Strong versus Weak Ties in Education*

Eleonora Patacchini† Edoardo Rainone‡ Yves Zenou§

Abstract

We develop a network model embedding the role of weak and strong ties in education decisions and consider the specification and estimation of social interaction models with different network structures. The empirical salience of the model is tested using a very detailed longitudinal dataset of adolescent friendship networks. 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 more than one waves) and not when they are weak ties (friends in only one wave). We also find that this result does not hold in the short run where both weak and strong ties have an impact on current grades.

Key words: Social networks, education, long-term peer effects, spatial autoregressive model.

JEL Classification: C31, D85, I21, Z13.

*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 and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

†Syracuse University, EIEF, IZA and CEPR. E-mail: epatacch@maxwell.syr.edu

‡Banca d'Italia and La Sapienza University of Rome E-mail: edoardo.rainone@uniroma1.it. The views expressed here do not necessarily reflect those of Banca d'Italia.

§Stockholm University, Research Institute of Industrial Economics (IFN) and GAINS. E-mail: yves.zenou@ne.su.se.

1 Introduction

The influence of peers on education outcomes has been widely studied both in economics and sociology (Sacerdote, 2011). Yet many questions remain unanswered, in particular concerning the role of neighborhoods and peers on education where no consensus has emerged. 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 level (Datcher, 1982; Corcoran et al., 1992) where individuals reside. The importance of peer effects as distinct from neighborhood influences is still a matter of debate in many fields (see, e.g., the literature surveys by Durlauf, 2004, Ioannides and Topa, 2010, and Ioannides, 2011, 2012). Besides, the mechanisms by which the peer effects affect education is unclear. Moreover, very little is known about the effect of school peers on the long-run outcomes of students. This is primarily due to the absence of information on peers together with long-run outcomes of individuals in most existing data.

In this paper, we focus on the long-run effects of high-school peers on years of schooling and put forward the role of strong and weak ties in educational outcomes.

To be more precise, we extend the network model proposed by CalvóArmengol et al. (2009) to incorporate strong and weak tie friendship relationships and to study how they affect the decision of college education. We then test this model using the unique information on friendship networks among students in the United States provided by the Addhealth data. We exploit three 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 variation in friendship network topology between Wave I and Wave II, which enables to distinguish between weak and strong ties, *(iii)* 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.

More specifically, we use the different 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 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.[1]

The empirical counterpart of our theoretical model has the specification of a spatial autoregressive (SAR) model with different spatial weight matrices. The specification and estimation of a traditional spatial autoregressive model with one weight matrix in the social space have been studied by Liu and Lee (2010). In this framework, the different positions of group members as measured by the Bonacich (1987) centrality provide additional information for identification and estimation. In this paper, we extend Liu and Lee (2010) 2SLS approach to a network model with two interaction matrices using appropriate IVs. First, we use two centralities, one for strong ties and one for weak ties. Second, we take advantage of the longitudinal structure of our data and include in the different instrumental matrices only values lagged in time (i.e. observed in wave I). The asymptotic consistency and efficiency of the proposed estimators are proved.

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. When looking at the role of *weak* and *strong ties* in educational decisions, it appears that the education decisions of weak ties have no significant effect on individual long-run outcomes, regardless on 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. While there is a large literature on the role of weak ties in the labor market (Granovetter, 1973, 1974, 1983), our study is the first to look at the impact of weak or strong ties on educational outcomes. Interestingly, 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.

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 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. Section 6 collects the empirical evidence and investigates the economic mechanisms

[1] Observe that our definition of weak and strong ties is slightly different from that of Granovetter (1973, 1974, 1983). Here strong ties are viewed as *stable* relationships and weak ties as *unstable* relationships. In his seminal papers, Granovetter 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.

behind our peer-effects results. Section 7 shows the robustness of our results with respect to network topology mispecification while Section 8 compares the short-run versus long-run effects of peers on education. Finally, Section 9 concludes and discusses some policy implications. All proofs of propositions and corollaries in the text can be found in Appendix A.

2 Related literature

Our paper lies at the intersection of different literatures.

Theory of peer effects in education There are some early theoretical contributions of peer effect models in education (De Bartolome, 1990; Benabou, 1993) but extremely few papers are using a network approach.[2] An exception is Calvó-Armengol et al. (2009), which is the first model with an explicit network that looks at the impact of peers on education outcomes. We extend here this model by incorporating weak and strong ties, i.e. by allowing peer effects to be heterogeneous.

Econometrics of networks The literature on identification and estimation of social network models has significantly progressed recently (see Blume et al., 2011, for a recent survey). In his seminal work, Manski (1993) introduces a linear-in-means social interaction model with endogenous effects, contextual effects, and correlated effects. Manski shows that the linear-in-means specification suffers from the "reflection problem" and the different social interaction effects cannot be separately identified. Bramoullé et al. (2009) generalize Manski's linear-in-means model to a general local-average social network model, whereas the endogenous effect is represented by the average outcome of the peers. They provide some general conditions for the identification of the local-average model and suggest using an indirect connection's characteristics as an instrument for the endogenous effect. Liu and Lee (2010) consider a local-aggregate network model where the endogenous effect is given by the aggregate outcome of the peers. They show that in the local-aggregate model, the Bonacich centrality (Bonacich, 1987) can be used as an additional instrument to improve estimation efficiency. Here, we go further by first extending the Liu and Lee (2010) network model specification to the case of two interaction matrices and then by designing appropriate estimators. We show that the proposed estimators are consistent and also asymptotically

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

efficient when the sample size grows fast enough relative to the number of instruments. Such a model and estimation strategy is suitable to study heterogenous peer effects and can be used in different contexts.

Empirical studies of peer effects in education 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). In this paper, we decompose the information on network structure in different types and apply this framework to look at the long-term effects of peers on education and to distinguish between weak and strong ties.

Long-run peer effects in education 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 early 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 (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) were 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, i.e. strong versus weak ties.

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. 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.[3] 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 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.[4] To the best of our knowledge, there is no study on the impact of weak or strong ties on educational outcomes. Furthermore, 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.

[3]See Zenou (2013) for a recent theoretical model of weak and strong ties in the labor market but where the social network is not explicitly modeled.

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

3 Theoretical framework

Consider a population of n individuals.

The network $N = \{1, \ldots, n\}$ is a finite set of agents. We keep track of social connections by a network g whose adjacency matrix is $\mathbf{G} = \{g_{ij}\}$, where $g_{ij} = 1$ if i and j are direct friends, and $g_{ij} = 0$, otherwise. Friendship are reciprocal so that $g_{ij} = g_{ji}$. We also set $g_{ii} = 0$ so that individuals are not linked to themselves. The adjacency matrix is thus a $0-1$ symmetric matrix.

The *neighbors* of an individual i in a network g are denoted by $N_i(g)$. The *degree* of an agent i in a network g is the number of neighbors (here friends) that i has in the network, so that $d_i(g) = |N_i(g)|$.

There are two types of relationship in the network, weak and strong tie relationships. The difference between weak and strong ties will be clear below. Denote the *strong-tie* adjacency matrix by $\mathbf{G}^S$ and the *weak-tie* adjacency matrix by $\mathbf{G}^W$, with $\mathbf{G}^S + \mathbf{G}^W = \mathbf{G}$, where the superscripts S and W denote a strong and a weak-tie relationship, respectively. To construct the *strong-tie* adjacency matrix $\mathbf{G}^S$, we put a 1 in $\mathbf{G}$ only if there is a strong-tie relationship between i and j, i.e. $g_{ij}^S = 1$ and a 0 otherwise. Similarly, to construct the *weak-tie* adjacency matrix $\mathbf{G}^W$, we put a 1 in $\mathbf{G}$ only if there is a weak-tie relationship between i and j, i.e. $g_{ij}^W = 1$ and a 0 otherwise. This means that the intersection between i's strong ties and i's weak ties is empty $\forall i$.

To illustrate these matrices, consider the following network:

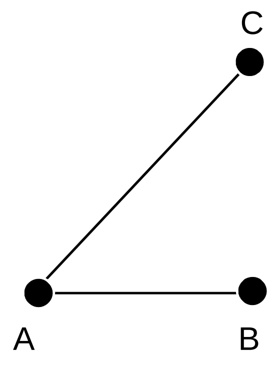


Figure 1: A star network

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

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

Assume that individual A has a *weak-tie* relationship with B but a *strong-tie* relationship with C. We easily obtain:

$$\mathbf{G}^S = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ and } \mathbf{G}^W = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Finally, we denote by $N_i^S(g)$ and $N_i^W(g)$ the set of strong and weak-tie friends each individual i has in network g and the cardinality of these sets by $d_i^S(g)$ and $d_i^W(g)$, respectively. In the network of Figure 1, we have, for example: $N_A^S(g) = \{C\}$, $N_A^W(g) = \{B\}$, $d_A^S(g) = 1$ and $d_1^W(g) = 1$.

Preferences Individuals decide how much effort to exert in education (e.g. how many years to study). We denote by y_i the educational effort level of individual i (i.e. years of schooling) and by $\mathbf{y} = (y_1, ..., y_n)'$ the population effort profile. We characterize the *weak* and *strong-tie* relationships between two individuals by the *strength* of their relationship, denoted by ϕ. This means that $\phi^S > \phi^W$. Even though this is not directly modeled, following the definition of Granovetter, a strong-tie relationship means that individuals often interact with each other (i.e. they are friends for a long period of time) while a weak-tie relationship indicates that individuals interact less with each other (i.e. they are friends for a shorter period of time). As stated above, friendship relationships are reciprocal so that if i has a strong (weak) tie relationship with j, then j has also a strong (weak) tie relationship with i. Each agent i selects an effort $y_i \geq 0$, and obtains a payoff $u_i(\mathbf{y}, g)$ that depends on the effort profile $\mathbf{y}$ and on the underlying network g, in the following way:

$$u_i(\mathbf{y}, g) = (a_i + \eta + \varepsilon_i)\; y_i - \frac{1}{2} y_i^2 + \left(\phi^S \sum_{j \in N_i^S(\mathbf{g})}^{n} y_j + \phi^W \sum_{j \in N_i^W(\mathbf{g})}^{n} y_j \right) y_i \qquad (1)$$

where $\phi^S, \phi^W > 0$, with $\phi^S > \phi^W$. The structure of this utility function is now relatively standard in games on networks (Ballester et al., 2006; Calvó-Armengol et al., 2009; Jackson and Zenou, 2013) where there is an idiosyncratic exogenous part $(a_i + \eta + \varepsilon_i)\ y_i - \frac{1}{2}y_i^2$ and an endogenous peer effect aspect $\phi^S \sum_{j \in N_i^S(\mathbf{g})}^n y_i y_j + \phi^W \sum_{j \in N_i^W(\mathbf{g})}^n y_i y_j$. The main difference with the standard approach is that we have here *heterogenous* peer effects since strong and weak ties have different impacts on own utility. Indeed, we have:

$$\frac{\partial u_i(\mathbf{y}, g)}{\partial y_i \partial y_j} = g_{ij}^S \phi^S + g_{ij}^W \phi^W \geq 0$$

where $g_{ij}^S = 1$ ($g_{ij}^W = 1$) if there exists a strong-tie (weak-tie) relationship between i and j and zero otherwise. To the best of our knowledge, this is the first paper that introduces different ϕ in a network model with strategic complementarities.

Observe that η denotes the unobservable network characteristics, ε_i is an error term (observable by all individuals but not by the researcher) and there is also an ex ante *idiosyncratic heterogeneity*, a_i, 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 friends, i.e. average level of parental education of i's friends, etc. (contextual effects). To be more precise, a_i can be written as:

$$a_i = \sum_{m=1}^{M} \beta_m x_i^m + \frac{1}{g_i^S} \sum_{m=1}^{M} \sum_{j=1}^{n_r^S} g_{ij}^S\, x_j^m \gamma_m^S + \frac{1}{g_i^W} \sum_{m=1}^{M} \sum_{j=1}^{n_r^W} g_{ij}^W\, x_j^m \gamma_m^W \tag{2}$$

where x_i^m is a set of M variables accounting for observable differences in individual characteristics of individual i, $\beta_m, \gamma_m^S, \gamma_m^W$ are parameters and $g_i^S = \sum_{j=1}^n g_{ij}^S$ and $g_i^W = \sum_{j=1}^n g_{ij}^W$ is the total number of strong-tie and weak-tie friends individual i has.

To summarize, when individual i exerts some effort in education, the benefits of the activity depends on own effects (i.e. on individual characteristics a_i, some network characteristics η and on some random element ε_i, which is specific to individual i and non observable by the researcher) and on peer effects, where the strength of interactions differ between strong and weak ties. Note that the utility (1) is concave in own decisions, and displays decreasing marginal returns in own effort levels. In sum,

$$u_i(\mathbf{y}, g) = \underbrace{(a_i + \eta + \varepsilon_i)\, y_i}_{\text{Benefits from own effort}} \underbrace{-\frac{1}{2}y_i^2}_{\text{Costs}} + \underbrace{\phi^S \sum_{j=1}^{n} g_{ij}^S\, y_i y_j}_{\text{Benefits from strong ties' effort}} + \underbrace{\phi^W \sum_{j=1}^{n} g_{ij}^W\, y_i y_j}_{\text{Benefits from weak ties' effort}}$$

Nash equilibrium We now characterize the Nash equilibrium of the game where agents choose their effort level $y_i \geq 0$ simultaneously. At equilibrium, each agent maximizes her utility (1) and we obtain the following best-reply function for each $i = 1, ..., n$:

$$y_i = \phi^S \sum_{j=1}^{n} g_{ij}^S\, y_j + \phi^W \sum_{j=1}^{n} g_{ij}^W\, y_j + a_i + \eta + \varepsilon_i \tag{3}$$

where a_i is given by (2). Denote $\alpha_i = a_i + \eta + \varepsilon_i$ and the corresponding $(1 \times n)$ vector by $\boldsymbol{\alpha}$. The matrix form equivalent of (3) is:

$$\mathbf{y} = \left(\mathbf{I} - \phi^S \mathbf{G}^S - \phi^W \mathbf{G}^W\right)^{-1} \boldsymbol{\alpha} \tag{4}$$

Denote by $\mu_1(\mathbf{G})$ the spectral radius of $\mathbf{G}$. We have:

Proposition 1 *If $\mu_1\left(\phi^S \mathbf{G}^S + \phi^W \mathbf{G}^W\right) < 1$, the peer effect game with payoffs (1) has a unique interior Nash equilibrium in pure strategies given by (3) or by (4).*

This proposition totally characterizes the Nash equilibrium and gives a condition that guarantees the existence, uniqueness and interiority of this equilibrium.

Corollary 1 *Assume that $\mathbf{G}$ is symmetric. A sufficient condition for the Nash equilibrium (3) or by (4) to exist, to be unique and to be interior is: $\left(\phi^S + \phi^W\right) \mu_1(\mathbf{G}) < 1$.*

This is an interesting result because it connects the adjacency matrix $\mathbf{G}$ to the split structure of peer effects ϕ^S and ϕ^W and it is directly comparable to the condition given in Ballester et al. (2006), i.e. $\phi\mu_1(\mathbf{G}) < 1$, where peer effects were assumed to be the same across all agents. The condition given in Proposition 1, $\mu_1\left(\phi^S \mathbf{G}^S + \phi^W \mathbf{G}^W\right) < 1$, is less restrictive than $\left(\phi^S + \phi^W\right) \mu_1(\mathbf{G}) < 1$ since $\mu_1\left(\phi^S \mathbf{G}^S + \phi^W \mathbf{G}^W\right) \leq \phi^S \mu_1\left(\mathbf{G}^S\right) + \phi^W \mu_1\left(\mathbf{G}^W\right)$ (see the proof of Corollary 1). However, it imposes a restriction on the exact topology of both weak and strong tie networks while $\left(\phi^S + \phi^W\right) \mu_1(\mathbf{G}) < 1$ allows $\mathbf{G}^S$ and $\mathbf{G}^W$ to be more free in terms of topology constraints but, more importantly, it does not require the matrix $\mathbf{G}$ (and thus $\mathbf{G}^S$ and $\mathbf{G}^W$) to be symmetric.

To illustrate this result, consider the network described in Figure 1. The largest eigenvalue of $\mathbf{G}$ is $\sqrt{2}$ while the largest eigenvalue of $\mathbf{G}^S$ and of $\mathbf{G}^W$ is 1. Thus, the sufficient condition given in Corollary 1 is: $\phi^S + \phi^W < 0.707$. In that case, there exists a unique Nash equilibrium given by:

$$\begin{aligned}
\mathbf{y} &= \left(\mathbf{I} - \phi^S \mathbf{G}^S - \phi^W \mathbf{G}^W\right)^{-1} \boldsymbol{\alpha} \\
&= \frac{1}{\left[1 - \left(\phi^S\right)^2 - \left(\phi^W\right)^2\right]} \begin{pmatrix} \alpha_A + \alpha_B \phi^W + \alpha_C \phi^S \\ \alpha_A \phi^W + \alpha_B \left[1 - \left(\phi^S\right)^2\right] + \alpha_C \phi^W \phi^S \\ \alpha_A \phi^S + \alpha_B \phi^W \phi^S + \alpha_C \left[1 - \left(\phi^W\right)^2\right] \end{pmatrix}
\end{aligned}$$

Interestingly, even though individuals B and C have each one link, it is easily verified that C's educational effort is much higher than B's and much closer to A, who has two links. This shows this importance of weak and strong ties in the relationship between different individuals.

Take, for example, $\alpha_1 = \alpha_2 = \alpha_3 = 1$. If for example, $\phi^S = 0.8$ and $\phi^W = 0.2$ (which implies that $\left(\phi^S\right)^2 + \left(\phi^W\right)^2 = 0.68 < 1$), then the condition given in Proposition 1 is satisfied since $\mu_1\left(\phi^S \mathbf{G}^S + \phi^W \mathbf{G}^W\right) = 0.825 < 1$ but the one given in Corollary 1 is not since $\phi^S + \phi^W = 1 > 0.707$. Since $\mu_1\left(\phi^S \mathbf{G}^S + \phi^W \mathbf{G}^W\right) = 0.825 < 1$, we can still derive the equilibrium and obtain:

$$y_A^* = 6.25,\ y_B^* = 4.125 \text{ and } y_C^* = 6$$

This shows again the importance of the position in the network and the role of weak and strong ties in educational outcomes. Strong ties influence each other so that their joint efforts increase independently of the network position. Even though this is not directly modeled here, one can interpret a link in the network between two students i and j, i.e. $g_{ij} = 1$ as an information exchange about education or even the establishment of a social norm in education. For example, if my friends think it is cool to study and to go to college, then I'm more likely to do so. This is even truer if the friends have a strong-tie relationship with each other. If they have only a weak-tie relationship, then the influence on each other is less important and we cannot speak about social norms. This is the main idea that we want to test in this paper.

4 Data description

Our analysis is made possible by the use of a unique database on friendship networks from the National Longitudinal Survey of Adolescent Health (AddHealth). The AddHealth 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).[5]

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).[6] This information is collected in Wave I and one year after, in 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).

Moreover, and this has not been done before, we would like to distinguish between strong and weak ties in the data. 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 believe that this is a good approximation of our measure in the model since, if i and j have nominated each other twice in two different waves, this clearly indicates that the relationship between them is strong. On the contrary, if they do not nominate each other twice, this indicates that their relationship is weak.

By matching the identification numbers of the friendship nominations to respondents'

[5] The AddHealth website describes survey design and data in details. http://www.cpc.unc.edu/projects/addhealth

[6] The limit in the number of nominations is not binding (even by gender). Less than 1% of the students in our sample show a list of ten best friends, both in Wave I and Wave II.

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 IV contains detailed information on the highest education qualification achieved. We measure education attainment in completed years of full time education.[7] 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 1,819 individuals distributed over 116 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.[8] In wave I, the mean and the standard deviation of network size are roughly 9.5 and 15, respectively. Roughly 61% of the nominations are not renewed in Wave II, and about 44% new ones are made. On average, these adolescents have roughly 30% strong ties and 70% weak ties. Further details on nomination data can be found in Appendix B. Appendix B gives also a precise definition of the variables used in our study[9] as well as their descriptive statistics.

5 Empirical model and estimation strategy

5.1 Empirical model

Let $\bar{r}$ be the total number of networks in the sample, n_r be the number of individuals in the rth network g_r, and $n = \sum_{r=1}^{\bar{r}} n_r$ be the total number of sample observations. Let $\boldsymbol{x}_{i,r} = (x_{i,r}^m, \cdots, x_{i,r}^M)'$. Using (2), the econometric model corresponding to the best-reply

[7] 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.

[8] 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 (see Calvó-Armengol et al., 2009).

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

function (3) of agent i in network g_r can be written as:

$$\begin{aligned} y_{i,r,t+1} &= \phi^S \sum_{j=1}^{n_r} g_{ij,r,t}^S y_{j,r,t+1} + \phi^W \sum_{j=1}^{n_r} g_{ij,r,t}^W y_{j,r,t+1} + \boldsymbol{x}'_{i,r}\delta \\ &+ \frac{1}{g_{i,r,t}^S} \sum_{j=1}^{n_r} g_{ij,r,t}^S x'_{j,r,t,t+1}\gamma^S + \frac{1}{g_{i,r,t}^W} \sum_{j=1}^{n_r} g_{ij,r,t}^W x'_{j,r,t,t+1}\gamma^W + \eta_{r,t} + \epsilon_{i,r,t+1}, \end{aligned} \tag{5}$$

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 and/or Wave II in 1995-96 (depending whether we consider weak or strong ties). 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 $g_{i,r,t}^S = \sum_{j=1}^n g_{ij,r,t}^S$ and $g_{i,r,t}^W = \sum_{j=1}^n g_{ij,r,t}^W$ are the total number of strong-tie and weak-tie friends each individual i has in network r at time t. Finally, $\epsilon_{i,r}$'s are i.i.d. innovations with zero mean and variance σ^2 for all i and r.

Let $\mathbf{Y}_r = (y_{1,r,t+1}, \cdots, y_{n_r,r,t+1})'$, $\mathbf{X}_r = (x_{1,r,t,t+1}, \cdots, x_{n_r,r,t,t+1})'$, and $\boldsymbol{\epsilon}_r = (\epsilon_{1,r}, \cdots, \epsilon_{n_r,r})'$. Denote the $n_r \times n_r$ adjacency matrix 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 $\mathbf{l}_{n_r}$. Let us split the adjacency matrix into two submatrices $\mathbf{G}_r^S$ and $\mathbf{G}_r^W$, which keep trace of weak and strong ties, respectively. Then model (5) can be written in matrix form as:

$$\mathbf{Y}_r = \phi^S \mathbf{G}_r^S \mathbf{Y}_r + \phi^W \mathbf{G}_r^W \mathbf{Y}_r + \mathbf{X}_r^* \beta + \eta_r \mathbf{l}_{n_r} + \boldsymbol{\epsilon}_r, \tag{6}$$

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

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

$$\mathbf{Y} = \phi^S \mathbf{G}^S \mathbf{Y} + \phi^W \mathbf{G}^W \mathbf{Y} + \mathbf{X}^* \beta + \boldsymbol{\iota} \cdot \boldsymbol{\eta} + \boldsymbol{\epsilon} \tag{7}$$

In this model, ϕ^S and ϕ^W represent *the endogenous effects*, i.e. the agent's outcome depends on that of his/her friends, while γ^S and γ^W represent *the contextual effect*, i.e. the agent's choice/outcome depends on the exogenous characteristics of his/her friends.

The vector of network fixed effects $\boldsymbol{\eta}$ captures *the correlated effect* where agents in the same network may behave similarly as they have similar unobserved individual characteristics or they face a similar (e.g. institutional) environment.

5.2 Estimation strategy

A number of papers using network data have dealt with the identification and estimation of peer effects with correlated effects (e.g., Clark and Loheac 2007; Lee 2007; Bramoullé et al., 2009; Liu and Lee, 2010, Calvó-Armengol et al., 2009; Lin, 2010; Lee et al., 2010) and we follow this approach. The common strategy is to exploit the architecture of network contacts to construct valid IVs for the endogenous effect (i.e. the characteristics of indirect friends) and to use network fixed effect as a remedy for the selection bias that originates from the possible sorting of individuals with similar unobserved characteristics into a network. The underlying assumption is that such unobserved characteristics are common to the individuals within each network. This is reasonable in our case study where the networks are quite small (see Section 4).

It should be noted that the traditional problem of reverse causality in the identification of peer effects arises when the peers and the decisions are taken at the same time. In our analysis, we do not have this problem since there is a time lag between when friends are chosen (Wave I and II in 1994-1996) and when the outcome (education) is observed (Wave IV in 2007-2008).

In addition, the longitudinal aspect of our analysis provides IVs at different points in time, i.e. the characteristics of indirect peers when they are at school and when they are adults. It is thus possible to use as instruments only variables lagged in time to ensure that the instruments are not correlated with the contemporaneous error term.

In the next section, we explain our identification and estimation strategy in more details.

5.2.1 2SLS estimation

Our econometric methodology extends Liu and Lee (2010)'s 2SLS estimation strategy to a social interaction model with two different network structures. Let us expose this approach and highlight the modification that is implemented in this paper. Model (7) can be written as:

$$\mathbf{Y} = \mathbf{Z}\theta + \boldsymbol{\iota} \cdot \boldsymbol{\eta} + \boldsymbol{\epsilon}, \tag{8}$$

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

We treat $\boldsymbol{\eta}$ 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:

$$\mathbf{JY} = \mathbf{JZ}\theta + \mathbf{J}\boldsymbol{\epsilon}. \tag{9}$$

Let $\mathbf{M} = (\mathbf{I} - \phi^S\mathbf{G}^S - \phi^W\mathbf{G}^W)^{-1}$. The equilibrium outcome vector $\mathbf{Y}$ in (8) is then given by the reduced form equation:

$$\mathbf{Y} = \mathbf{M}(\mathbf{X}^*\beta + \boldsymbol{\iota}\cdot\boldsymbol{\eta}) + \mathbf{M}\boldsymbol{\epsilon}. \tag{10}$$

It follows that $\mathbf{G}^S\mathbf{Y} = \mathbf{G}^S\mathbf{M}\mathbf{X}^*\beta + \mathbf{G}^S\mathbf{M}\boldsymbol{\iota}\boldsymbol{\eta} + \mathbf{G}^S\mathbf{M}\boldsymbol{\epsilon}$ and $\mathbf{G}^W\mathbf{Y} = \mathbf{G}^W\mathbf{M}\mathbf{X}^*\beta + \mathbf{G}^W\mathbf{M}\boldsymbol{\iota}\boldsymbol{\eta} + \mathbf{G}^W\mathbf{M}\boldsymbol{\epsilon}$. $\mathbf{G}^S\mathbf{Y}$ and $\mathbf{G}^W\mathbf{Y}$ are correlated with $\boldsymbol{\epsilon}$ because $\mathrm{E}[(\mathbf{G}^S\mathbf{M}\boldsymbol{\epsilon})'\boldsymbol{\epsilon}] = \sigma^2\mathrm{tr}(\mathbf{G}^S\mathbf{M}) \neq 0$ and $\mathrm{E}[(\mathbf{G}^W\mathbf{M}\boldsymbol{\epsilon})'\boldsymbol{\epsilon}] = \sigma^2\mathrm{tr}(\mathbf{G}^W\mathbf{M}) \neq 0$. Hence, in general, (9) cannot be consistently estimated by OLS.[10] 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. Liu and Lee (2010) use an instrumental variable approach and propose different estimators based on different instrumental matrices, denoted here by $\mathbf{Q}_1$ and $\mathbf{Q}_2$. They first consider the 2SLS estimator based on the conventional instrumental matrix for the estimation of (9): $\mathbf{Q}_1 = \mathbf{J}(\mathbf{GX}^*, \mathbf{X}^*)$ (*finite-IVs 2SLS*). They 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 $\mathbf{G}$ and are indicators of centrality in the networks. Liu and Lee (2010) 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). 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 procedure based on the estimated leading-order many-IV bias (*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

[10] 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.

case study.[11]

Let us now derive the best 2SLS estimator for equation (9). From the reduced form equation (8), we have $E(\mathbf{Z}) = [\mathbf{G}^S\mathbf{M}(\mathbf{X}^*\beta + \boldsymbol{\iota}\cdot\boldsymbol{\eta}), \mathbf{G}^W\mathbf{M}(\mathbf{X}^*\beta + \boldsymbol{\iota}\cdot\boldsymbol{\eta}), \mathbf{X}^*]$. The best IV matrix for $\mathbf{JZ}$ is given by

$$\mathbf{J}f = \mathbf{JE}(\mathbf{Z}) = J[\mathbf{G}^S\mathbf{M}(\mathbf{X}^*\beta + \boldsymbol{\iota}\cdot\boldsymbol{\eta}), \mathbf{G}^W\mathbf{M}(\mathbf{X}^*\beta + \boldsymbol{\iota}\cdot\boldsymbol{\eta}), \mathbf{X}^*] \tag{11}$$

which is an $n\times(3m+2)$ matrix. However, this matrix is infeasible as it involves unknown parameters. Note that f can be considered as a linear combination of the vectors in $\mathbf{Q_0} = J[\mathbf{G}^S\mathbf{M}(\mathbf{X}^*+\boldsymbol{\iota}), \mathbf{G}^W\mathbf{M}(\mathbf{X}^*+\boldsymbol{\iota}), \mathbf{X}^*]$. As $\boldsymbol{\iota}$ has $\bar{r}$ columns the number of IVs in $\mathbf{Q_0}$ increases as the number of groups increases. Furthermore as $\mathbf{M} = (\mathbf{I}-\phi^S\mathbf{G}^S-\phi^W\mathbf{G}^W)^{-1} = \sum_{j=0}^{\infty}(\phi^S\mathbf{G}^S+\phi^W\mathbf{G}^W)^j$ when $\sup||\phi^S\mathbf{G}^S+\phi^W\mathbf{G}^W||_\infty < 1$, $\mathbf{MX}^*$ and $\mathbf{M}\boldsymbol{\iota}$, can be approximated by linear combinations of

$$(\mathbf{G}^S\mathbf{X}^*, \mathbf{G}^W\mathbf{X}^*, \mathbf{G}^W\mathbf{G}^S\mathbf{X}^*, \left(\mathbf{G}^S\right)^2\mathbf{X}^*, \left(\mathbf{G}^W\right)^2\mathbf{X}^*, \left(\mathbf{G}^W\right)^2\mathbf{G}^S\mathbf{X}^*, \left(\mathbf{G}^W\right)^2\left(\mathbf{G}^S\right)^2\mathbf{X}^*, \cdots)$$

and

$$(\mathbf{G}^S\boldsymbol{\iota}, \mathbf{G}^W\boldsymbol{\iota}, \mathbf{G}^W\mathbf{G}^S\boldsymbol{\iota}, \left(\mathbf{G}^S\right)^2\boldsymbol{\iota}, \left(\mathbf{G}^W\right)^2\boldsymbol{\iota}, \left(\mathbf{G}^W\right)^2\mathbf{G}^S\boldsymbol{\iota}, \left(\mathbf{G}^W\right)^2\left(\mathbf{G}^S\right)^2\boldsymbol{\iota}, \cdots),$$

respectively. Hence, $\mathbf{Q}_0$ can be approximated by a linear combination of

$$\begin{aligned}\mathbf{Q}_\infty \;=\; & \mathbf{J}(\mathbf{G}^S(\mathbf{G}^S\mathbf{X}^*, \mathbf{G}^W\mathbf{X}^*, \mathbf{G}^W\mathbf{G}^S\mathbf{X}^*, \cdots, \mathbf{G}^S\boldsymbol{\iota}, \mathbf{G}^W\boldsymbol{\iota}, \mathbf{G}^W\mathbf{G}^S\boldsymbol{\iota}, \cdots), \\ & \mathbf{G}^W(\mathbf{G}^S\mathbf{X}^*, \mathbf{G}^W\mathbf{X}^*, \mathbf{G}^W\mathbf{G}^S\mathbf{X}^*, \cdots, \mathbf{G}^S\boldsymbol{\iota}, \mathbf{G}^W\boldsymbol{\iota}, \mathbf{G}^W\mathbf{G}^S\boldsymbol{\iota}, \cdots), \mathbf{X}^*)\end{aligned}$$

Let $\mathbf{Q_K}$ be an $n\times K$ submatrix of $\mathbf{Q}_\infty$ (with $K\geq 3m+2$) including $\mathbf{X}^*$. Let $\mathbf{Q_S}$ be a $n\times K_S$ submatrix of $\mathbf{Q_{S\infty}} = \mathbf{G}^S(\mathbf{G}^S\mathbf{X}^*, \mathbf{G}^W\mathbf{X}^*, \mathbf{G}^W\mathbf{G}^S\mathbf{X}^*, \cdots, \mathbf{G}^S\boldsymbol{\iota}, \mathbf{G}^W\boldsymbol{\iota}, \mathbf{G}^W\mathbf{G}^S\boldsymbol{\iota}, \cdots)$ and $\mathbf{Q_W}$ a $n\times K_W$ submatrix of $\mathbf{Q_{W\infty}} = \mathbf{G}^W(\mathbf{G}^S\mathbf{X}^*, \mathbf{G}^W\mathbf{X}^*, \mathbf{G}^W\mathbf{G}^S\mathbf{X}^*, \cdots, \mathbf{G}^S\boldsymbol{\iota}, \mathbf{G}^W\boldsymbol{\iota}, \mathbf{G}^W\mathbf{G}^S\boldsymbol{\iota}, \cdots)$. We assume $\frac{K_S}{K_W} = 1$. Let $\mathbf{P_K} = \mathbf{Q_K}(\mathbf{Q_K}'\mathbf{Q_K})^{-1}\mathbf{Q_K}'$ be the projector of $\mathbf{Q_K}$. The resulting 2SLS estimator is given by:

$$\widehat{\theta}_{2sls} = \left(\mathbf{Z}'\mathbf{P_K}\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{P_K}\mathbf{y} \tag{12}$$

5.2.2 Asymptotic properties of 2SLS estimator

As shown by Liu and Lee (2010), the 2SLS with a fixed number of IVs would be consistent but not efficient. Asymptotic efficiency can be achieved using a sequence of IVs in which the

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

number of IVs grows slow enough relative to the sample size. In general $\mathbf{K}$ may be seen as an increasing function of n. Following Liu and Lee (2010), we assume the following regularity conditions:

Assumption C1: The elements of ϵ are i.i.d with zero mean, variance σ^2 and a moment of order higher than four exists.

Assumption C2: The elements of $\mathbf{X}^*$ are uniformly bounded constants, $\mathbf{X}^*$ has the full rank k and $\lim_{n\to\infty} \mathbf{X}^{*\prime}\mathbf{X}^*$ exists and is nonsingular.

Assumption C3: The sequences of matrices $\left\{\mathbf{G}^S\right\}$, $\left\{\mathbf{G}^W\right\}$, $\{\mathbf{M}\}$ are uniformly bounded.

Assumption C4: $\overline{\mathbf{H}} = \lim_{n\to\infty} \frac{1}{n} f' f$ is a finite non singular matrix.

Assumption C5: There exists a $\mathbf{K}\times(3m+2)$ matrix π_K such that $\|f - \mathbf{Q_K}\pi_K\|_\infty \to 0$ as $n, K \to \infty$.

The 2SLS estimator with an increasing number of IVs approximating f can be asymptotically efficient under some conditions. However, when the number of instruments increases too fast, such an estimator could be asymptotically biased, which is known as the many-instrument problem. Let $\mathbf{\Psi}_{K,S} = \mathbf{P_K}\mathbf{G}^S\mathbf{M}$ and $\mathbf{\Psi}_{K,W} = \mathbf{P_K}\mathbf{G}^W\mathbf{M}$. The following proposition shows consistency and asymptotic normality of the 2SLS estimator (12).

Proposition 2 *Under assumptions C1-C5, if* $K/n \to 0$, *then* $\sqrt{n}\left(\widehat{\theta} - \theta - b_{2sls}\right) \xrightarrow{d} N\left(0, \sigma^2\overline{\mathbf{H}}^{-1}\right)$, *where* $b_{2sls} = \sigma^2\left(\mathbf{Z}'\mathbf{P_K}\mathbf{Z}\right)^{-1}\left[\mathrm{tr}\left(\mathbf{\Psi}_{K,S}\right), \mathrm{tr}\left(\mathbf{\Psi}_{K,W}\right), \mathbf{0}_{3m\times 1}\right]' = O_p\left(K/n\right)$.

Due to the increasing number of IVs $\sqrt{n}\left(\widehat{\theta} - \theta\right)$ has the bias $\sqrt{n}b_{2SLS}$, when $K^2/n \to 0$ the bias term converges to zero and the sequence of IV matrices $\mathbf{Q_K}$ gives the best IV estimator as $\sigma^2\overline{\mathbf{H}}^{-1}$ reaches the efficiency lower bound for the IV estimators.

Corollary 2 *Under assumptions C1-C5, if* $K^2/n \to 0$, *then* $\sqrt{n}\left(\widehat{\theta} - \theta\right) N\left(0, \sigma^2\overline{\mathbf{H}}^{-1}\right)$.

In summary, having a sequence of IV matrices $\{\mathbf{Q_K}\}$, the condition $K/n \to 0$ is fundamental for the estimator to be consistent, while having $K^2/n \to 0$ provides the asymptotically best estimator because $\sigma^2\overline{\mathbf{H}}^{-1}$ brings the lower bound for the IV estimators.

In this paper, we use 2SLS estimators and propose two innovations. First, we use two centralities, one for strong ties and one for weak ties in $\mathbf{Q}_2$ (*many-IVs 2SLS*). Second, we take advantage of the longitudinal structure of our data and include in the different instrumental matrices only values lagged in time (i.e. observed in wave I). Let $\boldsymbol{Q}_{1L}$ and $\boldsymbol{Q}_{2L}$ denote the set of instruments $\mathbf{Q}_1$ and $\mathbf{Q}_2$ which include only variables at Wave I (i.e. lagged in time).

Note that $[\mathbf{G}^S\boldsymbol{\iota}, \mathbf{G}^W\boldsymbol{\iota}]$ has $2\bar{r}$ columns, so if we include Bonacich centralities for both weak and strong ties from each of the $\bar{r}$ groups in $\mathbf{Q_K}$ then $2\bar{r}/K \to 0$. Hence, $K/n \to 0$

implies $2\bar{r}/n = 2/\bar{s} \to 0$ where $\bar{s}$ is the average group size. Then, as shown by Liu and Lee (2010) for the case of a single endogenous variable (i.e. coming from one interaction matrix), the average group size needs to be large enough, it should also be large relative to the number of groups because for the asymptotic efficiency it must be $K^2/n \to 0$ and it implies $(2\bar{r})^2/n = 2\bar{r}/\bar{s} \to 0$. If the network is not characterized by these properties a bias correction should be used. Given the topology of the Add Health network, which is composed by a quite large number of relatively small networks, the best (feasible) estimator is the *bias-corrected* one

$$\widehat{\theta}_{c2sls} = (\mathbf{Z}'\mathbf{P}_{\mathbf{Q_K}}\mathbf{Z})^{-1}\left[\mathbf{Z}'\mathbf{P}_{\mathbf{Q_K}}\mathbf{y} - \widetilde{\sigma}^2\left[\text{tr}\left(\widetilde{\mathbf{\Psi}}_{K,S}\right), \text{tr}\left(\widetilde{\mathbf{\Psi}}_{K,W}\right), \mathbf{0}_{3m\times 1}\right]'\right] \tag{13}$$

where $\widetilde{\mathbf{\Psi}}_{K,S} = \mathbf{P_K}\mathbf{G}^S\mathbf{M}(\widetilde{\phi}^S, \widetilde{\phi}^W)$ and $\widetilde{\mathbf{\Psi}}_{K,S} = \mathbf{P_K}\mathbf{G}^W\mathbf{M}(\widetilde{\phi}^S, \widetilde{\phi}^W)$ are estimated with initial $\sqrt{n}$-consistent estimators of $\widetilde{\sigma}$, $\widetilde{\phi}^S$ and $\widetilde{\phi}^W$. This estimator adjusts the 2SLS estimator by the estimated leading order bias b_{2sls}, which is presented in Proposition 2.

Proposition 3 *Under assumptions C1-C5, if $K/n \to 0$ and $\widetilde{\sigma}, \widetilde{\phi}^S$ and $\widetilde{\phi}^W$ are $\sqrt{n}-$consistent initial estimators of σ, ϕ^S and ϕ^W, then $\sqrt{n}\left(\widehat{\theta}_{c2sls} - \theta\right) \xrightarrow{d} N\left(0, \sigma^2\overline{\mathbf{H}}^{-1}\right)$.*

The 2SLS estimators of $\boldsymbol{\theta} = (\phi^S, \phi^W, \boldsymbol{\beta}')'$ considered in this paper are:

(i) *Finite-IV* : $\widehat{\boldsymbol{\theta}}_{2sls1} = (\mathbf{Z}'\mathbf{P}_1\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{P}_1\mathbf{y}$, where $\mathbf{P}_1 = \mathbf{Q}_1(\mathbf{Q}_1'\mathbf{Q}_1)_1^{-1}\mathbf{Q}_1'$ and $\mathbf{Q}_1$ contains the linearly independent columns of $\mathbf{J}[\mathbf{X}^*, \mathbf{GX}^*, \mathbf{GGX}^*]$.

(ii) *Many-IV* : $\widehat{\boldsymbol{\theta}}_{2sls2} = (\mathbf{Z}'\mathbf{P}_2\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{P}_2\mathbf{y}$, where $\mathbf{P}_2 = \mathbf{Q}_2(\mathbf{Q}_2'\mathbf{Q}_2)_2^{-1}\mathbf{Q}_2'$ and $\mathbf{Q}_2$ contains the linearly independent columns of $[\mathbf{Q}_1, \mathbf{JG}^S\boldsymbol{\iota}, \mathbf{JG}^W\boldsymbol{\iota}]$.

(iii) *Bias-corrected*: $\widehat{\boldsymbol{\theta}}_{c2sls} = (\mathbf{Z}'\mathbf{P}_2\mathbf{Z})^{-1}\{\mathbf{Z}'\mathbf{P}_2\boldsymbol{y} - \widetilde{\sigma}^2_{2sls1}[\text{tr}\left(\mathbf{P_2}\mathbf{G}^S\widetilde{\mathbf{M}}\right), \text{tr}\left(\mathbf{P_2}\mathbf{G}^W\widetilde{\mathbf{M}}\right), \mathbf{0}_{3m\times 1}]'\}$, where $\tilde{\mathbf{M}} = (\mathbf{I} - \widetilde{\phi}^S_{2sls1}\mathbf{G}^S - \widetilde{\phi}^W_{2sls1}\mathbf{G}^W)^{-1}$, and $\widetilde{\sigma}^2_{2sls1}, \widetilde{\phi}^S_{2sls1}$ and $\widetilde{\phi}^W_{2sls1}$ are $\sqrt{n}$-consistent initial estimators of σ^2, ϕ^S and ϕ^W obtained by *Finite-IV.*

(iv) *Finite-IV lagged*: $\widehat{\boldsymbol{\theta}}_{2sls1L} = (\mathbf{Z}'\mathbf{P}_3\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{P}_3\mathbf{y}$, where $\mathbf{P}_3 = \mathbf{Q}_{1L}(\mathbf{Q}_{1L}'\mathbf{Q}_{1L})_{1L}^{-1}\mathbf{Q}_{1L}'$ and $\mathbf{Q}_{1L}$ contains the linearly independent and lagged in time columns of $\mathbf{J}[\mathbf{X}^*, \mathbf{GX}^*, \mathbf{GGX}^*]$.

(v) *Many-IV lagged*: $\widehat{\boldsymbol{\theta}}_{2sls2L} = (\mathbf{Z}'\mathbf{P}_4\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{P}_4\mathbf{y}$, where $\mathbf{P}_4 = \mathbf{Q}_{2L}(\mathbf{Q}_{2L}'\ \mathbf{Q}_{2L})^{-1}\mathbf{Q}_{2L}'$ and $\mathbf{Q}_{2L}$ contains the linearly independent columns of $[\mathbf{Q}_{1L}, \mathbf{JG}^S\boldsymbol{\iota}, \mathbf{JG}^W\boldsymbol{\iota}]$

(vi) *Bias-corrected lagged*: $\widehat{\boldsymbol{\theta}}_{c2slsL} = (\mathbf{Z}'\mathbf{P}_4\mathbf{Z})^{-1}\{\mathbf{Z}'\mathbf{P}_4\mathbf{y} - \widetilde{\sigma}^2_{2sls1}[\text{tr}\left(\mathbf{P_4}\mathbf{G}^S\widetilde{\mathbf{M}}\right), \text{tr}\left(\mathbf{P_4}\mathbf{G}^W\widetilde{\mathbf{M}}\right), \mathbf{0}_{3m\times 1}]'$ where $\tilde{\mathbf{M}} = (\mathbf{I} - \widetilde{\phi}^S_{2sls1L}\mathbf{G}^S - \widetilde{\phi}^W_{2sls1L}\mathbf{G}^W)^{-1}$, and $\widetilde{\sigma}^2_{2sls1L}, \widetilde{\phi}^S_{2sls1L}$ and $\widetilde{\phi}^W_{2sls-1L}$ are $\sqrt{n}$-consistent initial estimators of σ^2, ϕ^S and ϕ^W obtained by *Finite-IV lagged.*

6 Estimation results

The aim of our empirical analysis is twofold, (i) to assess the presence of long-run peer effects in education and, (ii) following our theoretical model, to differentiate between the impact of weak ties and strong ties on education.

6.1 Long-run peer effects

Table 1 collects the estimation results of model (5), without distinguishing between strong and weak ties. The first three columns show the results when using the traditional set of instruments whereas, in the second column, the instrumental set contains only variables lagged in time (see Section 5.2.1). The first stage partial F-statistics (see Stock et al., 2002 and Stock and Yogo, 2005) reveals that our instruments are quite informative. We test the exogeneity of $\boldsymbol{G}$ using the over-identifying restrictions (OIR) test (Lee, 1992). If the OIR test cannot reject the null hypothesis that the moment conditions are correctly specified, then it provides evidence that $\boldsymbol{G}$ is uncorrelated with the error term, when $\boldsymbol{X}$, contextual effects and network fixed effects are controlled for. Table 1 shows that the p-value of the OIR test is larger than conventional significance levels, which provides evidence on *the exogeneity of network structure* (conditional on $\boldsymbol{X}$, $\boldsymbol{G}^{*}\boldsymbol{X}$ and network fixed effects).

[*Insert Table* 1 *here*]

The results in Table 1 do not change much across the columns and reveal that 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). According to the bias-corrected 2SLS estimator,[12] in a group of two friends, a standard deviation increase in the years of education of the friend translates into a roughly 5.4 percent increase of a standard deviation in the individual years of education (roughly 2 more months of education). If we consider an average group of 4 best friends (linked to each other in a network), a standard deviation increase in the level of education of each of the peers translates into a roughly 16 percent increase of a standard deviation in the individual's education attainment (roughly 7

[12] The bias-corrected 2SLS estimator is our preferred one since we have relatively small networks (see Appendix C).

more months of education). This is a non-negligible 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 seems to be carried over time.

6.2 The role of strong ties

We would like now to test our model, i.e., Proposition 1, especially equation (3) and its econometric equivalent (5), and thus determine how strong and weak ties affect educational choices by estimating the magnitude of ϕ^S and ϕ^W. As stated above, we split friendship links among teenagers in three categories: friendship links observed only in one wave, either Wave I or Wave II (*weak ties*), and friendship links observed in both waves (*strong ties*). Table 2 shows the estimation results of model (5).[13] Corollary 1 requires, as a sufficient condition, that $\phi_S + \phi_W$ is in absolute value smaller than the inverse of the largest eigenvalue of the block-diagonal network matrix $\mathbf{G}_r$, i.e. $\phi_S + \phi_W < 1/\mu_1(\mathbf{G}_r)$. Given that the largest eigenvalue of $\mathbf{G}_r$ is 6.48, the existence of the equilibrium in our model requires values for $\phi_S + \phi_W$ within the range $[0, 0.154)$. Table 2 shows that our results are within this parameter space.

We find that weak ties' educational choices have no significant impact on individual education outcomes (years of schooling) while strong ties' educational choices have a positive and significant effect on own educational outcome. 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 21 percent increase of a standard deviation in the individual's education attainment (roughly 8.3 more months of education). In an average group of 4 best friends (linked to each other in a network), a standard deviation increase of each of the peers translates into 2 more years of education. This is quite an important effect. This suggests that *strong ties* rather than *weak ties* matter for educational outcomes in the long run.

[*Insert Table* 2 *here*]

If we now go back to the theoretical model of Section 3, this result means that a strong relationship at school between two students (where $\phi = \phi^S$) has an important impact on their future educational choices. This is because the *strength of an interaction* between two students may affect how much they learn, the human capital accumulation and how

[13] We show the results for the bias-corrected 2SLS estimator, with the traditional set of instruments and when the instrumental set contains only variables lagged in time. The qualitative results when using the alternative estimators in Appendix C remain qualitatively unchanged. The latter remains available upon request.

much they value achievement. It also shapes social norms that accumulate over time, which affect years of schooling both directly and indirectly. This result is related to Akerlof's and Kranton's (2002) concept of identity in economics where learning at school can be viewed within a process of identity formation, resource allocation, and social interaction. In other words, following the sociology literature, Akerlof and Kranton (2002) postulate that students often care less about their studies than about what their friends think. Our empirical result confirms this intuition by showing that peers, especially strong ties, play an important role in a student's college decision.[14]

6.3 Understanding the mechanisms

Our results 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 ties as peers nominated in both Wave I and Wave II, which means that there is a strong and stable relationship between two students that is more likely to last.

One could put forward another explanation of why friends at school may influence education decisions: it could be the timing of friendship or *decision proximity* so that friends in the last grades (grades 10 to 12) are likely to impact on college decision, regardless on whether they are strong or weak ties. In other words, is it really the frequency and strength of social interactions (as shown in the previous section) or is it the timing of friendship formation that is crucial for future educational outcomes. We would therefore like to disentangle between the *decision proximity* effect and the *strength of interaction* effect. For that, we select students in the last grades (grades 10 to 12) and distinguish between weak and strong ties and examine the effect on college choices. We estimate a modified version of model (5), that is

$$\begin{aligned} y_{i,r,t+1} &= \phi^S \sum_{j=1}^{n_r} g^S_{ij,r,t} y_{j,r,t+1} + \phi^{W_1} \sum_{j=1}^{n_r} g^{W_1}_{ij,r,t} y_{j,r,t+1} + \phi^{W_2} \sum_{j=1}^{n_r} g^{W_2}_{ij,r,t} y_{j,r,t+1} + \frac{1}{g^S_{i,r,t}} \sum_{j=1}^{n_r} g^S_{ij,r,t} x'_{j,r,t,t+1} \gamma^S \\ &+ \frac{1}{g^{W_1}_{i,r,t}} \sum_{j=1}^{n_r} g^{W_1}_{ij,r,t} x'_{j,r,t,t+1} \gamma^{W_1} + \boldsymbol{x}'_{i,r}\delta + \frac{1}{g^{W_2}_{i,r,t}} \sum_{j=1}^{n_r} g^{W_2}_{ij,r,t} x'_{j,r,t,t+1} \gamma^{W_2} + \eta_{r,t} + \epsilon_{i,r,t+1}, \end{aligned}$$

We distinguish here between weak ties where best friends have only been nominated in Wave I (lower grades) and not in Wave II (later grades), i.e. $\phi^W = \phi^{W_1}$, from weak ties where

[14]This is also related to the empirical study of De Giorgi et al. (2010) who show that students from Bocconi University in Italy are more likely to choose a major if many of their peers make the same choice. They also show that peers can divert students from majors in which they have a relative ability advantage, with adverse consequences on academic performance.

best friends have only been nominated in Wave II and not in Wave I, i.e. $\phi^W = \phi^{W_2}$. If the decision proximity matters, then the coefficient ϕ^{W_2} should be significant while ϕ^{W_1} should not be.

Table 3 contains the estimation results. The empirical results reveal that the education decision of weak ties continue to show a non-significant effect on individual education outcomes, regardless on whether peers are interacting in lower or higher grades, highlighting the crucial role of strong ties in college decision.

[*Insert Table* 3 *here*]

To better highlight this result, Figure 2 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* (i.e. in terms of grade, how close is the student from making a college education decision) versus the *strength of interaction* effect (i.e. weak versus strong tie). 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 Table 3 show that what matters most for peer effects is the strength of interaction rather than the decision proximity effect (see also the last row of Figure 2). 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* 2 *here*]

As a result, our analysis so far seems to support entirely our theoretical mechanism. A last concern is that peers nominated in different time periods may have a different long-run effect because students value differently peer characteristics in friendship decisions made 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? To disentangle these effects, we check if students select peers differently between the first and the second wave. Table 4 compares 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.

[*Insert Table* 4 *here*]

To further investigate this issue, we test whether link formation is different between different waves. For that, we pool the data for Wave I ($t = 1$) and Wave II ($t = 2$) and estimate the following regression model:

$$g_{ij,r,t} = \alpha + \sum_{m=1}^{M} \beta_m |x_{i,r,t}^m - x_{j,r,t}^m| + \sum_{m=1}^{M} \gamma^m |x_{i,r,t}^m - x_{j,r,t}^m| \times d_{ij,r} + \epsilon_{ij,r,t}, \quad t = 1, 2, \tag{14}$$

where, as in (5), $g_{ij,r,t} = 1$ if there is a link between i and j belonging to network r at time t (where $t =$ Wave I, Wave II), $x_{i,r,t}^m$ indicates the individual characteristic m of individual i in network r at time t and $d_{ij,r}$ is a dummy which is equal to 1 if a 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 two students might be friends because they both like to play football and watch TV. The parameter in front of the dummy capture differences between the importance of these characteristics in link formation between Wave I and Wave II. Table 5 shows that most coefficients are not significant and that there are no observable differences in link formation process between Waves I and II. We have also performed an F test that tests the joint significance of the γ parameters.[15] Table 5 reports the p value of this test. It reveals that, controlling for network fixed effects, we cannot reject the null hypothesis of $\gamma^m = 0$, $\forall m = 1, \ldots, M$. In summary, Tables 4 and 5 provide evidence showing that there are no differences between peers in Waves I and II in terms of observable characteristics and that the link formation between the different waves is not different.

[*Insert Table* 5 *here*]

[15] The idea is similar to the *Chow test* in time series analysis to investigate the existence of a structural break (see e.g. Chow, 1960; Hansen, 2000, 2001).

7 Robustness checks

In this section we perform two different robustness checks.

7.1 Directed networks

Firstly, our empirical investigation have assumed that friendship relationships are symmetric, i.e. $g_{ij} = g_{ji}$. We check here how sensitive our results are to such an assumption, i.e. to a possible measurement error in the definition of the peer group. Indeed, our data make it possible to know exactly who nominates whom in a network and we find that 12 percent of relationships in our dataset are not reciprocal. Instead of constructing undirected network, in this section, we perform our analysis using *directed networks*. We focus on the choices made (outdegrees) and we denote a link from i to j as $g_{ij,r} = 1$ if i has nominated j as his/her friend in network r, and $g_{ij,r} = 0$, otherwise.[16] Table 6 shows the estimation results of model (5) for directed networks. The results remain qualitatively unchanged and only slightly higher in magnitude.

[*Insert Table* 6 *here*]

7.2 A simulation experiment

Secondly, our identification and estimation strategies depend on the correct identification of strong and weak ties. In this section, we test the robustness of our results with respect to misspecification of weak and strong network tie topologies. Indeed, in our theoretical model, we assume that $\phi^S > \phi^W$ and our empirical analysis confirms this assumption by finding a significant effect of strong ties (but not weak ties) on educational outcomes. These results clearly depend on the definition of a strong and a weak tie. In the present robustness check, we want to check if our results are robust even if we fail in exactly identifying strong ties. To be more precise, we use simulated data to answer questions such as: Do our results change if some links are not assigned in the right category (weak or strong ties)? To what extent? How many ties need to be misspecified before explaining away our results?

Numerical experiment

We use a simulation approach to randomly change (p_r) a certain percentage of ties from weak to strong (and vice versa) in each network r, one hundred times for each value of p_r going from 0 to 1 with a pace of 0.005. We thus draw one hundred network structures

[16] As highlighted by Wasserman and Faust (1994), centrality indices for directional relationships generally focus on choices made (outdegrees). The estimation results, however, remain qualitatively unchanged if we define link using the nominations received (indegrees).

(samples) of size equal to the real one ($n = 1,819$) 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. $p_r = p$.

Let S_r, with cardinality s_r, be the set of existent strong ties in network r and W_r, with cardinality w_r, the set of weak ties in the same network. The number of "possible changes" is consequently $c_r = \min(s_r, w_r)$.

In words, for each network r, we can exchange only a fraction of strong ties with weak ties (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 ties. The actual percentage will be $q_r = p_r \times c_r$.[17]

Simulated evidence

Our link replacement procedure enables us to simulate different network structures ($\mathbf{G}_r^S$ and $\mathbf{G}_r^W$ matrices in model (7)) that differ from the real ones by a given (increasing) number of misspecified ties. Note that we change here the structure of $\mathbf{G}_r^S$ and $\mathbf{G}_r^W$, maintaining the same $\mathbf{G}_r$. As mentioned before, for each percentage of randomly replaced ties, we draw 100 network structures (samples) of size and network density equal to the real one. We then estimate model (7) replacing the real $\mathbf{G}_r^S$ and $\mathbf{G}_r^W$ matrices with the simulated ones in turn, so that in total we estimate model (7) twenty thousand times for each type of estimator (see Section 5.2.1).

[*Insert Figure* 3 *here*]

Figure 3 shows the results of our simulation experiment. The upper panels depict the estimates of strong and weak ties effects, whereas the lower panels show the t-statistics with 90% confidence bands. It appears that the higher the percentage of misspecified links, the wider is the range of the peer effects estimates and the t-statistics fail more often to reject the hypothesis of no effects for strong ties and to accept it for weak ties.

The first important question concerns the percentage of network-structure misspecifica-

[17] An 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% 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 t_r is:

$$t_r = \begin{cases} [q_r] & if \;\; u > a \\ [q_r] + 1 & if \;\; u < a \end{cases}$$

where $a = q_r - [q_r]$, and u is a random extraction from a variable uniformly distributed on $(0,1)$.

tion needed for the *strong-tie* peer effects on college choice to disappear. Lets us focus on the lagged 2SLS bias-corrected estimator, which is the most appropriate one in our case (see Section 5.2.1).

[*Insert Figure* 4 *here*]

Figure 4 plots the averages of the estimates of strong and weak ties effects for each replacement rate with 90% confidence bands. Standard errors have been calculated assuming drawing independence and taking into account the variation between estimates for each replacement rate.[18] Panel (a) shows that strong-tie effects remain statistically significant up to a percentage of randomly replaced (interchangeable) links of about 30%. This implies that, even if we do not observe or we imprecisely observe a portion of each individual's strong ties, our results on the existence of this effect hold.

The second question we address is what is the percentage of replacement needed to have a significant effect of weak ties. Panel (b) in Figure 4 shows that this threshold is quite high since we find a significant effect for weak ties only after having replaced more that 90% of ties. In other words, the effects of weak ties is found to be important only when the large majority of our initial (true) strong ties are labeled as weak ties.

Finally, we show in Figure 5 the rejection rates[19] when using the *2SLS bias-corrected estimator* and the *2SLS bias-corrected lagged estimator.* This graph indicates that the *2SLS bias-corrected lagged estimator* tends to be more robust to a possible misspecification of strong and weak ties. Indeed, it appears that this estimator needs, on average, an higher percentage of misspecified ties to accept the hypothesis of no effects for strong ties and to reject it for weak ties.

[*Insert Figure* 5 *here*]

To wrap up, in this section, we have shown that the strength of strong ties ϕ^S is reduced and becomes insignificant when we have converted more than 30% of strong ties into weak ties (and vice versa) while the strength of weak ties ϕ^W is increasing and becomes significant

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

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

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

[19] Rejection refers to the null hypothesis of having $\phi^S = 0$ or $\phi^W = 0$ respectively for strong and weak ties effects. Each rate represents the frequency of rejection for the correspondent percentage of randomly replaced links.

after having replaced 90% of ties. This implies that, even if we did not observe or imprecisely observe a portion of each individual's strong ties, our results on the strength of strong ties will still hold and the effect of weak ties would only be important when the large majority of our initial (true) strong ties have been labeled as weak ties. To illustrate this result, consider a student i who has twenty friends, ten strong ties $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ and ten weak ties $\{11, 12, 13, 14, 15, 16, 17, 18, 19, 20\}$. Even if we wrongly assign 3 friends from one category (strong tie) to the other (weak tie), our results will still hold. For instance, if we instead observe $\{11, 12, 3, 4, 5, 6, 7, 8, 9, 10\}$ as strong ties (labeling 11 and 12 as strong when they are weak ties) and $\{1, 2, 13, 14, 15, 16, 17, 18, 19, 20\}$ as weak ties (labeling 1 and 2 as weak when they are strong ties), we would still have a significant effect of strong ties on education and a non-significant effect of weak ties since we have "only" converted 30% of the links. As a result, from 3 to 8 wrong assignments (which corresponds to 30% to 80% conversion of strong ties into weak ties or the contrary) both effects will still be insignificant. It is only after having converted 9 of ten ties (i.e. 90% of strong ties have been converted into weak ties or the contrary) that we find that weak ties have a significant effect on education while strong ties do not.

8 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) influence them in their educational choices. Using Addhealth data for Wave I only, Calvó-Armengol et al. (2009) have studied the *current* effect of peers on education, finding that peers do affect the current education activity (i.e. grades) of students. They did not differentiate between weak ties and strong ties.

To investigate further this issue, we would like now to oppose the long-run effects to the short-run effects of peers on education by differentiating between the effect of strong ties and weak ties 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). For that, we estimate the short-run counterpart of equation (5):

$$\begin{aligned} y_{i,r,t} &= \phi^S \sum_{j=1}^{n_r} g^S_{ij,r,t} y_{j,r,t} + \phi^W \sum_{j=1}^{n_r} g^W_{ij,r,t} y_{j,r,t} + \boldsymbol{x}'_{i,r}\delta \\ &\quad + \frac{1}{g^S_{i,r,t}} \sum_{j=1}^{n_r} g^S_{ij,r,t} x'_{j,r,t}\gamma^S + \frac{1}{g^W_{i,r,t}} \sum_{j=1}^{n_r} g^W_{ij,r,t} x'_{j,r,t}\gamma^W + \eta_{r,t} + \epsilon_{i,r,t}, \end{aligned} \tag{15}$$

where $y_{i,r,t}$ is now the grade of student i who belongs to network r at time t where t refers to Wave II. The rest of the notation remains 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 1995-1996). 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 (15) for peers who are only friends in Wave I, then for peers who are only friends in Wave II and finally for peers who are friends in both Waves I and II. We will look at how peers (as defined above) affect each student's grade obtained in Wave II. The identification and estimation strategy remain unchanged with the difference that we cannot now use IV variables lagged in time (see Section 5.2.1). 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.[20] It ranges between 0 and 6.09, with mean equal to 2.29 and standard deviation equal to 1.49.

The estimation results of model (15) are contained in Table 7. It appears that, *while in the long run, only strong ties matter, in the short run, both weak and strong ties are important in determining a student's performance at school.* A standard deviation increase in aggregate GPA of peers translates respectively into a 8.1 (for strong ties) and a 4.8 (for weak ties) percent increase of a standard deviation in the individual's GPA.

[*Insert Table* 7 *here*]

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).

9 Concluding remarks and policy implications

This paper provides a microfoundation for a spatial autoregressive model with different weight matrices. This model has features and implications directly relevant to social inter-

[20]The index explains roughly the 56 percent of the total variance and captures a general performance at school since it is positively and highly correlated to the scores in all subjects. Further details on this procedure are available upon request.

action issues. In particular, it allows to investigate *heterogenous peer effects*. We illustrate the practical importance of this issue by looking at the impact of friends made at school on college decision of students. We find that there are long-term peer effects. 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 and other school-based 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[21] 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 relationships and it is still possible for him/her to create such relationship in the new environment. This, of course, would be 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.

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

References

[1] Akerlof, G.A. and R.E. Kranton (2002), "Identity and schooling: Some lessons for the economics of education," *Journal of Economic Literature* 40, 1167-1201.

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

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

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

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

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

[7] 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.

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

[9] 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.

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

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

[12] Chow, G.C. (1960), "Test of equality between sets of coefficients in two linear regressions," *Econometrica* 28, 591-603.

[13] 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.

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

[15] 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.

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

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

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

[19] 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.

[20] Donald, S.G., and W.K. Newey (2001), "Choosing the number of Instruments," *Econometrica* 1161-1191.

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

[22] 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.

[23] 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.

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

[25] 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.

[26] 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.

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

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

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

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

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

[32] Hansen, B.E. (2000), "Testing for structural change in conditional models," *Journal of Econometrics* 97, 93-115.

[33] Hansen, B.E. (2001), "The new econometrics of structural change: Dating breaks in U.S. labor productivity," *Journal of Economic Perspectives* 15, 117-128.

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

[35] 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.

[36] Ioannides, Y.M. (2012), *From Neighborhoods to Nations: The Economics of Social Interactions*, Princeton: Princeton University Press.

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

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

[39] 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.

[40] 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.

[41] 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.

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

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

[44] 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.

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

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

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

[48] Lee, L-F. and Liu X. (2010) "Efficient GMM Estimation of High Order Spatial Autoregressive Models with Autoregressive Disturbances", *Econometric Theory* 26, 187-230.

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

[50] 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.

[51] 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.

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

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

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

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

[56] Meyer, C.D. (2000), *Matrix Analysis and Applied Linear Algebra*, Philadelphia: SIAM.

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

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

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

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

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

[62] 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.

[63] 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.

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

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

[66] 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.

[67] 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.

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

[69] Zenou, Y. (2013), "Spatial versus social mismatch," *Journal of Urban Economics* 74, 113-132.

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

Appendix A: Proofs of propositions and corollaries in the text

Proof of Proposition 1: We need to show that $\mathbf{I}-\mathbf{A}$ is non-singular (i.e. invertible), where $\mathbf{A}\equiv\phi^S\mathbf{G}^S+\phi^W\mathbf{G}^W$. We know that $\mathbf{I}-\mathbf{A}$ is non-singular if $\mu_1\left(\phi^S\mathbf{G}^S+\phi^W\mathbf{G}^W\right)<1$ (see, e.g. Meyer, 2000, page 618). To prove the interiority of the solution, we can use exactly the same arguments as in the proof of Theorem 1 in Ballester et al. (2006). ▮

Proof of Corollary 1: Let us start with two lemmas. Denote by $\mu_1(\mathbf{A})$ the spectral radius of $\mathbf{A}$.

Lemma 1 *If $\boldsymbol{A}$ is an $n\times n$ Hermitian matrix, then*

$$\mu_1(\boldsymbol{A})=\sup_{|\boldsymbol{v}|=\mathbf{1}}\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{v} \tag{16}$$

Proof: The *Rayleigh–Ritz quotient* is defined as:

$$R(\boldsymbol{A},\boldsymbol{x})=\frac{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}$$

where $\boldsymbol{x}^T$ is transpose of $\boldsymbol{x}$. This is equal to

$$R(\boldsymbol{A},\boldsymbol{x})=\frac{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}=\frac{\sum_{i=1}^n\sum_{j=1}^n x_ix_ja_{ij}}{\boldsymbol{x}^T\boldsymbol{x}}$$

That is

$$R(\boldsymbol{A},\boldsymbol{x})=\frac{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}=\frac{\frac{\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2}}{\frac{\boldsymbol{x}^T\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2}}=\frac{\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)^T\boldsymbol{A}\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)}{\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)^T\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)}$$

where $\|\boldsymbol{x}\|_2\equiv\sqrt{\left(\sum_{i=1}^n|x_i|^2\right)}$ is the Euclidian norm (or vector 2-norm). We want to compute

$$\sup_{\boldsymbol{x}\in\mathbb{R}^n\setminus\{\mathbf{0}\}}R(\boldsymbol{A},\boldsymbol{x})=\sup_{\boldsymbol{x}\in\mathbb{R}^n\setminus\{\mathbf{0}\}}\frac{\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)^T\boldsymbol{A}\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)}{\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)^T\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)}$$

Define $\boldsymbol{y}\equiv\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}$ so that $\boldsymbol{y}^T\boldsymbol{y}=\mathbf{1}$, then

$$\sup_{\boldsymbol{x}\in\mathbb{R}^n\setminus\{\mathbf{0}\}}R(\boldsymbol{A},\boldsymbol{x})=\sup_{\boldsymbol{y}\in\mathbb{R}^n\setminus\{\mathbf{0}\}}\left\{\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}\mid\|\boldsymbol{y}\|_2=1\right\}$$

We want to show that $\sup_{\boldsymbol{y}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \left\{\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} \mid \|\boldsymbol{y}\|_2 = 1\right\} = \mu_1(\boldsymbol{A})$. Observe that the function $\boldsymbol{y} \to \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}$ is continuous with compact domain ($n-1$ dimensional sphere). Every continuous function attains a maximum and a minimum on a compact set. There is thus a $\boldsymbol{y}$ so that the sup is a maximum. The problem is to find the critical points of the function $\boldsymbol{y} \to \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}$ subject to the constraint $\|\boldsymbol{y}\|_2^2 = \boldsymbol{y}^T\boldsymbol{y} = 1$, i.e. to find the critical points of the following lagrangian

$$\mathcal{L}(\boldsymbol{y}) = \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} - \mu\left(\boldsymbol{y}^T\boldsymbol{y} - 1\right)$$

where μ is a Lagrange multiplier. The stationary points of $\boldsymbol{y} \to \mathcal{L}(\boldsymbol{y})$ are given by: $\frac{d\mathcal{L}(\boldsymbol{y})}{d\boldsymbol{y}} = \mathbf{0}$. We obtain:

$$\frac{d\mathcal{L}(\boldsymbol{y})}{d\boldsymbol{y}} = \left(\boldsymbol{A}+\boldsymbol{A}^T\right)\boldsymbol{y} - \mu\left(\boldsymbol{I}+\boldsymbol{I}^T\right)\boldsymbol{y} = \mathbf{0}$$

Using the fact that $\boldsymbol{A}$ is *symmetric*, i.e. $\boldsymbol{A}=\boldsymbol{A}^T$, then

$$\frac{d\mathcal{L}(\boldsymbol{y})}{d\boldsymbol{y}} = 2\boldsymbol{A}\boldsymbol{y} - 2\mu\boldsymbol{y} = \mathbf{0}$$

which is equivalent to

$$\boldsymbol{A}\boldsymbol{y} = \mu\,\boldsymbol{y}$$

This means that $\boldsymbol{y}$ is an eigenvector and μ the associated eigenvalue. Therefore, the eigenvectors $\boldsymbol{y}_1,...,\boldsymbol{y}_n$ of $\boldsymbol{A}$ are the critical points of the Rayleigh Quotient and their corresponding eigenvalues $\mu_1, ..., \mu_n$, with $\mu_1 \geq ... \geq \mu_n$, are the stationary values of $R(\boldsymbol{A}, \boldsymbol{x})$.

At the stationay points, we have

$$\boldsymbol{y}^T\mathbf{A}\mathbf{y} = \boldsymbol{y}^T\mu\,\boldsymbol{y} = \boldsymbol{\mu}\,\boldsymbol{y}^T\boldsymbol{y} = \mu$$

This implies that

$$\sup_{\boldsymbol{x}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} R(\boldsymbol{A}, \boldsymbol{x}) = \sup_{\boldsymbol{y}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \left\{\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} \mid \|\boldsymbol{y}\|_2 = 1\right\} = \mu_1(\boldsymbol{A})$$

We have thus shown that $\mu_1(\boldsymbol{A}) = \sup_{|\boldsymbol{v}|=\mathbf{1}} \boldsymbol{v}^T\boldsymbol{A}\boldsymbol{v}$. ∎

Lemma 2 *If* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are two* $n \times n$ *symmetric matrices, then*

$$\mu_1(\boldsymbol{A}+\boldsymbol{B}) \leq \mu_1(\boldsymbol{A}) + \mu_1(\boldsymbol{B}) \tag{17}$$

Proof: In Lemma 1, we have shown that

$$\mu_1(\boldsymbol{A}) = \sup_{|\boldsymbol{v}|=\mathbf{1}} \boldsymbol{v}^T \boldsymbol{A} \boldsymbol{v}$$

which implies that

$$\mu_1(\boldsymbol{B}) = \sup_{|\boldsymbol{v}|=\mathbf{1}} \boldsymbol{v}^T \boldsymbol{B} \boldsymbol{v}$$

and

$$\mu_1(\boldsymbol{A}+\boldsymbol{B}) = \sup_{|\boldsymbol{v}|=\mathbf{1}} \boldsymbol{v}^T (\boldsymbol{A}+\boldsymbol{B}) \boldsymbol{v}$$

We need now to show that

$$\mu_1(\boldsymbol{A}+\boldsymbol{B}) \leq \mu_1(\boldsymbol{A}) + \mu_1(\boldsymbol{B})$$

Using the sub-additivity of the sup function, we have

$$\begin{aligned}
\mu_1(\boldsymbol{A}+\boldsymbol{B}) &= \sup_{|\boldsymbol{v}|=\mathbf{1}} \boldsymbol{v}^T (\boldsymbol{A}+\boldsymbol{B}) \boldsymbol{v} \\
&= \sup_{|\boldsymbol{v}|=\mathbf{1}} \left\{ \boldsymbol{v}^T \boldsymbol{A} \boldsymbol{v} + \boldsymbol{v}^T \boldsymbol{B} \boldsymbol{v} \right\} \\
&\leq \sup_{|\boldsymbol{v}|=\mathbf{1}} \boldsymbol{v}^T \boldsymbol{A} \boldsymbol{v} + \sup_{|\boldsymbol{v}|=\mathbf{1}} \boldsymbol{v}^T \boldsymbol{B} \boldsymbol{v} \\
&= \mu_1(\boldsymbol{A}) + \mu_1(\boldsymbol{B})
\end{aligned}$$

which is the statement of the lemma. ∎

In Proposition 1, we have shown that $\mathbf{I} - \mathbf{A}$ is non-singular if $\mu_1\left(\phi^S \mathbf{G}^S + \phi^W \mathbf{G}^W\right) < 1$. Assume now that $\mathbf{G}$ is symmetric so that both $\mathbf{G}^S$ and $\mathbf{G}^W$ also are symmetric. We can use Lemma 2, which states that (given $\phi^S > 0$ and $\phi^W > 0$):

$$\begin{aligned}
\mu_1\left(\phi^S \mathbf{G}^S + \phi^W \mathbf{G}^W\right) &\leq \mu_1\left(\phi^S \mathbf{G}^S\right) + \mu_1\left(\phi^W \mathbf{G}^W\right) \\
&= \phi^S \mu_1\left(\mathbf{G}^S\right) + \phi^W \mu_1\left(\mathbf{G}^W\right)
\end{aligned}$$

For $\mathbf{A} = \{a_{ij}\}$ and $\mathbf{B} = \{b_{ij}\}$, we say $\mathbf{A} \leq \mathbf{B}$ if $a_{ij} \leq b_{ij}$ for all i, j. In our context, this means that $0 \leq \mathbf{G}^S \leq \mathbf{G}$ and $0 \leq \mathbf{G}^W \leq \mathbf{G}$ (since $g_{ij}^S \leq g_{ij}$ and $g_{ij}^W \leq g_{ij}$ for all i, j). As a result, using Theorem I*, page 600 in Debreu and Herstein (1953), we have: $\mu_1\left(\mathbf{G}^S\right) \leq \mu_1(\mathbf{G})$ and $\mu_1\left(\mathbf{G}^W\right) \leq \mu_1(\mathbf{G})$. This implies that: $\phi^S \mu_1\left(\mathbf{G}^S\right) + \phi^W \mu_1\left(\mathbf{G}^W\right) \leq \left(\phi^S + \phi^W\right) \mu_1(\mathbf{G})$ and the condition for $\mathbf{I}-\phi^{S}\mathbf{G}^{S}-\phi^{W}\mathbf{G}^{W}$ to be non-singular is given by:

$$\left(\phi^{S}+\phi^{W}\right)\mu_{1}\left(\mathbf{G}\right)<1$$

which is the condition given in Corollary 1. ▮

Proof of Proposition 2 Let $\mathbf{JZ}=\mathbf{J}(f+v)$, where $v=[\mathbf{G^{S}M}\boldsymbol{\epsilon},\mathbf{G^{W}M}\boldsymbol{\epsilon},\mathbf{0}_{n\times 3m}]$. Assuming Lemma B.1-3 in Liu and Lee (2010) and Lemma A.3 in Donald and Newey (2001), we have

$$\begin{aligned}\frac{1}{n}\mathbf{Z}'\mathbf{P_{K}Z} &= \mathbf{H}-e_{f}+\frac{1}{n}v'\mathbf{P_{K}}f+\frac{1}{n}f'\mathbf{P_{K}}v+\frac{1}{n}v'\mathbf{P_{K}}v\\ &= \mathbf{H}+\mathbf{O}(\mathrm{tr}(e_{f}))+\mathbf{O}_{p}(\sqrt{K/n})+\mathbf{O}_{p}(K/n)\\ &= \overline{\mathbf{H}}+o_{p}(1)\end{aligned}$$

where $\mathbf{H}=\frac{1}{n}f'f$ and $e_{f}=\frac{1}{n}f'(I-\mathbf{P_{K}})f$, because $e_{f}=\mathbf{O}(\mathrm{tr}(e_{f}))$, $\frac{1}{n}v'\mathbf{P_{K}}v=\mathbf{O}_{p}(K/n)$ and $\frac{1}{n}v'\mathbf{P_{K}}f=\mathbf{O}_{p}(\sqrt{K/n})$. Furthermore we have

$$\begin{aligned}&(\mathbf{Z}'\mathbf{P_{K}}\boldsymbol{\epsilon}-\sigma^{2}\left[\mathrm{tr}\left(\boldsymbol{\Psi}_{K,S}\right),\ \ \mathrm{tr}\left(\boldsymbol{\Psi}_{K,W}\right),\mathbf{0}_{3m\times 1}\right]')/\sqrt{n}\\ =\ & h-f'(I-\mathbf{P_{K}})\boldsymbol{\epsilon}/\sqrt{n}+(\frac{1}{n}v'\mathbf{P_{K}}\boldsymbol{\epsilon}-\sigma^{2}\left[\mathrm{tr}\left(\boldsymbol{\Psi}_{K,S}\right),\mathrm{tr}\left(\boldsymbol{\Psi}_{K,W}\right),\mathbf{0}_{3m\times 1}\right]')/\sqrt{n}\\ =\ & h+\mathbf{O}_{p}(\sqrt{\mathrm{tr}(e_{f})})+\mathbf{O}_{p}(\sqrt{K/n})\\ =\ & h+o_{p}(1)\xrightarrow{d}N\left(0,\sigma^{2}\overline{\mathbf{H}}\right)\end{aligned}$$

where $h=f'\boldsymbol{\epsilon}/\sqrt{n}$. Then, applying the Slutzky theorem, the proposition follows. ▮

Proof of Proposition 3. We need to show that

$$\{\widetilde{\sigma}^{2}\left[\mathrm{tr}\left(\mathbf{P_{K}G}^{S}\widetilde{\mathbf{M}}\right),\mathrm{tr}\left(\mathbf{P_{K}G}^{W}\widetilde{\mathbf{M}}\right)\right]'-\sigma^{2}\left[\mathrm{tr}\left(\boldsymbol{\Psi}_{K,S}\right),\mathrm{tr}\left(\boldsymbol{\Psi}_{K,W}\right)\right]'\}/\sqrt{n}=o_{p}(1)$$

where $\widetilde{\mathbf{M}} = \mathbf{M}(\widetilde{\phi}^S, \widetilde{\phi}^W)$. Given Proposition 2, this is quite straightforward since

$$\begin{aligned}
& \{\widetilde{\sigma}^2 \left[\mathrm{tr}\left(\mathbf{P_K G}^S \widetilde{\mathbf{M}}\right), \mathrm{tr}\left(\mathbf{P_K G}^W \widetilde{\mathbf{M}}\right)\right]' - \sigma^2 \left[\mathrm{tr}\left(\mathbf{\Psi}_{K,S}\right), \mathrm{tr}\left(\mathbf{\Psi}_{K,W}\right)\right]'\}/\sqrt{n} \\
= \; & \sqrt{n}(\widetilde{\sigma}^2 - \sigma^2) \left[\mathrm{tr}\left(\mathbf{P_K G}^S \widetilde{\mathbf{M}}\right), \mathrm{tr}\left(\mathbf{P_K G}^W \widetilde{\mathbf{M}}\right)\right]'/n \\
& + \sqrt{n}\sigma^2 \left\{\mathrm{tr}\left[\mathbf{P_K G}^S (\widetilde{\mathbf{M}} - \mathbf{M})\right], \mathrm{tr}\left[\mathbf{P_K G}^W (\widetilde{\mathbf{M}} - \mathbf{M})\right]\right\}'/n \\
= \; & \sqrt{n}(\widetilde{\sigma}^2 - \sigma^2) \left[\mathrm{tr}\left(\mathbf{P_K G}^S \widetilde{\mathbf{M}}\right), \mathrm{tr}\left(\mathbf{P_K G}^W \widetilde{\mathbf{M}}\right)\right]'/n \\
& + \sqrt{n}\sigma^2 \left[(\widetilde{\phi}^S - \phi^S)\mathrm{tr}\left(\mathbf{P_K G}^S \widetilde{\mathbf{M}} \mathbf{G}^S \mathbf{M}\right), 0\right]'/n \\
& + \sqrt{n}\sigma^2 \left[(\widetilde{\phi}^W - \phi^W)\mathrm{tr}\left(\mathbf{P_K G}^S \widetilde{\mathbf{M}} \mathbf{G}^W \mathbf{M}\right), (\widetilde{\phi}^S - \phi^S)\mathrm{tr}\left(\mathbf{P_K G}^W \widetilde{\mathbf{M}} \mathbf{G}^S \mathbf{M}\right)\right]'/n \\
& + \sqrt{n}\sigma^2 \left[0, (\widetilde{\phi}^W - \phi^W)\mathrm{tr}\left(\mathbf{P_K G}^W \widetilde{\mathbf{M}} \mathbf{G}^W \mathbf{M}\right)\right]'/n \\
= \; & \mathbf{O}_p(\sqrt{K/n}) \\
= \; & o_p(1)
\end{aligned}$$

because $\widetilde{\mathbf{M}} - \mathbf{M} = \widetilde{\mathbf{M}}[(\widetilde{\phi}^S - \phi^S)\mathbf{G}^S\mathbf{M} + (\widetilde{\phi}^W - \phi^W)\mathbf{G}^W\mathbf{M}]$, as a special case of Lemma C.11 in Lee and Liu (2010). ∎

Appendix B: Data Appendix

Table B1 provides a detailed description of the variables used in our study as well as the summary statistics for our sample. 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 the highest mark in mathematics. On average, these adolescents declare having the same number of best friends both in wave I and II (about 2.50 friends), although the composition of the friends changes.

[*Insert Table B*1 *here*]

Figure 2: Strength of interactions or college decision proximity?

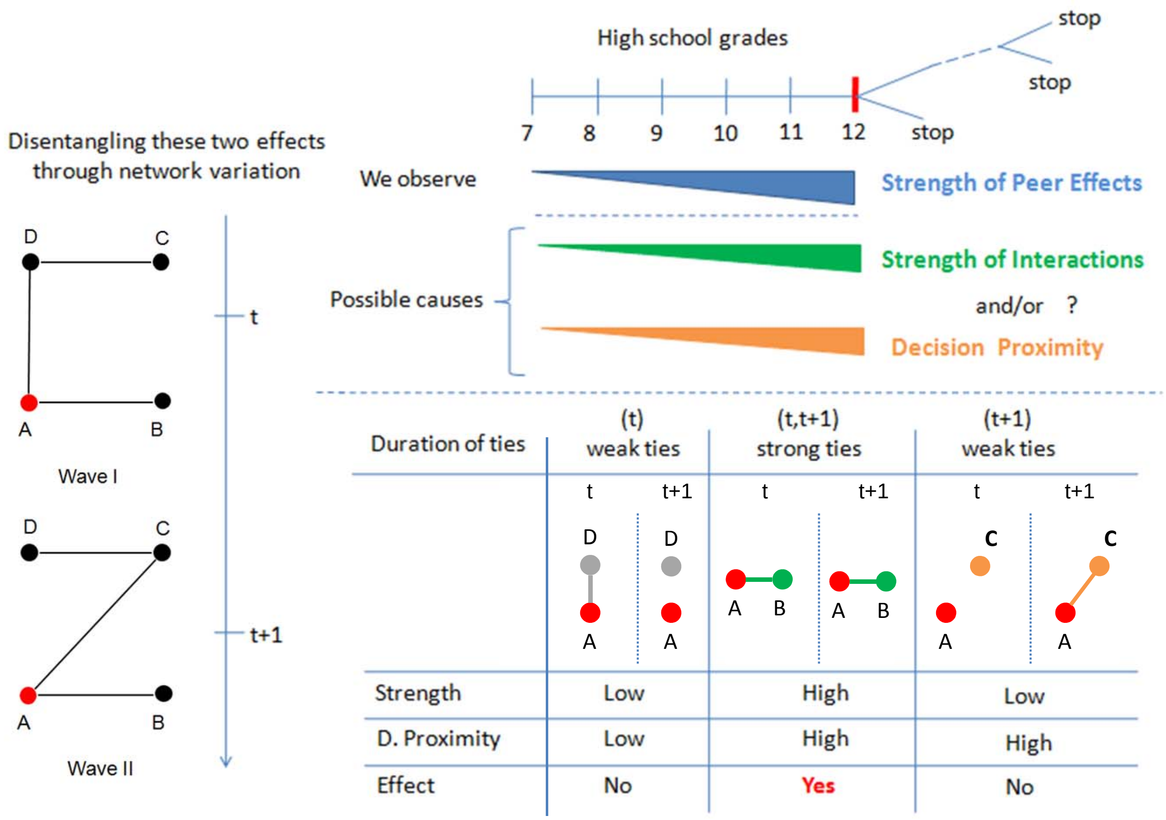
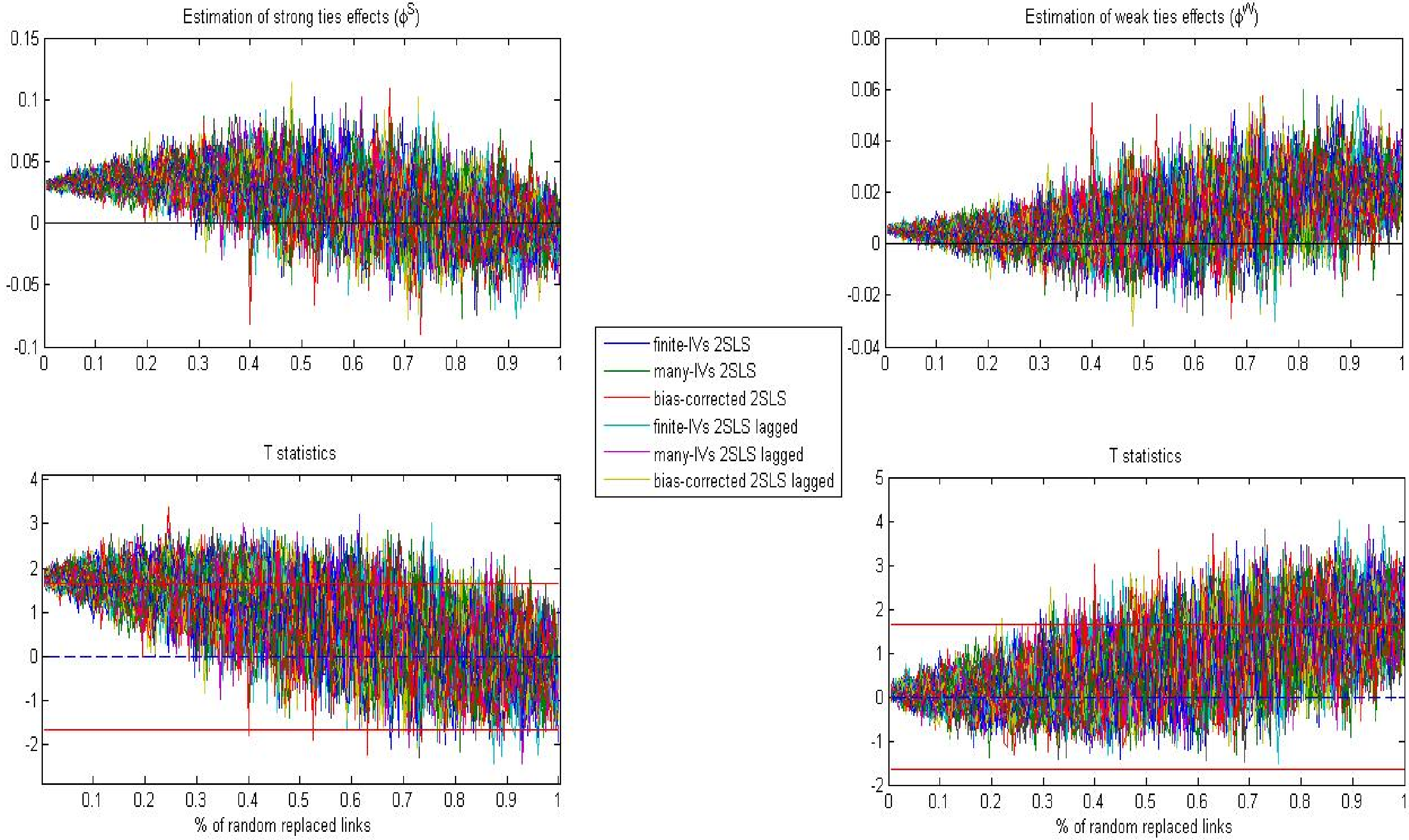
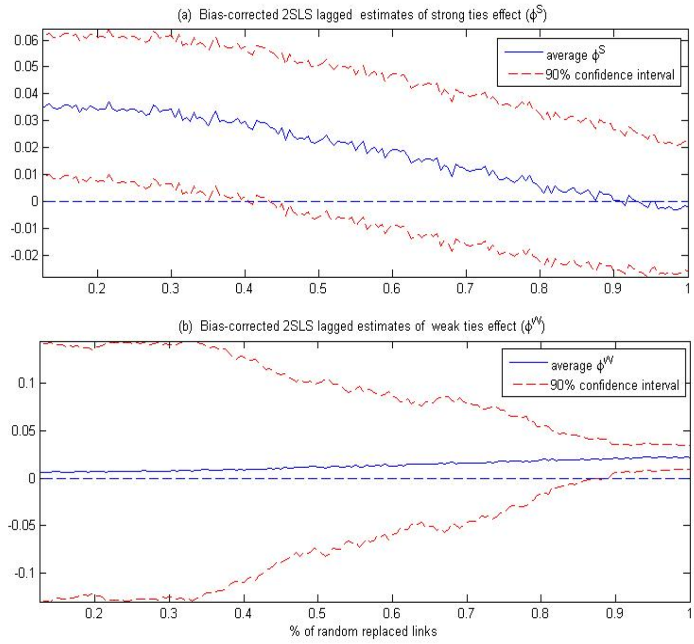


Figure 3. Misspecification of strong and weak ties
Numerical simulations



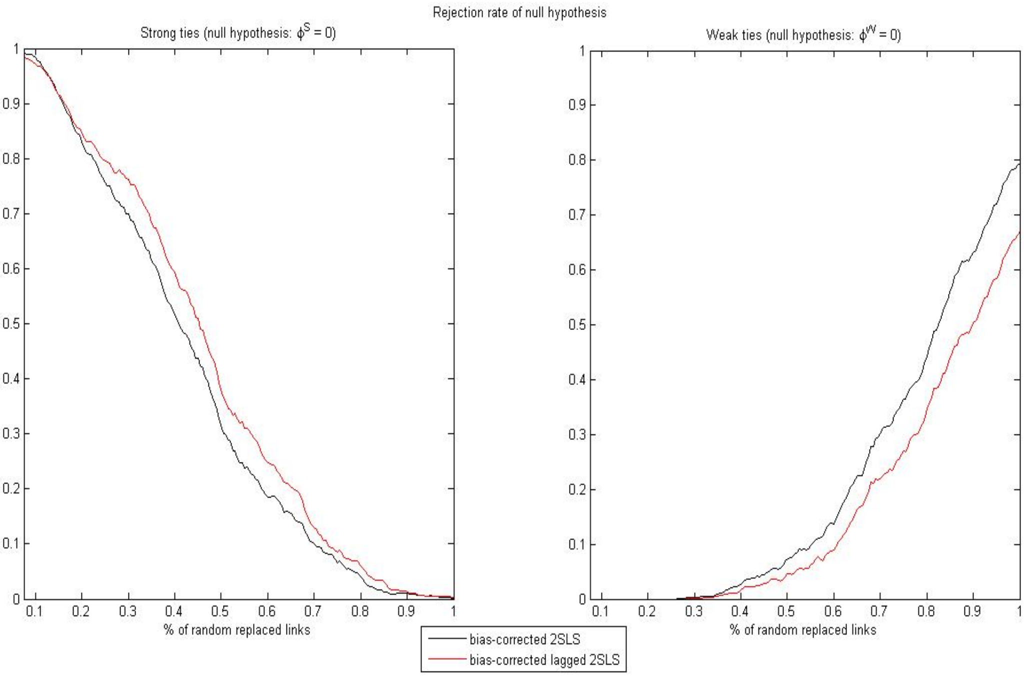
Notes: For each percentage of randomly replaced links, we draw 100 samples of size and network density equal to the real one and show the estimated weak and strong ties effects and t-statistics.

Figure 4. Simulation experiment
Summarizing the evidence



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 and between sample variation and assuming drawing independence.

Figure 5. Rejection rates of the null hypothesis
Comparing estimators



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 rejection rate of null hypothesis for weak and strong ties effects.

Table B1: Data description and and summary statistics

Variables	Description	Average (Std.Dev.)	Min - Max
Wave II (grade 7-12)			
Individual socio-demographic			
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
GPA	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. The final composite index is the first principal component score.	2.29 (1.49)	0 - 6.09
GPA of peers	Sum of GPA attained by respondent's peers	11.36 (8.85)	0 - 53.06
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
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			
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 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			
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, 3= poorly kept , 2= fairly well kept, 1= very well kept.	1.52 (0.80)	1 - 4
Contextual effects	Average of peers' characteristics for all of listed variables.		
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's peers.	40.93 (29.48)	9 - 326
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
Networks			
Links in Wave I	Number of individual links in Wave I.	2.60 (2.57)	1 - 21
Links in Wave II	Number of individual links in Wave II.	2.49 (2.50)	1 - 26
Deleted links	Percentage of nominations in Wave I not renewed in Wave II.	0.61 (0.37)	0 - 1
New links	Percentage of new nominations in Wave II.	0.44 (0.36)	0 - 1
Strong Ties	Percentage of Strong ties on total individual links.	0.28 (0.28)	0 - 1
Weak Ties	Percentage of Weak ties on total individual links.	0.72 (0.29)	0 - 1

Table 1: Long-run peer effects

Dep.Var. Years of Education	2SLS			2SLS Lagged		
	Finite IV	Many IV	Bias Corrected	Finite IV	Many IV	Bias Corrected
Peer effects (ϕ)	0.0057 ***	0.0052 ***	0.0052 ***	0.0064 ***	0.0058 ***	0.0059 ***
	(0.0020)	(0.0019)	(0.0019)	(0.0026)	(0.0020)	(0.0020)
Female	0.9702 ***	0.9718 ***	0.9718 ***	0.7383	1.0433 ***	1.0434 ***
	(0.2004)	(0.2004)	(0.2004)	(0.6123)	(0.2408)	(0.2408)
Black or African American	-0.1346	-0.1341	-0.1341	-0.2933	-0.1833	-0.1830
	(0.4088)	(0.4088)	(0.4088)	(0.7722)	(0.4464)	(0.4464)
Other races	-0.3445	-0.3447	-0.3447	-0.3873	-0.2930	-0.2927
	(0.2912)	(0.2911)	(0.2911)	(0.4561)	(0.3111)	(0.3111)
Religion Practice	0.2515 ***	0.2521 ***	0.2521 ***	0.2515 ***	0.2521 ***	0.2521 ***
	(0.0540)	(0.0540)	(0.0540)	(0.1373)	(0.0599)	(0.0599)
Household Size	0.0325	0.0327	0.0327	0.0247	0.0355	0.0355
	(0.0574)	(0.0574)	(0.0574)	(0.0900)	(0.0605)	(0.0605)
School Attachment	-0.0588	-0.0605	-0.0605	-0.0308	-0.0534	-0.0533
	(0.0776)	(0.0775)	(0.0775)	(0.0998)	(0.0805)	(0.0805)
Parent education	0.2890 ***	0.2895 ***	0.2895 ***	0.2275 *	0.2262 ***	0.2262 ***
	(0.0969)	(0.0969)	(0.0969)	(0.1528)	(0.1051)	(0.1051)
Mathematics score A	1.3550 ***	1.3578 ***	1.3578 ***	0.8917 *	1.2388 ***	1.2385 ***
	(0.2551)	(0.2551)	(0.2551)	(0.5065)	(0.2816)	(0.2816)
Mathematics score B	0.9691 ***	0.9683 ***	0.9683 ***	0.6889 *	0.8941 ***	0.8941 ***
	(0.2399)	(0.2399)	(0.2399)	(0.3966)	(0.2532)	(0.2532)
Mathematics score C	0.5254 ***	0.5272 ***	0.5272 ***	0.3422	0.5110 **	0.5107 **
	(0.2583)	(0.2583)	(0.2583)	(0.4019)	(0.2709)	(0.2709)
Mathematics score missing	0.6214	0.6210	0.6210	1.1487	0.6678	0.6677 *
	(0.4286)	(0.4286)	(0.4286)	(0.8355)	(0.4471)	(0.4471)
Resid. building qual.	-0.1771 **	-0.1798 **	-0.1797 **	-0.0950 **	-0.1264 **	-0.1262 **
	(0.0966)	(0.0966)	(0.0966)	(0.1669)	(0.1053)	(0.1053)
Student grade	0.4781 ***	0.4766 ***	0.4766 ***	0.4051 ***	0.5042 ***	0.5043 ***
	(0.0851)	(0.0850)	(0.0850)	(0.1654)	(0.0914)	(0.0914)
children	-0.4171 ***	-0.4157 ***	-0.4157 ***	-1.5107	-1.2429 **	-1.2438**
	(0.1668)	(0.1668)	(0.1668)	(1.7631)	(0.6263)	(0.6263)
Religion Practice (Wave 4)	0.1511 ***	0.1514 ***	0.1514 ***	1.0470	0.3429 **	0.3427 **
	(0.0514)	(0.0514)	(0.0514)	(1.1264)	(0.1873)	(0.1873)
Married	-0.1168	-0.1166	-0.1166	-0.8750	-0.6020	-0.6024
	(0.1655)	(0.1655)	(0.1655)	(2.113)	(0.6381)	(0.6381)
Parental occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes
Contextual effects	Yes	Yes	Yes	Yes	Yes	Yes
Network fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
First stage F statistic	1210.64	700.35		1240.51	708.83	
OIR test p-value	0.503	0.461		0.554	0.503	
Observations	1819	1819	1819	1819	1819	1819
Networks	116	116	116	116	116	116

Notes: Estimation has been performed using Matlab.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Strong ties and weak ties: Long-run effects

Dep.Var. Years of Education	Total IV	Lagged IV
Strong ties (ϕ^S)	0.0317***	0.0345***
	(0.0137)	(0.0155)
Weak ties (ϕ^W)	0.0062	0.0080
	(0.0055)	(0.0063)
Individual socio-demographic	yes	yes
Family Background	yes	yes
Protective Factors	yes	yes
Residential neighborhood	yes	yes
Contextual Effects	yes	yes
Network Fixed Effects	yes	yes
Observations	1819	1819
Networks	116	116

Notes: Estimation has been performed using Matlab. We report bias-corrected 2SLS estimates. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Weak ties in different grades (10 - 12)

Dep.Var. Years of Education	Total IV	Lagged IV
Strong ties (ϕ^S)	0.0419***	0.0485**
	(0.0187)	(0.0212)
Weak ties in lower grades (ϕ^{W1})	0.0015	0.0027
	(0.0237)	(0.0331)
Weak ties in higher grades (ϕ^{W2})	0.0011	0.0049
	(0.0207)	(0.0257)
Individual socio-demographic	yes	yes
Family Background	yes	yes
Protective Factors	yes	yes
Residential neighborhood	yes	yes
Contextual Effects	yes	yes
Network Fixed Effects	yes	yes
Observations	628	628
Networks	41	41

Notes: Estimation has been performed using Matlab. We report bias-corrected 2SLS estimates. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: 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
GPA	2.89	0.87	0	4.85	2.91	0.94	0	4.81	2.88	0.97	0	4.82
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 5: Link formation
Model (8) OLS estimation results

VARIABLE	γ 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 6: Directed networks

Dep.Var. Years of Education		
	Total IV	Lagged IV
Strong ties (ϕ^S)	0.0393***	0.0474***
	(0.0160)	(0.0183)
Weak ties (ϕ^W)	0.0049	0.0052
	(0.0063)	(0.0070)
Individual socio-demographic	yes	yes
Family Background	yes	yes
Protective Factors	yes	yes
Residential neighborhood	yes	yes
Contextual Effects	yes	yes
Network Fixed Effects	yes	yes
Observations	1819	1819
Networks	116	116

Notes: Estimation has been performed using Matlab. We report bias-corrected 2SLS estimates. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Strong ties and weak ties: Short-run effects

Dep.Var. GPA	
	Total IV
Strong ties (ϕ^S)	0.0238***
	(0.0097)
Weak ties (ϕ^W)	0.0079*
	(0.0046)
Individual socio-demographic	yes
Family Background	yes
Protective Factors	yes
Residential neighborhood	yes
Contextual Effects	yes
Network Fixed Effects	yes
Observations	1819
Networks	116

Notes: Estimation has been performed using Matlab. We report bias-corrected 2SLS estimates. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.