | 0.0222*** | 0.0343*** | 0.0371*** |
| | (0.0082) | (0.0088) | (0.0119) | (0.0128) |
| Individual socio-demographic variables | yes | yes | yes | yes |
| Family Background Variables | yes | yes | yes | yes |
| Protective Factors | yes | yes | yes | yes |
| Contextual Effects | yes | yes | yes | yes |
| Network Fixed Effects | yes | yes | yes | yes |
| Observations | 615 | 615 | 615 | 615 |

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Long-run peer effects in education for weak ties, grades 10-12
(peers only in Wave I)

| Dep.Var. Years of Education | | | | |
|---|---|---|---|---|
| | Unweighted Networks | | Weighted Networks | |
| | Total IV | Lagged IV | Total IV | Lagged IV |
| 2SLS finite IVs | -0.0022 | -0.0056 | 0.0158 | 0.0169 |
| | (0.0141) | (0.0243) | (0.0221) | (0.0336) |
| 2SLS many Ivs | 0.0063 | 0.0002 | 0.0228 | 0.0163 |
| | (0.0109) | (0.0127) | (0.0162) | (0.0183) |
| Bias-corrected Ivs | 0.0078 | 0.0035 | 0.0254 | 0.0213 |
| | (0.0109) | (0.0126) | (0.0162) | (0.0183) |
| Individual socio-demographic variables | yes | yes | yes | yes |
| Family Background Variables | yes | yes | yes | yes |
| Protective Factors | yes | yes | yes | yes |
| Contextual Effects | yes | yes | yes | yes |
| Network Fixed Effects | yes | yes | yes | yes |
| Observations | 408 | 408 | 408 | 408 |

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Long-run peer effects in education for weak ties, grades 10-12
(peers only in Wave II)

| Dep.Var. Years of Education | | | | |
|---|---|---|---|---|
| | Unweighted Networks | | Weighted Networks | |
| | Total IV | Lagged IV | Total IV | Lagged IV |
| 2SLS finite IVs | -0.0012 | 0.0082 | 0.0048 | 0.0123 |
| | (0.0151) | (0.033) | (0.0222) | (0.0494) |
| 2SLS many Ivs | 0.0134 | 0.0154 | 0.0271 | 0.0411 |
| | (0.0121) | (0.0141) | (0.0175) | (0.0248) |
| Bias-corrected Ivs | 0.0136 | 0.0162 | 0.0273 | 0.0418 |
| | (0.0121) | (0.0141) | (0.0175) | (0.0248) |
| Individual socio-demographic variables | yes | yes | yes | yes |
| Family Background Variables | yes | yes | yes | yes |
| Protective Factors | yes | yes | yes | yes |
| Contextual Effects | yes | yes | yes | yes |
| Network Fixed Effects | yes | yes | yes | yes |
| Observations | 297 | 297 | 297 | 297 |

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7: Long-run peer effects in education for strong ties, grades 10-12
(peers in both Wave I and Wave II)

| Dep.Var. Years of Education | Unweighted Networks | | Weighted Networks | |
|---|---|---|---|---|
| | Total IV | Lagged IV | Total IV | Lagged IV |
| 2SLS finite IVs | 0.0193* | 0.0394** | 0.0285* | 0.0906** |
| | (0.01) | (0.0214) | (0.017) | (0.045) |
| 2SLS many Ivs | 0.0175* | 0.0385** | 0.0261 | 0.0746*** |
| | (0.0098) | (0.0158) | (0.0167) | (0.0304) |
| Bias-corrected Ivs | 0.0172* | 0.0377** | 0.0253 | 0.0711*** |
| | (0.0098) | (0.0157) | (0.0167) | (0.03) |
| Individual socio-demographic variables | yes | yes | yes | yes |
| Family Background Variables | yes | yes | yes | yes |
| Protective Factors | yes | yes | yes | yes |
| Contextual Effects | yes | yes | yes | yes |
| Network Fixed Effects | yes | yes | yes | yes |
| Observations | 300 | 300 | 300 | 300 |

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: Peer characteristics

| | Wave I | | | | Wave II | | | | Wave I and Wave II | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | std | min | max | mean | std | min | max | mean | std | min | max |
| Years of education | 16.96 | 3.48 | 11.00 | 25.00 | 16.62 | 3.35 | 11.00 | 25.00 | 16.91 | 3.39 | 11.00 | 25.00 |
| Female | 0.48 | 0.50 | 0.00 | 1.00 | 0.43 | 0.50 | 0.00 | 1.00 | 0.47 | 0.50 | 0.00 | 1.00 |
| Black or African American | 0.06 | 0.24 | 0.00 | 1.00 | 0.01 | 0.12 | 0.00 | 1.00 | 0.04 | 0.20 | 0.00 | 1.00 |
| Other races | 0.09 | 0.29 | 0.00 | 1.00 | 0.06 | 0.23 | 0.00 | 1.00 | 0.05 | 0.23 | 0.00 | 1.00 |
| Physical Development | 3.32 | 1.04 | 1.00 | 5.00 | 3.35 | 0.95 | 1.00 | 5.00 | 3.39 | 0.98 | 1.00 | 5.00 |
| Religion Practice | 3.98 | 1.82 | 1.00 | 7.00 | 3.88 | 1.89 | 1.00 | 7.00 | 3.84 | 1.86 | 1.00 | 7.00 |
| Household Size | 3.34 | 1.37 | 1.00 | 10.00 | 3.32 | 1.34 | 1.00 | 10.00 | 3.47 | 1.38 | 1.00 | 10.00 |
| Two married parents family | 0.75 | 0.44 | 0.00 | 1.00 | 0.77 | 0.42 | 0.00 | 1.00 | 0.80 | 0.40 | 0.00 | 1.00 |
| School Attachment | 2.10 | 0.98 | 1.00 | 5.00 | 2.07 | 1.00 | 1.00 | 5.00 | 1.94 | 0.95 | 1.00 | 5.00 |
| Relationship with teachers | 0.89 | 0.92 | 0.00 | 4.00 | 0.83 | 0.86 | 0.00 | 4.00 | 0.77 | 0.81 | 0.00 | 4.00 |
| Self esteem | 3.97 | 1.09 | 1.00 | 6.00 | 3.94 | 1.01 | 1.00 | 6.00 | 3.98 | 1.10 | 1.00 | 6.00 |
| Parent education | 3.33 | 0.90 | 1.00 | 5.00 | 3.22 | 0.83 | 1.00 | 5.00 | 3.19 | 0.83 | 1.00 | 5.00 |
| Social inclusion | 4.28 | 0.77 | 1.00 | 5.00 | 4.33 | 0.73 | 1.00 | 5.00 | 4.38 | 0.73 | 1.00 | 5.00 |
| Mathematics score A | 0.24 | 0.42 | 0.00 | 1.00 | 0.21 | 0.41 | 0.00 | 1.00 | 0.25 | 0.43 | 0.00 | 1.00 |
| Mathematics score B | 0.28 | 0.45 | 0.00 | 1.00 | 0.32 | 0.47 | 0.00 | 1.00 | 0.31 | 0.46 | 0.00 | 1.00 |
| Mathematics score C | 0.25 | 0.43 | 0.00 | 1.00 | 0.24 | 0.43 | 0.00 | 1.00 | 0.24 | 0.43 | 0.00 | 1.00 |
| Mathematics score missing | 0.08 | 0.27 | 0.00 | 1.00 | 0.08 | 0.28 | 0.00 | 1.00 | 0.07 | 0.26 | 0.00 | 1.00 |
| GPA | 2.33 | 1.53 | 0.00 | 6.09 | 2.27 | 1.48 | 0.00 | 6.09 | 2.25 | 1.55 | 0.00 | 6.09 |
| Residential uilding quality | 1.56 | 0.78 | 1.00 | 4.00 | 1.59 | 0.79 | 1.00 | 4.00 | 1.52 | 0.79 | 1.00 | 4.00 |
| Student grade | 10.44 | 0.50 | 10.00 | 11.00 | 10.43 | 0.50 | 10.00 | 11.00 | 10.40 | 0.49 | 10.00 | 11.00 |
| Children | 0.48 | 0.50 | 0.00 | 1.00 | 0.48 | 0.50 | 0.00 | 1.00 | 0.45 | 0.50 | 0.00 | 1.00 |
| Religion Practice (Wave 4) | 1.44 | 1.62 | 0.00 | 5.00 | 1.32 | 1.56 | 0.00 | 5.00 | 1.46 | 1.66 | 0.00 | 5.00 |
| Married | 0.50 | 0.50 | 0.00 | 1.00 | 0.50 | 0.50 | 0.00 | 1.00 | 0.49 | 0.50 | 0.00 | 1.00 |

Notes: Differences between means are never statistical significant at conventional levels of significance

Table 9: Link formation
Model (16) OLS estimation results

| VARIABLE | $\gamma$ Coefficient | Std. error |
|---|---|---|
| Female | 0.0025 | (0.019) |
| Black or African American | -0.0107 | (0.044) |
| Other races | -0.0265 | (0.036) |
| Student grade | -0.0226 | (0.016) |
| Religion Practice | 0.0032 | (0.008) |
| Self esteem | 0.0090 | (0.009) |
| Mathematics score A | -0.0063 | (0.017) |
| Mathematics score B | 0.0178 | (0.018) |
| Mathematics score C | 0.0122 | (0.019) |
| Mathematics score missing | -0.0030 | (0.022) |
| School Attachment | -0.0108 | (0.010) |
| Physical Development | -0.0060 | (0.008) |
| Social inclusion | -0.0102 | (0.013) |
| Parent education | -0.0073 | (0.013) |
| Household Size | 0.0125 | (0.008) |
| Parent occupation professional/technical | 0.0103 | (0.024) |
| Parent occupation manual | 0.0028 | (0.023) |
| Parent occupation office or sales worker | 0.0162 | (0.027) |
| Parent occupation other | 0.0460** | (0.022) |
| Two married parents family | -0.0026 | (0.020) |
| Relationship with teachers | -0.0106 | (0.010) |
| Residential building quality | 0.0085 | (0.012) |
| Network fixed effects | yes | |
| Chow test p value | 0.6083 | |
| Observations | 6,932 | |

Notes: Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Short-run peer effects in education for weak ties, grades 10-12
(peers only in Wave I)

| Dep.Var. GPA | | |
| --- | --- | --- |
| | Unweighted Networks | Weighted Networks |
| | Total IV | Total IV |
| 2SLS finite IVs | 0.1525* | 0.1256* |
| | (0.0906) | (0.0726) |
| 2SLS many Ivs | 0.0123 | 0.0498 |
| | (0.0295) | (0.0401) |
| Bias-corrected Ivs | 0.0121* | 0.0199* |
| | (0.0072) | (0.0121) |
| Individual socio-demographic variables | yes | yes |
| Family Background Variables | yes | yes |
| Protective Factors | yes | yes |
| Contextual Effects | yes | yes |
| Network Fixed Effects | yes | yes |
| Observations | 408 | 408 |

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 11: Short-run peer effects in education for weak ties, grades 10-12
(peers only in Wave II)

| Dep.Var. GPA | | |
| --- | --- | --- |
| | Unweighted Networks | Weighted Networks |
| | Total IV | Total IV |
| 2SLS finite IVs | 0.1189* | 0.1725** |
| | (0.0321) | (0.0845) |
| 2SLS many Ivs | 0.0144 | 0.0152 |
| | (0.0351) | (0.0351) |
| Bias-corrected Ivs | 0.0215 | 0.0281* |
| | (0.0231) | (0.0161) |
| Individual socio-demographic variables | yes | yes |
| Family Background Variables | yes | yes |
| Protective Factors | yes | yes |
| Contextual Effects | yes | yes |
| Network Fixed Effects | yes | yes |
| Observations | 297 | 297 |

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 12: Short-run peer effects in education for strong ties, grades 10-12
(peers in both Wave I and Wave II)

| Dep.Var. GPA | Unweighted Networks | Weighted Networks |
|---|---|---|
| | Total IV | Total IV |
| 2SLS finite IVs | 0.0429 | 0.1324* |
| | (0.0563) | (0.0790) |
| 2SLS many Ivs | 0.0195 | 0.0253 |
| | (0.0171) | (0.0273) |
| Bias-corrected Ivs | 0.0173* | 0.0243* |
| | (0.0271) | (0.0144) |
| Individual socio-demographic variables | yes | yes |
| Family Background Variables | yes | yes |
| Protective Factors | yes | yes |
| Contextual Effects | yes | yes |
| Network Fixed Effects | yes | yes |
| Observations | 300 | 300 |

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1