Peer Effects in Education, Sport, and Screen Activities: Local Aggregate or Local Average?*

Xiaodong Liu† Eleonora Patacchini‡ Yves Zenou§

October 15, 2011

Abstract

We develop two different social network models with different economic foundations. In the local-aggregate model, it is the sum of friends' efforts in some activity that affects the utility of each individual while, in the local-average model, it is costly to deviate from the average effort of friends. Even though the two models are fundamentally different in terms of behavioral foundation, their implications in terms of Nash equilibrium are relatively close since only the adjacency (social interaction) matrix differs in equilibrium, one being the row-normalized version of the other. We test these alternative mechanisms of social interactions to study peer effects in education, sport and screen activities for adolescents in the United States using the AddHealth data. We extend Kelejian's (2008) J test for spatial econometric models helping differentiate between these two behavioral models. We find that peer effects are not significant for screen activities (like e.g. video games). On the contrary, for sport activities, we find that students are mostly influenced by the aggregate activity of their friends (local-aggregate model) while, for education, we show that both the aggregate performance at school of friends and conformism matter, even though the magnitude of the effect is higher for the latter.

Key words: Social networks, conformism, peer effects, econometrics of networks.

JEL Classification: A14, D85, Z13

*We thank Lung-fei Lee for very helpful comments and discussions. We are also grateful to Edoardo Rainone for excellent research assistance.

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

‡Corresponding author. La Sapienza University of Rome, EIEF and CEPR. E-mail: eleonora.patacchini@uniroma1.it

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

1 Introduction

In many circumstances, the decision of agents to exert effort in education or some other activity cannot adequately be explained by their characteristics and by the intrinsic utility derived from it. Rather, its rationale may be found in how peers and others value this activity. There is indeed strong evidence that the behavior of individual agents is affected by that of their peers. This is particularly true in education, crime, labor markets, fertility, participation to welfare programs, etc. (for surveys, see, Glaeser and Scheinkman, 2001; Moffitt, 2001; Durlauf, 2004; Ioannides and Loury, 2004; Ioannides, 2011). The way peer effects operate is, however, unclear. Are students working hard at school because their friends work hard or because they do not want to be different from the majority of their peers?

The aim of this paper is to help our understanding of social interactions mechanisms by studying peer effects in education, sport and screen activities for adolescents in the United States.

For that, we develop two *social network models* aiming at capturing the different ways peer effects operate.[1] In the *local-aggregate model*, peer effects are captured by the *sum of friends' efforts* in some activity so that the higher is the number of active friends an individual has, the higher is her marginal utility of exerting effort. In the *local-average model*, peer effects are viewed as a social norm and individuals pay a cost from deviating from this norm. In this model, each individual wants to *conform* as much as possible to the social norm of her reference group, which is defined as the average efforts of her friends. Conformism is the idea that the easiest and hence best life is attained by doing one's very best to blend in with one's surroundings, and to do nothing eccentric or out of the ordinary in any way.[2] It may well be best expressed in the old saying, "When in Rome, do as the Romans do".

We characterize the Nash equilibrium of each model and show under which condition an interior Nash equilibrium exists and is unique. Even though the two models are fundamentally different in terms of behavioral foundation, it turns out that their implications in terms

[1] There is a growing literature on networks in economics. See the recent literature surveys by Goyal (2007) and Jackson (2008).

[2] In economics, different aspects of conformism and social norms have been explored from a theoretical point of view. To name a few, (*i*) peer pressures and partnerships (Kandel and Lazear, 1992) where peer pressure arises when individuals deviate from a well-established group norm, e.g., individuals are penalized for working less than the group norm, (*ii*) religion (Iannaccone 1992, Berman 2000) since praying is much more satisfying the more average participants there are, (*iii*) social status and social distance (Akerlof 1980, 1997; Bernheim 1994; Battu et al., 2007, among others) where deviations from the social norm (average action) imply a loss of reputation and status.

of Nash equilibrium are relatively close since only the adjacency matrix (which keeps track on whom each individual is friend with) differs in equilibrium, one being the row-normalized version of the other.

We then test these two models using the U.S. National Longitudinal Survey of Adolescent Health (AddHealth), which contains unique detailed information on friendship relationships among teenagers. Empirical tests of models of social interactions are quite problematic because of well-known issues that render the identification and measurement of peer effects quite difficult: (*i*) reflection, which is a particular case of simultaneity (Manski, 1993) and (*ii*) endogeneity, which may arise for both peer self-selection and unobserved common (group) correlated effects. Our econometric strategy utilizes the structure of the network as well as network fixed effects and high quality individual information to clearly identify the peer effects from the contextual affects and from the correlated effects. This approach for the identification of peer effects, i.e. the use of network fixed effects in combination with high quality data on social contacts has been used in a number of recent studies based on the AddHealth data (e.g. Liu et al. 2011; Lin 2010; Patacchini and Zenou, 2012).

We extend Kelejian's (2008) J test for spatial econometric models to differentiate between the local-aggregate and the local-average effects in a social-interaction model with network fixed-effects. We propose an hybrid model encompassing both local-aggregate and local-average effects and develop appropriate IV-based estimators. The traditional 2SLS estimator does not work well in our empirical analysis as the first-stage F test suggests the available IVs are weak. To fix the weak IV problem, we follow Lee (2007a) by generalizing the 2SLS estimator to a GMM estimator with additional quadratic moment conditions based on the correlation structure of the error term in the reduced-form equation. The GMM approach is easy to implement, asymptotic efficient, and allows us to test the hypothesis that the network formation (i.e. the adjacency matrix) is exogenous (conditional on covariates and network fixed effects) using an over-identifying restriction test.

We find that peer effects are not significant for screen activities (like e.g. video games). On the contrary, for sport activities, we find that students are mostly influenced by the aggregate activity of their friends (local-aggregate model) while, for education, we show that both the aggregate activity of their friends and the deviation from the social norms matter, even though the magnitude of the effect is higher for the latter. This indicates that students tend to *conform* to the social norm of their friends in terms of grades.

There are several papers that have formalized the local-aggregate model using a network approach (see, in particular, Ballester et al., 2006, 2010; Bramoullé and Kranton, 2007; Galeotti et al., 2009) and have tested this model for education (Calvó-Armengol et al.,

2009) and crime (Calvó-Armengol et al., 2005; Patacchini and Zenou, 2008; Liu et al., 2011). There are fewer papers that have explicitly modeled the local-average model (Glaeser and Scheinkman, 2003; Patacchini and Zenou, 2012) and have tested it for education (Lin, 2010; Boucher et al., 2010) and crime (Patacchini and Zenou, 2012). Finally, Ghiglino and Goyal (2010) develop a theoretical model where they compare the local aggregate and the local average model in the context of a pure exchange economy where individuals trade in markets and are influenced by their neighbors. They found that with *aggregate* comparisons, networks matter even if all people have same wealth. With *average* comparisons networks are irrelevant when individuals have the same wealth. The two models are, however, similar if there is heterogeneity in wealth.[3]

To the best of our knowledge, this is the first paper that provides two different social network models of peer effects (the local-aggregate versus the local-average model) and tests which one prevails in different activities. This is an important issue since it highlights a possibly important difference between "quantity" and "quality" of peers.

Let us explain in more detail the differences between the two models and what kinds of mechanisms they imply. For that, we will start with the study of Whyte (1955) on the Italian North End of Boston in the late 1930's. Whyte studied the behavior of a streetcorner gang, especially that of their leader Doc. Whyte wondered why Doc, a highly intelligent and curious individual, was not upwardly mobile and, instead, dropped out from school. Whyte was puzzled by Doc's behavior because school would have been easy for Doc given his exceptional ability and intelligence. Whyte concluded that Doc did not seek extra education out of loyalty to his group, whom he would be abandoning where he to advance beyond them educationally. The behavior of Doc is in accordance with the local-average model where it is costly to deviate from the group's social norm. Even if Doc is much more intelligent than the members of his gang, it would be too costly for him to acquire a higher level of education since this would mean interacting less with his friends or even abandoning them. Contrary to a model with no social interactions, where educational costs are mainly tuition fees, lost wages, etc., here it is the cost of lost contacts with own friends that is crucial. Now, if Doc had preferences according to the local-aggregate model, he would have acted differently. His decision to seek extra education would have been driven by his formidable ability and the sum of his friends' educational level, which is going to be quite high as Doc, a leader, has many friends. What is crucial, however, is that there would not be a cost from deviating from his friends' decisions and he would certainly have decided to pursue education, despite

[3] An other interesting paper is that of Clark and Oswald (1998) who propose a choice-theoretical justification for the local-average (i.e. conformist) model.

the lower average education level of his peers. Observe that, in the local-average model, the cost of deviating from the social norm depends on the distance from the average behavior of the group the individual belongs to. This is important from a policy viewpoint. Indeed, if only one person deviates from a large group of friends, then the cost could be very high. Anson (1985) told the story of an African-American youth, Eddie Perry, who left his black poor neighborhood to enter in a very prestigious prep school but at the cost of considerable psychological pain, because he did not fit naturally into either his old world of the inner city or his new world of the prep school. In particular, when Eddie went back to his old neighborhood, his friends ridiculed him because he couldn't play basket properly. As a mentor of Eddie put it: "They ridiculed him for going away to school, they ridiculed him for turning white".

This means that the policy implications of the two models are quite different. In the local-average model, the only way to affect individuals' behavior and thus their outcomes is to change the social norm of the group. In other words, one needs to affect most persons of the group for the policy to be effective. On the other hand, in the local-aggregate model, one can target only one individual and still have positive effects because she, in turn, will affect her peers. In other words, in the local-aggregate model there is a *learning process* from peers that is not true in the local-average model. Akerlof (1997) discussed the Eugene Lang's famous offer to give a college scholarship to *every student* at the sixth grade class in Harlem. Of the 51 students who remained in the New York area, 40 were considered likely to go to college six years later. Akerlof (1997) explained the success of this policy by the fact that it affected *all* students not some of them. As Akerlof put it: "The experiment was successful because the students formed a cohesive group in which each member received reinforcement from others who, like themselves, were on the academic track toward graduation from high school". In the language of the local-average model, this policy worked well because it changes the norm's group by affecting all its members. After the policy experiment, graduating and going to college was not anymore considered as "bad" or "acting white" but as the social norm of the group, i.e. what should be done. In the context of the local-aggregate model, one does not need to perform such a costly policy. It suffices to give a college scholarship to some students who, by increasing their performance, will increase the total effort of the peers' reference group of their friends, who will, in turn, affect their the total effort of their own friends, etc. This implies that, in a conformist group, changing people's behavior is much more difficult than in a "local-aggregate" group.

The paper is organized as follows. In the next section, we expose the two theoretical models, characterize the Nash equilibrium and show under which condition it exists and is unique. In Section 3, we describe our data while Section 4 exposes all the econometric issues we are facing and explains how we solve them. In Section 5, we discuss our results, both from a statistical and an economic viewpoint. We perform a robustness check of our results for undirected networks in Section 6. Finally, Section 7 concludes.

2 Theoretical framework

2.1 The model with local aggregates

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

The network $N_r = \{1, \ldots, n_r\}$ is a finite set of agents in network g_r $(r = 1, ..., \overline{r})$, where $\overline{r}$ is the total number of networks. We keep track of social connections by a network $G_r = [g_{ij,r}]$, where $g_{ij,r} = 1$ if i and j are direct friends, and $g_{ij,r} = 0$, otherwise. Friendship are reciprocal so that $g_{ij,r} = g_{ji,r}$. All our results hold for non-symmetric (*directed*) networks but, for the ease of the presentation, we focus on symmetric (*undirected*) networks in the theoretical model. We also set $g_{ii,r} = 0$.

Preferences Individuals in network g_r decide how much effort to exert in some activity (education, sport, and screening activities in the empirical analysis). We denote by $y_{i,r}$ the effort level of individual i in network g_r and by $Y_r = (y_{1,r}, ..., y_{n_r,r})'$ the population effort profile in network g_r. Each agent i selects an effort $y_{i,r} \geq 0$, and obtains a payoff $u_{i,r}(Y_r, g_r)$ that depends on the effort profile Y_r and on the underlying network g_r, in the following way:

$$u_{i,r}(Y_r, g_r) = (a_{i,r} + \eta_r + \varepsilon_{i,r})\ y_{i,r} - \frac{1}{2} y_{i,r}^2 + \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{i,r} y_{j,r} \tag{1}$$

where $\phi_1 > 0$. This utility has two parts: own and peer characteristics, $(a_{i,r} + \eta_r + \varepsilon_{i,r})\ y_{i,r} - \frac{1}{2} y_{i,r}^2$, and peer effects, $\phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{i,r} y_{j,r}$. In the first part of (1), η_r denotes the unobservable network characteristics and $\varepsilon_{i,r}$ is an error term, meaning that there is some uncertainty in the benefit part of the utility function. Both η_r and $\varepsilon_{i,r}$ are observed by the individuals but not by the econometrician. There is also an ex ante *idiosyncratic heterogeneity*, $a_{i,r}$, 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, age, 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, $a_{i,r}$ can be written as:

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

where $g_{i,r} = \sum_{j=1}^{n_r} g_{ij,r}$ is the number of direct links of individual i, x_i^m is a set of M variables accounting for observable differences in individual characteristics of individual i, and β_m, γ_m are parameters. The benefits from the utility are given by $(a_{i,r} + \eta_r + \varepsilon_{i,r})\ y_{i,r}$ while the cost is $\frac{1}{2}y_{i,r}^2$; both are increasing in own effort $y_{i,r}$. The second part of the utility function is: $\phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{i,r} y_{j,r}$, which reflects the influence of friends' behavior on own action. The peer effect component can also be heterogeneous, and this *endogenous heterogeneity* reflects the different locations of individuals in the friendship network g_r and the resulting effort levels. More precisely, bilateral influences are captured by the following cross derivatives, for $i \neq j$:

$$\frac{\partial^2 u_{i,r}(Y_r, g_r)}{\partial y_{i,r} \partial y_{j,r}} = \phi_1 g_{ij,r} \geq 0. \tag{3}$$

When i and j are direct friends, the cross derivative is $\phi_1 > 0$ and reflects strategic complementarity in efforts. When i and j are not direct friends, this cross derivative is zero. In particular, $\phi_1 > 0$ means that if two students are friends, i.e. $g_{ij,r} = 1$, and if j increases her effort, then i will experience an increase in her (marginal) utility if she also increases her effort. Interestingly, utility increases with the *number* of friends each person has, weighted by efforts $y_{j,r}$. This is the *local aggregate model* since more active friends implies higher utility.

To summarize, when individual i exerts some effort in some activity, the benefits of the activity depends on own ability $a_{i,r}$, some network characteristics η_r and on some random element $\varepsilon_{i,r}$, which is specific to individual i. In other words, $a_{i,r}$ is the observable part (by the econometrician) of i's characteristics while $\varepsilon_{i,r}$ captures the unobservable characteristics of individual i. Note that the utility (1) is concave in own decisions, and displays decreasing marginal returns in own effort levels.

Nash equilibrium

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

$$\frac{u_{i,r}(Y_r, g_r)}{\partial y_{i,r}} = a_{i,r} + \eta_r + \varepsilon_{i,r} - y_{i,r} + \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} = 0.$$

We obtain the following best-reply function for each $i = 1, ..., n_r$:

$$y_{i,r} = \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + a_{i,r} + \eta_r + \varepsilon_{i,r} \tag{4}$$

Denote by $\omega_1(G_r)$ the spectral radius of G_r. We have:[4]

Proposition 1 *If $\phi_1 \omega_1(G_r) < 1$, the peer effect game with payoffs (1) has a unique interior Nash equilibrium in pure strategies given by (4).*

2.2 The model with local averages

So far, we have seen that the sum of active direct friends had an impact on own utility. This was referred to as the *local aggregate* model. Let us now develop the *local average* model where the *average effort level* of direct friends affects utility. For that, let us denote the set of individual i's best friends (direct connections) as:

$$N_{i,r}(g_r) = \{j \neq i \mid g_{ij,r} = 1\}$$

which is of size $g_{i,r}$ (i.e. $g_{i,r} = \sum_{j=1}^{n_r} g_{ij,r}$ is the number of direct links of individual i). This means in particular that, if i and j are best friends, then in general $N_{i,r}(g_r) \neq N_{j,r}(g_r)$ unless the network is complete. This also implies that groups of friends may overlap if individuals have common best friends. To summarize, the *reference group* of each individual i in network r is $N_{i,r}(g_r)$, i.e. the set of her best friends, which does not include herself.

Let $g^*_{ij,r} = g_{ij,r}/g_{i,r}$, for $i \neq j$, and set $g^*_{ii,r} = 0$. By construction, $0 \leq g^*_{ij,r} \leq 1$. Note that g^*_r is a row-normalization of the initial friendship network g_r, as illustrated in the following example, where G_r and G^*_r are the adjacency matrices of, respectively, g_r and g^*_r.

Example 1 *Consider the following friendship network g_r:*

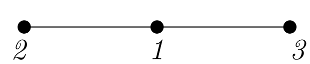

Figure 1

Then,

$$G_r = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad G^*_r = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

[4] All proofs of the propositions in the theoretical model can be found in Appendix A.

Preferences As above, $y_{i,r}$ denotes the effort level of individual i in network r. Denote by $\overline{y}_{i,r}$ the average effort of individual i's best friends. It is given by:

$$\overline{y}_{i,r} = \frac{1}{g_{i,r}} \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} \tag{5}$$

Each individual i selects an effort $y_{i,r} \geq 0$ and obtains a payoff given by the following utility function:

$$u_{i,r}(Y_r, g_r) = (a_{i,r} + \eta_r + \varepsilon_{i,r})\ y_{i,r} - \frac{1}{2} y_{i,r}^2 - \frac{d}{2} \left(y_{i,r} - \overline{y}_{i,r}\right)^2 \tag{6}$$

with $d > 0$. All the parameters have the same interpretation as in (1). Let us now interpret the peer-effect part of this utility function since it is the only aspect that differ from (1). Indeed, the last term $d\left(y_{i,r} - \overline{y}_{i,r}\right)^2$ reflects the influence of friends' behavior on own action. It is such that each individual wants to minimize the *social distance* between herself and her reference group, where d is the parameter describing the *taste for conformity*. Here, the individual loses utility $d\left(y_{i,r} - \overline{y}_{i,r}\right)^2$ from failing to conform to others. This is the standard way economists have been modelling conformity (see, among others, Akerlof, 1980, Bernheim, 1994, Kandel and Lazear, 1992, Akerlof, 1997, Fershtman and Weiss, 1998; Patacchini and Zenou, 2012).

Observe that beyond the idiosyncratic heterogeneity, $a_{i,r}$, there is a second type of heterogeneity, referred to as *peer heterogeneity*, which captures the differences between individuals due to network effects. Here it means that individuals have different types of friends and thus different reference groups $\overline{y}_{i,r}$. As a result, the social norm each individual i faces is endogenous and depends on her location in the network as well as the structure of the network. Indeed, in a star-shaped network (as the one described in Figure 1) where each individual is at most distance 2 from each other, the value of the social norm will be very different than a circle network, where the distance between individuals can be very large.

Nash equilibrium

We now characterize the Nash equilibrium of the game where agents choose their effort level $y_{i,r} \geq 0$ simultaneously. The corresponding first-order conditions are:

$$\frac{u_{i,r}(Y_r, g_r)}{\partial y_{i,r}} = a_{i,r} + \eta_r + \varepsilon_{i,r} - y_{i,r} - d(y_{i,r} - \overline{y}_{i,r}) = 0$$

Therefore, using (5), we obtain the following best-reply function for each $i = 1, ..., n_r$:

$$y_{i,r} = \phi_2 \frac{1}{g_{i,r}} \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + \frac{a_{i,r} + \eta_r + \varepsilon_{i,r}}{(1 - \phi_2)} \tag{7}$$

where $\phi_2 = d/(1 + d)$. We have:

Proposition 2 *Assume $\phi_2 < 1$. Then, the peer effect game with payoffs (6) has a unique interior Nash equilibrium in pure strategies given by (7).*

2.3 Local aggregate or local average? Theoretical considerations

In the local aggregate model, it is the *sum of the efforts of her peers* that affects the utility of individual i. So the more individual i has active (i.e. providing effort) friends, the higher is her utility. On the contrary, in the *local-average* model, it is the deviation from the *average of efforts of her peers* that affects the utility of individual i. So the closer is i's effort from the average of her friends' efforts, the higher is her utility.

Consequently, the two models are quite different from an economic viewpoint, even though, from a pure technical point of view, they are not that different (compare the best-reply functions (4) and (7)). In particular, the adjacency matrix G_r of direct links of the network totally characterizes the peer effects in the *local aggregate* model whereas it is a transformation of this matrix G_r to a weighted stochastic matrix G_r^* that characterizes the peer effects in the *local-average* model. This means that, in equilibrium, in the former model, individuals are positively affected by the sum of their friends' effort (non row-normalized G_r) while, in the latter, they are positively affected by the average effort of their friends (row-normalized G_r). Observe, however, that in the local-aggregate model, a condition ($\phi_1 \omega_1(G_r) < 1$) is needed for the Nash equilibrium to exist while $\phi_2 < 1$ is required in the local-average model, which is a much weaker condition. Indeed, the condition $\phi_1 \omega_1(G_r) < 1$ stipulates that local complementarities must be small enough compared to own concavity, which prevents multiple equilibria to emerge and, in the same time, rules out corner solutions (i.e., negative or zero solutions). When $\phi_1 \omega_1(G_r) > 1$, an equilibrium fails to exist because the positive feedback from other agents' contributions is too high and contributions increase without bound. This is not true in the local-average model since an increase in the effort of friends also increases the individual marginal utility of exerting own effort but it is the average effort that matters.

From an economic viewpoint, in the *local-aggregate* model, even if individuals were ex ante identical (in terms of $a_{i,r}$ and $\varepsilon_{i,r}$), different positions in the network would imply different effort levels, because it is the sum of efforts that matter. This would not be true in the *local-average* model since, in that case, the position in the network would not matter since it is the deviation from the average effort of friends that affects the utility.

Take for example the star-shaped network with 3 ex ante individuals in Figure 1. In the local *aggregate* model, individual 1 will exert the highest effort since she has two direct friends and will thus receive high local complementarities, given by $y_{2,r} + y_{3,r}$, whereas the two other individuals have only one friend and each will only receive $y_{1,r}$. In the *local-average* model, this is not anymore true since the peer effect component of individual 1 is $-\left[y_{1,r}-\left(y_{2,r}+y_{3,r}\right)/2\right]^2$ whereas, for individuals 2 and 3, we have: $-\left(y_{2,r}-y_{1,r}\right)^2$ and $-\left(y_{3,r}-y_{1,r}\right)^2$, respectively. The differences in the direct links are already small and, in equilibrium, where both direct and indirect links are taken into account, these peer-effect aspects turn out to be the same for all individuals in the network. Indeed, denote $\alpha_{i,r}\equiv a_{i,r}+\eta_r+\varepsilon_{i,r}$ and assume that all three individuals are ex ante identical so that $\alpha_{1,r}=\alpha_{2,r}=\alpha_{3,r}=1$ and that $\phi_1=\phi_2=\phi$. Then, it is easily verified that, for the *local-aggregate* model, if $\phi<1/\sqrt{2}$, then $y_{1,r}^*=(1+2\phi)/\left(1-2\phi^2\right)$ and $y_{2,r}^*=y_{3,r}^*=(1+\phi)/\left(1-2\phi^2\right)$, with $y_{1,r}^*>y_{2,r}^*=y_{3,r}^*$. On the other hand, for the *local-average* model, if $\phi<1$, then $y_{1,r}^*=y_{2,r}^*=y_{3,r}^*=1/\left(1-\phi\right)^2$.

In our empirical analysis we would like to determine which model matters the most for different activities. This will help understand if students are conformist or not. For that, we will estimate the best-reply functions (4) and (7) using an appropriate econometric strategy.

3 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).[5]

The AddHealth database has been designed to study the impact of the social environment (i.e. friends, family, neighborhood and school) on adolescents' behavior in the United States by collecting data on students in grades 7-12 from a nationally representative sample of roughly 130 private and public schools in years 1994-95. Every pupil attending the sampled schools on the interview day is asked to compile a 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

[5]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.

and household information (in-home and parental data).[6] Those subjects of the subset are interviewed again in 1995–96 (wave II), in 2001–2 (wave III), and again in 2007-2008 (wave IV).[7] For the purpose of our analysis, we focus on wave I because the network information is only available in the first wave.

From a network perspective, the most interesting aspect of the AddHealth data is the information on friendships. Indeed, the friendship information is based upon actual friends nominations. Pupils were asked to identify their best friends from a school roster (up to five males and five females).[8] Knowing exactly who nominates whom in a network, we exploit the directed nature of the nominations data.[9] We focus on choices made and we denote a link from i to j as $g_{ij,r} = 1$ if i has nominated j as her friend in network r, and $g_{ij,r} = 0$, otherwise.[10] By matching the identification numbers of the friendship nominations to respondents' identification numbers, one can obtain information on the characteristics of nominated friends. More importantly, one can reconstruct the whole geometric structure of the friendship networks. For each school, we thus obtain all the networks of (best) friends.[11]

We exploit this unique data set to understand the impact of peer pressure on individual behavior for three different outcomes: (*i*) school performance; (*ii*) sport activities, such as playing baseball, softball, basketball, soccer, or football; (*iii*) screen activities, such as playing video or computer games.

For each individual, we calculate an index of performance (or involvement) in each category using the answers to different related questions.[12] More specifically, the *school performance* is measured using the respondent's scores received in the more recent grading period in several subjects, namely English or language arts, history or social science, mathematics,

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

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

[8] 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.

[9] We also exploit the undirected nature of the friendship data in Section 6.

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

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

[12] This is a standard factor analysis, where the factor loadings of the different variables are used to derive the total score.

and science. The scores are coded as 1=D or lower, 2=C, 3=B, 4=A. The final composite score (labeled as *GPA index* or grade point average index) ranges between 0 and 4.40, with mean equals to 3.03 and standard deviation equals to 1.10.

The involvement in *sport activities* is derived using questions on how often students go wheeling, play a team sport or do more general physical exercise during the past week. Each response is coded using an ordinal scale ranging from 0 (i.e. never participate) to 1 (i.e. participate 1 or 2 times), 2 (participate 3 or 4 times) up to 3 (i.e. participate 5 or more times). The final index of involvement in sport activities (labeled as *sport participation index*) ranges between 0 and 2.95, with mean equals to 1.53 and standard deviation equals to 1.05.

Similarly, the involvement in *screen activities* (labeled as *screen participation index*) is derived using questions on how many hours a week students watch television, videos or play video or computer games. The answers take values between 0 and 99 hours for each activity. The final composite ranges between 0 and 13.01, with mean equals to 1.37 and standard deviation equals to 1.23.[13] Precise definitions of the remaining variables (control variables) used in our empirical analysis, as well as a summary description of our final sample of students distinguishing between the different outcomes, can be found in Table C.1 in Appendix C. This table shows that, on average, the characteristics of the students are remarkably similar. The large reduction in sample size with respect to the original AddHealth sample is mainly due to missing values in variables and to the network construction procedure. Indeed, roughly 20% of the students do not nominate any friends and another 20% cannot be correctly linked (for example because the identification code is missing or misreported). Also, we do not consider networks at the extremes of the network size distribution to avoid the possibility that, in these extreme networks, the strength of peer effects can have extreme values (too low or too high) that may be a matter of concern. Indeed, our theoretical model and hence our empirical strategy consider homogenous peer effects across networks. The use of network fixed effects, which is an important feature of our identification strategy (Section 4.1) prevents us to deal with this issue. We focus our analysis on networks with network size between 50 and 150 students for all outcomes.[14]

[13] Endogenous selection into GPA, sport and screen activititities in not an issue here. Less than 1% and 5% of the students never participated in screen and sport activities, respectively. A final assessment of the comprehension within the different subjects studied at school (GPA data) is mandatory.

[14] Our results, however, do not depend crucially on these network size thresholds. They remain qualitatively unchanged when slightly moving the network size window.

4 Econometric issues

The test of our theoretical framework requires an appropriate estimate of peer effects for the local-aggregate model ($\widehat{\phi}_1$) and for the local-average model ($\widehat{\phi}_2$). Our econometric strategy is organized as follows. In Section 4.1, we expose our identification strategy. Section 4.2 presents a J test for model selection aiming at showing which behavioral mechanism prevails. In Section 4.3, we present the hybrid model, which combines both the local-aggregate and the local-average model. Finally, in Section 4.4, we detail the appropriate 2SLS and GMM estimators designed to tackle the issue of simultaneity in spatial/network econometric models.

4.1 Identification strategy

The identification of peer effects raises different challenges that are common to the local-aggregate and the local-average model. As stated above, from a behavioral point of view, in the *local-aggregate* model, it is the *sum of the efforts of the peers* that affects the individual utility whereas in the *local-average* model, it is the deviation from the *average efforts of the peers* that affects the individual utility. From a pure technical point of view, this behavioral distinction implies to deal with a non row-standardized matrix of social contacts if local aggregates matters (i.e. G_r) and with a row-standardized social interaction matrix (i.e. G_r^*) if instead local averages matter. Although the information contained in the variation of row sums of the social interaction matrix may help model identification and estimation efficiency (Liu and Lee, 2010), the basic empirical issues that arise when one seeks to *separately* identify peer or endogenous effects from contextual or exogenous effects are common to both models. We begin this section by discussing each of them in turn.

4.1.1 The reflection problem

In *linear-in-means* models, simultaneity in behavior of interacting agents introduces a perfect collinearity between the expected mean outcome of the group and its mean characteristics. Therefore, it is difficult to differentiate between the effect of peers' choice of effort and peers' characteristics that do impact on their effort choice. This is the so-called *reflection problem*, first formulated by Manski (1993). Basically, the reflection problem arises because, in the standard approach, individuals interact in groups, that is individuals are affected by all individuals belonging to the same group and by nobody outside the group. In other words, groups completely overlap. In the case of social networks, instead, this is nearly never true since the reference group has individual-level variation. In other words, if i and j are friends and j and k are friends, it does not necessarily imply that i and k are also friends. This corresponds to the case of a non-complete network, i.e. everybody is not connected with everybody else. When this occurs, it creates the ability to identify network effects and to solve the reflection problem. Indeed, the characteristics of k affects the decisions of i only indirectly through their effects on the decisions of j. As a result they act as a valid instrument to determine if i is influenced by the decisions of j, i.e. for the identification of peer effects. To identify peer effects, one needs only one such intransitivity. In most real-world social network, there is a very large number of non-overlapping peer groups and the Manski's (1993) reflection problem is thus eluded. Peer effects in social networks are thus identified and can be estimated using 2SLS or maximum likelihood (Lee 2007; Calvó-Armengol et al., 2009; Lin, 2010).[15] The conditions on the parameters that guarantee identification of peer effects are formally derived in Bramoullé et al. (2009) (Proposition 3) and Calvó-Armengol et al. (2009) (Proposition 2) for the case of a row-standardized interaction matrix. The results for the case of a non row-standardized interaction matrix are instead contained in Liu et al. (2011).

4.1.2 Sorting into groups: Endogeneity of network formation

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

[15] A more technical exposition of these results can be found in Liu and Lee (2010).

from the individual-level variables.[16] As detailed in the next section, this paper follows this approach.

One might also question the presence of problematic unobservable factors that are not network-specific, but rather individual-specific. In this respect, the richness of the information provided by the AddHealth questionnaire on adolescents' behavior allow us to find proxies for typically unobserved individual characteristics that may be correlated with our variable of interest. Specifically, to control for differences in leadership propensity across adolescents, we include an indicator of *self-esteem,* and we use *mathematics score* as an indicator of ability. Also, we attempt to capture differences in parenting and more general social influences by including parental care and indicators of the student's school attachment and relationship with teachers.

Our identification strategy is based on the assumption that any troubling source of heterogeneity (if any), which is left unexplained by our unusually large set of observed (individual and peers) characteristics can be captured at the network level, and thus taken into account by the inclusion of network fixed effects. In other words, our particularly large information on individual and peer variables, which also includes behavioral characteristics, *should reasonably explain the process of selection into groups whereas network fixed effects might capture any remaining source of selection on unobservables.*

To be more precise, we allow link formation (as captured by our matrix G_r) to be correlated with observed individual characteristics, contextual effects and unobserved network characteristics (captured by the network fixed effects). We will test the hypothesis that G_r is exogenous (conditional on covariates and network fixed effects) using an over-identifying restrictions (OIR) test, as described in Lee (2007a). The moment conditions used in the GMM estimation of a spatial autoregressive model (such as the ones used in our empirical investigation) are based on the assumption that G_r is exogenous. If the OIR test cannot reject the null hypothesis that the moment conditions (or restrictions) are correctly specified, then it provides evidence that G_r can be considered as exogenous. If the number of restrictions is small relative to the sample size, the OIR test statistic given by the GMM objective function (23) evaluated at the GMM estimator follows a chi-squared distribution with degrees of freedom equal to the number of over-identifying restrictions (Lee, 2007a, Proposition 2).

[16] Bramoullé et al. (2009) also deal with this problem in the case of a row-normalized interaction matrix. In their Proposition 5, they provide formal conditions on network topology structure under which, by subtracting from the variables the network average, social effects are again identified. These conditions are typically satisfied in real-world social networks.

Given that G_r is exogenous (conditional on covariates and network fixed effects), the identification of the local aggregate effect ($\widehat{\phi}_1$) versus local average effect ($\widehat{\phi}_2$) hinges on the variation in row sums of G_r (see the following section). In the case that each student has the same number of friends as everyone else in the same group, we cannot separately identify those two effects.

Our econometric strategy thus utilizes the structure of the network as well as network fixed effects and high quality individual information to clearly identify the peer effects from the contextual affects and from the correlated effects. This approach for the identification of peer effects, i.e. the use of network fixed effects in combination with high quality data on social contacts has been used in a number of recent studies based on the AddHealth data (e.g. Patacchini and Zenou, 2012; Liu et al. 2011; Lin 2010).

4.1.3 Local aggregate or local average? Econometric considerations

We propose a test for model selection design to detect which behavioral mechanism better represents the data at hand.

In standard linear regression models, the J test is used to compare non-nested model specifications (Davidson and MacKinnon, 1981). The idea of the J test is as follows. If a given model contains the correct set of regressors, then including the fitted values of an alternative (or of a fixed number of competing models) into the set of regressors should provide no significant improvement. The J test statistic is simply the marginal test of the fitted values of the augmented model.

Kelejian (2008) extends the J test to a spatial framework. He shows that the test could, but need not, relate solely to the specification of the spatial weighting matrix. Importantly, since the J test relies on whether the prediction based on an alternative model significantly increases the explanatory power of the null model, it is important to use all the available information in the alternative model. However, Kelejian (2008) does not use the information in an efficient way to determine the predictions (Kelejian and Piras, 2011). The J test of Kelejian (2008) and Kelejian and Piras (2011) is implemented using the spatial 2SLS estimation procedure. Our contributions in the present paper can be summarized as follows.

(1) We generalize their J test to a network model with group fixed effect. The two spatial weight matrices that we consider are the row standardization of the other. Our source of identification is the variation in the row sums (i.e. peer group size).[17]

[17] Lee (2007b) shows that it is possible to identify the endogenous effect from contextual effects if there is variation in the network sizes in the sample.

(2) The augmented null model of the J test is similar to the model in Liu and Lee (2010), with, however, an important difference. While in Liu and Lee (2010) the endogeneity comes solely from the endogenous social interaction effect, in the augmented null model considered here, the endogeneity stems from both the endogenous effect and the predicted value. Besides the IVs proposed by Kelejian and Prucha (1998), we consider additional IVs based on the row sums of the weight matrix to improve the efficiency in the estimation of the augmented null model. At the same time, when we consider those additional IVs, we could introduce many-IV bias into the estimates if the number of IVs is large relative to our sample size. Hence, we propose a bias-correction procedure to eliminate the leading order many-IV bias.

(3) If the first stage, the F test suggests that the IVs are weak and thus the J test proposed by Kelejian (2008) based on the 2SLS estimator would not be reliable. We propose a GMM estimator to implement the J test. The GMM estimator uses additional quadratic moment conditions, which are especially helpful when the IVs are weak.

We present our generalization of the J test for network models with group fixed effects in the next section.

4.2 J test for model selection

4.2.1 Empirical models

Let $\bar{r}$ be the total number of networks in the sample, n_r be the number of individuals in the rth network, and $n = \sum_{r=1}^{\bar{r}} n_r$ be the total number of sample observations. Let us define the ex ante heterogeneity $a_{i,r}$ of each individual in network r as (see (2)):

$$a_{i,r} = x'_{i,r}\beta + \frac{1}{g_{i,r}} \sum_{j=1}^{n_r} g_{ij,r} x'_{j,r}\gamma.$$

The empirical model corresponding to (4), i.e., the *local-aggregate model*, can be written as:

$$y_{i,r} = \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + x'_{i,r}\beta_1 + \frac{1}{g_{i,r}} \sum_{j=1}^{n_r} g_{ij,r} x'_{j,r}\gamma_1 + \eta_{1r} + \epsilon_{1i,r}, \tag{8}$$

for $i = 1, \cdots, n_r$ and $r = 1, \cdots, \bar{r}$, where $x_{i,r} = (x^1_{i,r}, \cdots, x^m_{i,r})'$, $g_{i,r} = \sum_{j=1}^{n_r} g_{ij,r}$ and $\epsilon_{1i,r}$'s are i.i.d. innovations with zero mean and variance σ_1^2 for all i and r.

The empirical model corresponding to (7), i.e., the *local-average model*, is:

$$y_{i,r} = \phi_2 \frac{1}{g_{i,r}} \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + x'_{i,r}\beta_2 + \frac{1}{g_{i,r}} \sum_{j=1}^{n_r} g_{ij,r} x'_{j,r}\gamma_2 + \eta_{2r} + \epsilon_{2i,r}, \tag{9}$$

where $\eta_{2r} = \eta_r/(1-\phi_2)$ and $\epsilon_{2i,r}$'s are i.i.d. innovations with zero mean and variance σ_2^2 for all i and r.

Let $Y_r = (y_{1,r}, \cdots, y_{n_r,r})'$, $X_r = (x_{1,r}, \cdots, x_{n_r,r})'$, $\epsilon_{1r} = (\epsilon_{11,r}, \cdots, \epsilon_{1n_r,r})'$, $\epsilon_{2r} = (\epsilon_{21,r}, \cdots, \epsilon_{2n_r,r})'$. Denote the $n_r \times n_r$ sociomatrix (adjacency matrix) by $G_r = [g_{ij,r}]$, the row-normalized G_r by G_r^*, and an n_r-dimensional vector of ones by 1_{n_r}. Then, models (8) and (9) can be written in matrix form as:

$$\begin{aligned} H_1 &: Y_r = \phi_1 G_r Y_r + X_r^* \delta_1 + \eta_{1r} 1_{n_r} + \epsilon_{1r}, \\ H_2 &: Y_r = \phi_2 G_r^* Y_r + X_r^* \delta_2 + \eta_{2r} 1_{n_r} + \epsilon_{2r}, \end{aligned}$$

where $X_r^* = (X_r, G_r^* X_r)$, $\delta_1 = (\beta_1', \gamma_1')'$ and $\delta_2 = (\beta_2', \gamma_2')'$. For a sample with $\bar{r}$ groups, stack up the data by defining $Y = (Y_1', \cdots, Y_{\bar{r}}')'$, $X^* = (X_1^{*\prime}, \cdots, X_{\bar{r}}^{*\prime})'$, $G = \mathrm{D}(G_1, \cdots, G_{\bar{r}})$, $G^* = \mathrm{D}(G_1^*, \cdots, G_{\bar{r}}^*)$ and $\iota = \mathrm{D}(1_{n_1}, \cdots, 1_{n_{\bar{r}}})$, where $\mathrm{D}(A_1, \cdots, A_K)$ is a block diagonal matrix in which the diagonal blocks are $m_k \times n_k$ matrices A_k's. Furthermore, define $\epsilon_l = (\epsilon_{l1}', \cdots, \epsilon_{l\bar{r}}')'$ and $\eta_l = (\eta_{l1}, \cdots, \eta_{l\bar{r}})'$ for $l = 1, 2$. For the entire sample, the two models are, respectively,

$$H_1 : Y = \phi_1 GY + X^* \delta_1 + \iota \cdot \eta_1 + \epsilon_1, \quad (10)$$

$$H_2 : Y = \phi_2 G^* Y + X^* \delta_2 + \iota \cdot \eta_2 + \epsilon_2, \quad (11)$$

4.2.2 Augmented models

The test of model H_1 against model H_2 To test against the model specification H_2, one can estimate the following augmented model of H_1,

$$Y = \alpha_1 Y_{H_2} + \phi_1 GY + X^* \delta_1 + \iota \cdot \eta_1 + \epsilon_1, \quad (12)$$

where Y_{H_2} is a predictor of Y under H_2 such that $Y_{H_2} = \phi_2 G^* Y + X^* \delta_2 + \iota \cdot \eta_2$ (see Kelejian and Prucha, 2007; Kelejian and Piras, 2011). Thus, a test of the null model (10) against the alternative one (11) would be in terms of the hypotheses $H_0 : \alpha_1 = 0$ against $H_a : \alpha_1 \neq 0$.

Substitution of the predictor Y_{H_2} into (12) gives

$$\begin{aligned} Y &= \alpha_1(\phi_2 G^* Y + X^* \delta_2) + \phi_1 GY + X^* \delta_1 + \iota \cdot (\eta_1 + \alpha_1 \eta_2) + \epsilon_1 \quad (13) \\ &= Z_1^* \vartheta_1 + \iota \cdot (\eta_1 + \alpha_1 \eta_2) + \epsilon_1, \end{aligned}$$

where $Z_1^* = [(\phi_2 G^* Y + X^* \delta_2), GY, X^*]$ and $\vartheta_1 = (\alpha_1, \phi_1, \delta_1')'$.

In this model, we treat η_1 and η_2 as vectors of unknown parameters. When the number of groups $\bar{r}$ is large, we have the incidental parameter problem. Let $J = \mathrm{D}(J_1, \cdots, J_{\bar{r}})$, where $J_r = I_{n_r} - \frac{1}{n_r} 1_{n_r} 1_{n_r}'$. The within transformation of the second line of (13) gives

$$JY = JZ_1^* \vartheta_1 + J\epsilon_1. \quad (14)$$

The proposed J test can be implemented by the following two steps:

(1) Estimate model H_2 by the quasi-maximum-likelihood (QML) method of Lee et al. (2010). Let the preliminary QML estimators of ϕ_2 and δ_2 be denoted by $\tilde{\phi}_2$ and $\tilde{\delta}_2$.

(2) Estimate the feasible counterpart of model (14)

$$JY = J\tilde{Z}_1^* \vartheta_1 + J\epsilon_1, \tag{15}$$

where $\tilde{Z}_1^* = [(\tilde{\phi}_2 G^* Y + X^* \tilde{\delta}_2), GY, X^*]$, by the 2SLS or GMM method described in Section 4.4. If the estimated α_1 is insignificant, then this is evidence against model H_2.

The test of model H_2 against model H_1 The test of model H_2 against model H_1 can be carried out in a similar manner. Consider the following augmented model of H_2,

$$H_2 : Y = \alpha_2 Y_{H_1} + \phi_2 G^* Y + X^* \delta_2 + \iota \cdot \eta_2 + \epsilon_2, \tag{16}$$

where Y_{H_1} is a predictor of Y under H_1 such that $Y_{H_1} = \phi_1 GY + X^* \delta_1 + \iota \cdot \eta_1$. Thus, the test of the null model (11) against the alternative (10) would be in terms of the hypotheses $H_0 : \alpha_2 = 0$ against $H_a : \alpha_2 \neq 0$. The within transformation of (16) gives

$$JY = JZ_2^* \vartheta_2 + J\epsilon_2. \tag{17}$$

where $Z_2^* = [(\phi_1 GY + X^* \delta_1), G^* Y, X^*]$ and $\vartheta_2 = (\alpha_2, \phi_2, \delta_2)'$.

The proposed J test can be implemented by the following two steps:

(1) Estimate model H_1 by the 2SLS with IVs $J[X, G^* X, GX]$. Let the preliminary 2SLS estimators of ϕ_1 and δ_1 be denoted by $\tilde{\phi}_1$ and $\tilde{\delta}_1$.

(2) Estimate the feasible counterpart of model (17)

$$JY = J\tilde{Z}_2^* \vartheta_2 + J\epsilon_2, \tag{18}$$

where $\tilde{Z}_2^* = [(\tilde{\phi}_1 GY + X^* \tilde{\delta}_1), G^* Y, X^*]$, by the 2SLS or GMM method described in Section 4.4. If the estimated α_2 is significant, then that is evidence against model H_2.

4.3 The hybrid network model

As an alternative to the augmented models of the J test, we consider a hybrid model encompassing both *local-aggregate* and *local-average* effects,

$$y_{i,r} = \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + \phi_2 \frac{1}{g_{i,r}} \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + x_{i,r}' \beta + \frac{1}{g_{i,r}} \sum_{j=1}^{n_r} g_{ij,r} x_{j,r}' \gamma + \eta_r + \epsilon_{i,r}. \tag{19}$$

For the entire sample, model (19) can be written in matrix form

$$Y = \phi_1 GY + \phi_2 G^* Y + X^* \delta + \iota \cdot \eta + \epsilon, \tag{20}$$

where $\delta = (\beta', \gamma')'$. Let $J = \mathrm{D}(J_1, \cdots, J_{\bar{r}})$, where $J_r = I_{n_r} - \frac{1}{n_r} 1_{n_r} 1'_{n_r}$. The network fixed effect can be eliminated by the within transformation with J such that

$$JY = \phi_1 JGY + \phi_2 JG^* Y + JX^* \delta + J\epsilon. \tag{21}$$

The within model (21) can be estimated by the 2SLS and GMM methods described in the next subsection.

Under the restriction $\phi_1 = 0$, model (21) becomes

$$JY = \phi_2 JG^* Y + JX^* \delta + J\epsilon.$$

This restricted model only captures *local-average* effects, and it can be estimated by the QML method proposed by Lee et al. (2010).[18] On the other hand, under the restriction $\phi_2 = 0$, model (21) becomes

$$JY = \phi_1 JGY + JX^* \delta + J\epsilon,$$

which only captures *local-aggregate* effects. This model can be estimated by the 2SLS and GMM methods proposed by Liu and Lee (2010).

Observe that, in our case, where the only difference between the local-average and local-aggregate models is captured by a single parameter, the J test procedure has little advantage over the hybrid model (19) as the latter nests the two alternative network models. In other words, in our case where the only difference between the local aggregate and the local average model is the different adjacency matrices (non row-normalized and row-normalized, respectively), a testing procedure based on the augmented model or the hybrid model would be similar. The J test presented in Section 4.2 has, however, a more general validity for testing non-nested models (see Davidson and MacKinnon, 1993, pp. 381-388).

4.4 The 2SLS and GMM estimators

For the estimation of the hybrid network model (21) and the augmented model (15) or (18) in the second step of the J test, we consider the following estimators by generalizing the 2SLS and GMM methods in Liu and Lee (2010):

[18] For the QML method to be applicable, all rows of G^* have to be normalized to sum to unity. For the case of undirected networks, G^* satisfies this condition. However, for the case of directed networks, some rows of G^* are all zeros, and, hence, cannot be nomalized to sum to one. So for the estimation of directed networks, we use the 2SLS and GMM estimators given in (a)-(c) and (d$''$)-(f$''$) in Section 4.4.

(a) "2SLS-1": a 2SLS estimator with IVs $Q_1 = J[X, G^*X, GX]$.

(b) "2SLS-2": a 2SLS estimator with IVs $Q_2 = J[X, G^*X, GX, G\iota]$. This estimator has been proposed by Liu and Lee (2010). The additional IVs $G\iota$ corresponds to the information on different positions of group members measured by Bonacich (1987) centrality. The additional IVs improves asymptotic efficiency of the estimator and helps achieve identification when the "conventional" IVs Q_1 are weak.

(c) "C2SLS": a bias-corrected 2SLS estimator with IVs Q_2. Note that, the additional IVs in Q_2, $G\iota$, has $\bar{r}$ columns, where $\bar{r}$ is the number of networks in the data. Therefore, if there are many groups, the "2SLS-2" estimator may have an asymptotic bias, which is known as the many-instrument bias.[19] The "C2SLS" estimator adjusts the "2SLS-2" estimator by an estimated leading-order many-instrument bias.

The 2SLS estimators are based on moment conditions that are *linear* in the model coefficients. However, when the IVs are weak, the inference based on the 2SLS estimation may be unreliable. Lee (2007a) has suggested to generalize the 2SLS method to a comprehensive GMM framework with additional *quadratic* moment conditions based on the covariance structure of the reduced form equation to improve identification and estimation efficiency. The added quadratic moment conditions are especially helpful when the IVs are weak. In this paper, we consider the following GMM estimators for the estimation of the empirical model:

(d) "GMM-1": an optimal GMM estimator with linear moment conditions with Q_1 and quadratic moment conditions.

(e) "GMM-2": an optimal GMM estimator with linear moment conditions with Q_2 and the same quadratic moment conditions as in GMM-1.

(f) "CGMM": a bias-corrected optimal GMM estimator with the same moment functions as in GMM-2. Similar to the corresponding 2SLS estimator, the additional IVs in Q_2 may introduce many-instrument bias into the GMM estimator. The "CGMM" estimator adjusts "GMM-2" by an estimated leading-order bias.

The details of the 2SLS and GMM methods including the explicit form of the quadratic moment condition are given in Appendix B.

[19] This is less of a concern in the data used in this paper, as the number of groups are small relative to the sample size.

5 Estimation results

5.1 Statistical analysis

Let us now test our theoretical local-aggregate and local-average models for different outcomes using the J test described above. Tables 1, 2, 3 report the estimation results for the different outcomes, i.e. sports (Table 1), screen activities (Table 2) and GPA (Table 3) using directed networks. We present the complete list of estimation results for the hybrid network model (19) and the results for the target parameters for the augmented models (12) and (16) in the last rows. The chosen estimator here is the bias-corrected optimal GMM (CGMM) described in Section 4.4 above, which is reliable even if the IVs are weak and there are many small groups. We present our results with an increasing set of controls, which helps us validate our identification strategy. More specifically, our identification strategy is based on the assumption that conditional on observed individual and peers characteristics and network unobserved characteristics, peers choice is random (within networks). Under this assumption, it should not matter too much which further controls are included: the estimated effects of peers' outcome on individual outcome should remain roughly unchanged. If some controls do matter, it implies that these covariates are correlated with peers' outcome and also influence individual decisions. As a result, one could worry that there may be other unobservable variables that are similarly correlated with our peer-level target variable and our dependent variable. We thus start by including a set of individual and peers characteristics that should reasonably explain the sorting of children into peer groups (peers' choice), such as parental education, sex, grade, ethnicity, mathematics score and indicators of school attachment and relationship with teachers (specification (1) in Tables 1, 2, 3) and then gradually introduce other possibly relevant factors affecting peers' choice and individual outcome. In specification (2), we add neighborhood quality and indicators of the social structures of families, namely number of components and whether the parents are married or not. Next, we consider an indicator of parental care and our proxy of self-esteem (specification (3)), and, finally, in specification (4), we add parental occupation dummies. Tables 1, 2, 3 show that the estimates of peer effects remain roughly unchanged across columns, thus supporting our confidence on *the exogeneity of network structure* (conditional on controls and network fixed effects).

[*Insert Tables* 1 *to* 3 *here*]

In order to appreciate the differences between the different estimators proposed in Section 4.4, we report in Tables 4, 5, 6 the estimation results for the hybrid network model (with the most extensive sets of controls) using alternative estimators.[20] These tables also show in the lasts rows the first stage F tests for weak IVs and the OIR tests.

[*Insert Tables* 4 *to* 6 *here*]

The values of the F-statistics reveal that our instruments in the linear moment conditions are quite weak.[21,22] This is the reason why we prefer the GMM with additional quadratic moment conditions (see Section 4.4 and Appendix B) to the 2SLS estimates. The different types of GMM estimators deliver similar estimation results. Indeed, the number of additional moment conditions used in GMM-2 is equal to the number of the groups, which are not many in our case. The correction used in CGMM is greatly effective if we have many small groups, which is not our case. Therefore in our context where the number of groups is limited and the average group size is not small with respect to total sample size, the statistical performance of the three different GMM estimators is similar.

Because of the weak IV problem, we cannot trust the OIR test results based on the 2SLS. We consider the OIR test statistic (Lee, 2007a) for the estimator GMM-1 which is based on a relatively smaller set of instruments (Q1). In fact, such a test could be biased if there are a large number of IVs, as in the IV matrix Q2.[23] We find that the $p-$values of the over-identifying restrictions test is large for all outcomes, which means that G_r is exogenous (i.e. we cannot reject the null hypothesis that the moment conditions based on an exogenous G_r are valid). This evidence provides further confidence on *the exogeneity of network structure* (conditional on controls and network fixed effects).

5.2 Interpretation of results in economics terms

Do peer effects matter? Which model is more adequate for each activity? Looking at the first two rows of Table 1, 2, 3 (tests of ϕ_1 and ϕ_2 of the hybrid network model) and at the last row of Table 1, 2, 3 (tests of the augmented models where the null hypothesis is $\alpha_1 = 0$,

[20] The qualitative results remain unchanged when using the augmented models.

[21] Stock et al. (2002) have suggested that "evidently the first-stage F statistic must be large, typically exceeding 10, for 2SLS inference to be reliable."

[22] The construction and statistical properties of the first stage F-test for weak instruments can be found in by Stock et al. (2002). The partial F-test used in this analysis is further detailed in Stock and Yogo (2005). The matlab codes that implement the 2SLS and GMM estimators, J test, F test and OIR test remain available upon request.

[23] For the case with independent observations, Chao et al. (2010) have proposed an OIR test that is robust to many IVs. However, no robust OIR test with many IVs is available when observations are spatially correlated.

i.e., the local average model does not matter for model (12), and $\alpha_2 = 0$, i.e., the local aggregate model does not matter for model (16), we see clearly that, for sport activities, it is the *sum of the effort of the friends* (i.e. the local aggregate model) and not *their average effort* that matters for explaining own sport activity while, for education (i.e. GPA index), both matter. Observe, however, that, even if both matter, the magnitude of the effects is higher for the local-average model than for the local-aggregate one. Indeed, a one-standard deviation increase in the average activity of individual i's reference group translates roughly into a 0.29 increase in standard deviations of individual i's GPA score while it is only 0.10 for the sum of activity of friends. Finally, for screen activities, our results suggest that peer effects are not important in explaining own screen activity. The latter appears to be explained by own characteristics and contextual effects. For example, male, black and lower grade students are more likely to participate in screen activities than other students.

Our results are interesting and new. First, it is not that surprising that peer effects have no significant impact on screening activities. Think, for example, of video games. There is, obviously, a social aspect to it but there is also an addicted one. If teenagers are very much into video games, then addiction might prevail and it does not matter very much if their friends are also into it. Second, for sport activities, it does not seem that social norms play an important role so that being different from friends have a cost. In other words, a student who is not into sport may have some friends who are athletes (or "jocks"). However, individuals with many sporty friends will very likely be themselves sporty. Finally, for education, we find that both social norms and total activity of friends matter. In other words, if a student have on *average* friends who have good grades, then he/she will not deviate from this norm. However, he/she will be likely to have good grades even if the peers are heterogenous with respect to grades but there are some friends with very good grades. In terms of magnitude, it seems that social norms are important actors in education so that students tend to conform to these norms.

6 Robustness checks

Our identification and estimation strategy depend on the correct specification of network links. In particular, our identification strategy hinges upon nonlinearities in group membership, i.e. on the presence of intransitive triads. In this section, we test the robustness of our results with respect to misspecification of network topology. So far, we have measured peer groups as precisely as possibly by exploiting the direction of the nomination data. However, friendship relationships are reciprocal in nature, and even if a best friend of a given student does not nominate this student as her best friend, one may think that social interactions take place. Under this circumstance, there can be some "unobserved" network link that, if considered, would change the network topology and break some intransitivities in network links. Therefore, in this section, we repeat our analysis by considering *undirected networks*, i.e. we assume that a link exists between two friends if both students have named each other, that is $g_{ij,r} = g_{ji,r} = 1$.

Tables 7, 8, 9 have the same structure as Tables 1, 2, 3 and report the results for undirected networks. The qualitative results remain unchanged. The magnitude of the effects is only slightly lower.

[*Insert Tables* 7 *to* 9 *here*]

For completeness, we also report in Tables 10, 11, 12 the different estimators, the first-stage F-tests and the OIR tests (as in Tables 4, 5, 6) for undirected networks. The qualitative evidence also remains roughly unchanged.

[*Insert Tables* 10 *to* 12 *here*]

7 Concluding remarks

In this paper, we have proposed two different social network models with different economic foundations. In the first one, the local-aggregate model, each individual is positively affected by the sum of effort of her friends while, in the local-average model, it is the deviation from the average effort of her friends that has a negative impact on her utility. We show that there exists a unique interior Nash equilibrium in each model.

We then test each model using the AddHealth data, which provides detailed information on adolescent friendships in the United States. We find that peer effects have no significant impact on screen activities (like e.g. video games). On the contrary, for sport activities, we find that students are mostly influenced by the sum of activities of their friends (local-aggregate model) while, for education, we show that both the aggregate school performance of friends and conformism matter, even though the magnitude of the effect is higher for the latter.

People tend to be very conformist in their behavior since it is well known that humans readily conform to the wishes or beliefs of others. Asch (1955, 1956) found that people will do this even in cases where they can obviously determine that others are incorrect. Asch presented subjects with two cards, one contained a single reference line and the other contained three lines of various lengths (one was the same length as the reference line). Asch manipulated the social situation by occasionally having two confederates publicly answer incorrectly prior to the subject providing an answer. The subject heard the incorrect responses of the others and was asked to publicly declare his answer as well. Asch found that the degree of conformity was relatively high. It is therefore surprising that, in the real-world, conformity do not always matter (like here for video games) or are less important (like here for sport activities) than the impact of the sum of active friends.

We believe that it is important to be able to disentangle between different behavioral peer-effect models because, as stated in the Introduction, it implies different policy implications. For example, in terms of education, to understand whether it is a local average or a local aggregate model, the prevailing mechanism of interactions would be helpful for policy makers to optimally design the composition of each classroom. Having around one very smart friend might not be equivalent to having three less smart friends. If the local aggregate mechanism of peer effects prevails, then classes should be heterogenous with respect to students' test scores, with the highly performing students distributed among the classes, if the objective is an increase of the average school performance. Under this scenario, a class composition where students are more homogeneous within each class would be less effective to reach the target. If a conformism mechanism of peer effects dominates, then heterogeneous classes can be used to reduce the variance of the students' grades within each class and homogenous classes to design excellence schemes where some classes are composed with the smartest students.

An effective policy in the local-average model would be to change people's perceptions of "normal" behavior (i.e. their social norm) while, in the local-aggregate model, this would not be necessary. Consider, for example, crime policies. In the local-aggregate model, a key-player policy (Ballester et al., 2006; Liu et al., 2011), whose aim is to remove the criminal that reduces the most total crime in a network, would be quite effective since the effort of each criminal and thus the sum of one's friends crime efforts will be reduced. On the contrary, a key-player policy would have nearly no effect in the local-average model since it will not affect the social norm of each group of friends in the network. To be effective, one would have to remove many key players in order to change the perception of crime in each peer group. In the Introduction, we discussed Eugene Lang's famous offer to give a college scholarship to *every student* at the sixth grade class in Harlem. This policy worked well because it changed the norm's group by affecting all its members. In the context of crime, in a conformist group, one needs to change people's perception of what is a "good" and what is a "bad" behavior. This is clearly a difficult goal to achieve.

This paper has uncovered some of the tricky aspects of how peers affect individual effort in different activities by highlighting that the difference between the "quantity" and "quality" of friends may play an important role. We have showed that the economic mechanism behind each model was quite different, leading to distinct policy remedies. This is a first step but there is much more work to be done. We leave this for future research.

References

[1] Akerlof, G.A. (1980), "A theory of social custom of which unemployment may be one consequence," *Quarterly Journal of Economics* 94, 749-775.

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

[3] Anson, R.S. (1985), *Best Intentions: The Education and Killing of Edmund Perry*, New York: Random House.

[4] Asch, S.E. (1955), "Opinions and social pressure," *Scientific American* 193, 31-35.

[5] Asch, S.E. (1956), "Studies of independence and submission to group pressure: I. A minority of one against a unanimous majority," *Psychological Monographs* 70.

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

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

[8] Battu, H., M. Mwale, and Y. Zenou (2007), "Oppositional identities and the labor market," *Journal of Population Economics* 20, 643-667.

[9] Berman, E. (2000), "Sect, subsidy, and sacrifice: An economist's view of ultra-orthodox Jews," *Quarterly Journal of Economics* 115, 905-953.

[10] Bernheim, B.D. (1994), "A theory of conformity," *Journal of Political Economy* 102, 841-877.

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

[12] Boucher, V., Bramoullé, Y., Djebbari, H. and B. Fortin (2010), "Do peers affect student achievement? Evidence from Canada using group size variation," IZA Discussion Paper No. 4723.

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

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

[15] Calvó-Armengol, A., Patacchini, E. and Y. Zenou (2005), "Peer effects and social networks in education and crime," CEPR Discussion Paper No. 5244.

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

[17] 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.

[18] Clark, A.E. and A.J. Oswald (1998), "Comparison-concave utility and following behaviour in social and economic settings," *Journal of Public Economics* 70, 133-155.

[19] Chao, J, C., Hausman, J., Newey, W.K., Swanson, N.R. and T.Woutersen (2010), "Testing overidentifying restrictions with many instruments and heteroskedasticity," University of Maryland Working Paper.

[20] Davidson R. and J.G. MacKinnon (1981), "Several tests for model specification in the presence of alternative hypotheses," *Econometrica* 49, 781-793.

[21] Davidson R. and J.G. MacKinnon (1993), *Estimation and inference in Econometrics*, New York: Oxford University Press.

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

[23] Fershtman, C., and Y. Weiss (1998), "Social rewards, externalities and stable preferences," *Journal of Public Economics* 70, 53-73.

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

[25] Ghiglino, C. and S. Goyal (2010), "Keeping up with the neighbors: Social interaction in a market economy," *Journal of the European Economic Association* 8, 90-119.

[26] Glaeser, E.L. and J. Scheinkman (2001), "Measuring social interactions," In: S. Durlauf and H.P. Young (Eds.), *Social Dynamics*, Cambridge: MIT Press.

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

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

[29] Horn, R. and C. Johnson (1990), *Matrix Analysis*, Cambridge: Cambridge University Press.

[30] Iannaccone, L. R. (1998), "Introduction to the economics of religion," *Journal of Economic Literature* 36, 1465-1496.

[31] Ioannides, Y.M. and D.L. Loury (2004), "Job information networks, neighborhood effects, and inequality," *Journal of Economic Literature* 42, 1056-1093.

[32] Ioannides, Y.M. (2011), *From Neighborhoods to Nations: The Economics of Social Interactions*, forthcoming.

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

[34] Kandel, E. and, E.P. Lazear (1992), "Peer pressure and partnerships," *Journal of Political Economy* 100, 801-817.

[35] Kelejian, H.H. (2008), "A spatial J-test for model specification against a single or a set of nonnested alternatives," *Letters in Spatial and Resources Sciences* 1, 3-11.

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

[37] Kelejian, H.H. and I.R. Prucha (2007), "The relative efficiencies of various predictors in spatial econometric models containing spatial lags," *Regional Science and Urban Economics* 37, 363-374.

[38] Kelejian, H.H. and G. Piras (2011), "An extension of Kelejian's J-test for non-nested spatial models," *Regional Science and Urban Economics* 41, 281-292

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

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

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

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

[43] Lee, L.F. and X. Liu (2010), "Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances," *Econometric Theory* 26, 187-230.

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

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

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

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

[48] Moffitt, R. (2001), "Policy interventions low-level equilibria, and social interactions," In: S. Durlauf and P. Young (Eds.), *Social Dynamics*, Cambridge, MA: MIT Press, pp. 45-82.

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

[50] Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economic, and Organization*, forthcoming.

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

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

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

[54] Whyte, W.F. (1955), *Street Corner Society*, Second Edition, Chicago: University of Chicago Press.

APPENDICES

A Proofs of propositions in the theoretical model

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

Proof of Proposition 2: Our utility function (6) can be written as:

$$\begin{aligned} u_{i,r}(Y_r, g_r) &= (a_{i,r} + \eta_r + \varepsilon_{i,r})\ y_{i,r} - \frac{1}{2} y_{i,r}^2 - \frac{d}{2}\left(y_{i,r} - \overline{y}_{i,r}\right)^2 \\ &= (a_{i,r} + \eta_r + \varepsilon_{i,r})\ y_{i,r} - \frac{(1+d)}{2} y_{i,r}^2 + d y_{i,r}\overline{y}_{i,r} - \frac{d}{2}\overline{y}_{i,r}^2 \end{aligned}$$

Using (5), this can be written as:

$$u_{i,r}(Y_r, g_r) = (a_{i,r} + \eta_r + \varepsilon_{i,r})\ y_{i,r} - \frac{(1+d)}{2} y_{i,r}^2 + d\sum_{j=1}^{n_r} g_{ij,r}^* y_{i,r} y_{j,r} - \frac{d}{2}\left(\sum_{j=1}^{n_r} g_{ij,r}^* y_{j,r}\right)^2$$

We can now apply Theorem 1, part b, in Calvó-Armengol et al. (2009) to our problem.[24] The condition on eigenvalue (that guarantees that the Nash equilibrium is unique and interior) can now be written as: $1 + d > d\omega_1(G_r^*)$. Observe that G_r^* is a stochastic matrix so that its largest eigenvalue is 1, i.e., $\omega_1(G_r^*) = 1$. As a result, the condition $1 + d > d\omega_1(G_r^*)$ can be written as $1 + d > d$, which is always true. As a result, the only condition needed is $\phi_2 < 1$ for the Nash equilibrium (7) to be well-defined. ▮

B 2SLS and GMM Estimation

We consider 2SLS and GMM estimators for the estimation of an empirical hybrid network model, and for the estimation of augmented models in the J test. This appendix presents the derivation and asymptotic properties of the estimators.

For any $n \times n$ matrix $A = [a_{ij}]$, let $\text{vec}_D(A) = (a_{11}, \cdots, a_{nn})'$, $A^s = A + A'$, $A^t = A - \text{tr}(A)J/\text{tr}(J)$, and A^- denote a generalized inverse of a square matrix A. For a parameter θ, let θ_0 denote the true parameter value in the data generating process. Let μ_3 and μ_4 denote, respectively, the third and fourth moments of the error term.

[24] Observe that the term $-\frac{d}{2}\left(\sum_{j=1}^{n_r} g_{ij,r}^* y_{j,r}\right)^2$ does not matter since the derivative of this term with respect to $y_{i,r}$ is equal to zero.

B.1 Estimation of a hybrid network model

B.1.1 2SLS estimation

Let $M_0 = (I - \phi_{10}G - \phi_{20}G^*)^{-1}$. From the reduced form equation (20), $\mathrm{E}(Y) = M_0(X^*\delta_0 + \iota \cdot \eta)$.[25] For $Z = [GY, G^*Y, X^*]$, the ideal IV matrix for the explanatory variables JZ in (21) is given by

$$f = \mathrm{E}(JZ) = J[G\mathrm{E}(Y), G^*\mathrm{E}(Y), X^*]. \tag{22}$$

However, this IV matrix is infeasible as it involves unknown parameters. Note that f can be considered as a linear combination of the IVs in $Q_\infty = J[GM_0X^*, GM_0\iota, G^*M_0X^*, G^*M_0\iota, X^*]$. As ι has $\bar{r}$ columns, the number of IVs in Q_∞ increases as the number of groups $\bar{r}$ increases. Furthermore, if $|\phi_{10}\max_i(\sum_j g_{ij})| + |\phi_{20}| < 1$,[26] we have $M_0 = (I - \phi_{10}G - \phi_{20}G^*)^{-1} = \sum_{j=0}^{\infty}(\phi_{10}G + \phi_{20}G^*)^j$. Hence, M_0 in Q_∞ can be approximated by a linear combination of $[I, G, G^*, G^2, GG^*, G^*G, G^{*2}, \cdots]$.

To achieve asymptotic efficiency, we assume the number of IVs increases with the sample size so that the ideal IV matrix f can be approximated by a feasible IV matrix Q_K with an approximation error diminishing to zero. That is, for an $n \times K$ IV matrix Q_K premultiplied by J, there exists some conformable matrix π_K such that $||f - Q_K\pi_K||_\infty \to 0$ as $n, K \to \infty$. Let $P_K = Q_K(Q_K'Q_K)^- Q_K'$, the 2SLS estimator consider is $\hat{\theta}_{2sls} = (Z'P_KZ)^{-1}Z'P_KY$.

Let $\theta_0 = (\phi_{10}, \phi_{20}, \delta_0')'$. If $K/n \to 0$, then it follows by a similar argument as in Liu and Lee (2010) that $\sqrt{n}(\hat{\theta}_{2sls} - \theta_0 - b_{2sls}) \xrightarrow{d} N(0, \sigma^2\bar{H}^{-1})$, where $\bar{H} = \lim_{n\to\infty}\frac{1}{n}f'f$ and $b_{2sls} = \sigma^2(Z'P_KZ)^{-1}[\mathrm{tr}(P_KGM_0), \mathrm{tr}(P_KG^*M_0), 0_{1\times 2m}]' = O_p(K/n)$. The 2SLS estimator has an asymptotic bias term due to the large number of IVs. When $K^2/n \to 0$, the leading order bias term $\sqrt{n}b_{2sls}$ converges to zero and the proposed 2SLS estimator is efficient as the variance matrix $\sigma^2\bar{H}^{-1}$ attains the efficiency lower bound for the class of IV estimators.

[25] For simplicity, we assume G and X are nonstochastic. If G and X are stochastic, then the following results can be considered as conditional on G and X.

[26] The model represents an equilibrium so $I - \phi_{10}G - \phi_{20}G^*$ is assumed to be invertible. A sufficient condition for the invertibility assumption can be derived as follows. Let g_i being the ith row sum of G. Since G^* is the row-normalized G, we have $G = RG^*$, where R is a diagonal matrix with the ith diagonal element being g_i. The invertibility of $I - \phi_{10}G - \phi_{20}G^*$ requires $||\phi_{10}G + \phi_{20}G^*|| < 1$ for some matrix norm $||\cdot||$ (see Horn and Johnson, 1990). Let $||\cdot||_\infty$ denote the row-sum matrix norm. As $||G^*||_\infty = 1$, we have $||\phi_{10}G + \phi_{20}G^*||_\infty = ||\phi_{10}RG^* + \phi_{20}G^*||_\infty \le ||\phi_{10}R + \phi_{20}I||_\infty$. Therefore, a sufficient condition for the invertibility assumption would be $|\phi_{10}\max_i(g_i)| + |\phi_{20}| < 1$.

On the other hand, a sufficient condition for the the invertibility of $I - \phi_{10}G$ for the local aggregate model is $|\phi_{10}| < 1/\max_i(g_i)$ (see Liu and Lee, 2010) and a sufficient condition for the the invertibility of $I - \phi_{20}G^*$ for the local average model is $|\phi_{20}| < 1$. Both of them are weaker than the invertibility condition of the hybrid model.

To correct for the many-instrument bias in the 2SLS estimator, one can estimate the leading order bias term and adjust the 2SLS estimator by the estimated leading-order bias $\tilde{b}_{2sls}$. With $\sqrt{n}$-consistent initial estimates $\check{\sigma}^2, \check{\phi}_1, \check{\phi}_2$, the bias-corrected 2SLS (C2SLS) is given by $\hat{\theta}_{c2sls} = \hat{\theta}_{2sls} - \tilde{b}_{2sls}$, where $\tilde{b}_{2sls} = \check{\sigma}^2(Z'P_K Z)^{-1}[\text{tr}(P_K GM), \text{tr}(P_K G^* M), 0_{1\times 2m}]'$ and $M = (I - \check{\phi}_1 G - \check{\phi}_2 G^*)^{-1}$. The C2SLS is efficient when $K/n \to 0$.

B.1.2 GMM estimation

The 2SLS estimator can be generalized to the GMM with additional quadratic moment equations. Let $\epsilon(\theta) = J(Y - Z\theta)$. The IV moment conditions $Q_K'\epsilon(\theta) = 0$ are linear in ϵ at θ_0. As $\text{E}(\epsilon' U_1 \epsilon) = \text{E}(\epsilon' U_2 \epsilon) = 0$ for $U_1 = (JGM_0 J)^t$ and $U_2 = (JG^* M_0 J)^t$, the quadratic moment conditions for estimation are given by $[U_1\epsilon(\theta), U_2\epsilon(\theta)]'\epsilon(\theta) = 0$. The proposed quadratic moment conditions can be shown to be optimal (in terms of efficiency of the GMM estimator) under normality (see Lee and Liu, 2010). The vector of linear and quadratic empirical moments for the GMM estimation is given by $g(\theta) = [Q_K, U_1\epsilon(\theta), U_2\epsilon(\theta)]'\epsilon(\theta)$.

In order for inference based on the following asymptotic results to be robust, we do not impose the normality assumption for the following analysis. The variance matrix of $g(\theta_0)$ is given by

$$\Omega = \text{Var}[g(\theta_0)] = \begin{pmatrix} \sigma^2 Q_K' Q_K & \mu_3 Q_K' \omega \\ \mu_3 \omega' Q_K & (\mu_4 - 3\sigma^4)\omega'\omega + \sigma^4 \Delta \end{pmatrix},$$

where $\omega = [\text{vec}_D(U_1), \text{vec}_D(U_2)]$ and $\Delta = \frac{1}{2}[\text{vec}(U_1^s), \text{vec}(U_2^s)]'[\text{vec}(U_1^s), \text{vec}(U_2^s)]$. By the generalized Schwarz inequality, the optimal GMM estimator is given by

$$\hat{\theta}_{gmm} = \arg\min g'(\theta)\Omega^{-1}g(\theta). \tag{23}$$

Let $B^{-1} = (\mu_4 - 3\sigma^4)\omega'\omega + \sigma^4\Upsilon - \frac{\mu_3^2}{\sigma^2}\omega' P_K \omega$,

$$D = -\sigma^2 \begin{pmatrix} \text{tr}(U_1^s G M_0) & \text{tr}(U_1^s G^* M_0) & 0_{1\times 2m} \\ \text{tr}(U_2^s G M_0) & \text{tr}(U_2^s G^* M_0) & 0_{1\times 2m} \end{pmatrix},$$

$\bar{D} = D - \frac{\mu_3}{\sigma^2}\omega' f$, and $\check{D} = D - \frac{\mu_3}{\sigma^2}\omega' P_K Z$. When $K^{3/2}/n \to 0$, the optimal GMM estimator[27] has the asymptotic distribution

$$\sqrt{n}(\hat{\theta}_{gmm} - \theta_0 - b_{gmm}) \xrightarrow{d} N(0, (\sigma^{-2}\bar{H} + \lim_{n\to\infty}\frac{1}{n}\bar{D}' B \bar{D})^{-1}), \tag{24}$$

[27] The weighting matrices for quadratic moments U_1, U_2 and the optimal weighting matrix for the objective function Ω^{-1} involves unknown parameters $\phi_1, \phi_2, \sigma_0^2, \mu_3$ and μ_4. With consistent preliminary estimators of those unknown parameters, the feasible optimal GMM estimator can be shown to have the same asymptotic distribution given by (24).

where $b_{gmm} = (\sigma^{-2}Z'P_KZ + \check{D}'B\check{D})^{-1}[\text{tr}(P_KGM_0), \text{tr}(P_KG^*M_0), 0_{1\times 2m}]' = O(K/n)$.

As the asymptotic bias $\sqrt{n}b_{gmm}$ is $O(K/\sqrt{n})$, the asymptotic distribution of the GMM estimator $\hat{\theta}_{gmm}$ will be centered at θ_0 only if $K^2/n \to 0$. With a consistently estimated leading order bias $\tilde{b}_{gmm}$, the bias-corrected GMM (CGMM) estimator $\hat{\theta}_{cgmm} = \hat{\theta}_{gmm} - \tilde{b}_{gmm}$ has a proper centered asymptotic normal distribution as given in (24) if $K^{3/2}/n \to 0$.

The asymptotic variance matrix of the many-IV GMM estimator can be compared with that of the many-IV 2SLS estimator. As $\bar{D}'B\bar{D}$ is nonnegative definite, the asymptotic variance of the many-IV GMM estimator is relatively smaller than that of the 2SLS estimator. Thus, the many-IV GMM estimator with additional quadratic moments improves efficiency upon the 2SLS estimator.

B.2 Estimation of augmented models in the J test

In this subsection, we focus on the estimation of the augmented model in the test of model H_1 against model H_2. The estimator for the test of model H_2 against model H_1 can be derived in a similar manner.

B.2.1 2SLS estimation of the augmented model

First, we consider the 2SLS estimator of the augmented model (14). Let $M_{10} = (I - \alpha_{10}\phi_{20}G^* - \phi_{10}G)^{-1}$. The ideal IV matrix for JZ_1^* in (14) is given by $f_1 = \text{E}(JZ_1^*) = J[\phi_{20}G^*\text{E}(Y) + X^*\delta_{20}, G\text{E}(Y), X^*]$, where $\text{E}(Y) = M_{10}[X^*(\alpha_{10}\delta_{20} + \delta_{10}) + \iota \cdot (\eta_1 + \alpha_{10}\eta_2)]$. The ideal IV matrix f_1 is infeasible as it involves unknown parameters. We note that f_1 can be considered as a linear combination of the IVs in $Q_\infty = J[G^*M_{10}X^*, G^*M_{10}\iota, GM_{10}X^*, GM_{10}\iota, X^*]$. Furthermore, under some regularity conditions, $M_{10} = (I - \alpha_{10}\phi_{20}G^* - \phi_{10}G)^{-1} = \sum_{j=0}^{\infty}(\alpha_{10}\phi_{20}G^* + \phi_{10}G)^j$. Hence, M_{10} in Q_∞ can be approximated by polynomials of I, G and G^*.

To achieve asymptotic efficiency, we consider an $n \times K$ feasible submatrix of Q_∞, denoted by Q_K, such that the ideal IV matrix f_1 can be approximated by a linear combination of Q_K with an approximation error diminishing to zero as the number of IVs K increases. Let $P_K = Q_K(Q_K'Q_K)^-Q_K'$ and $\tilde{Z}_1^* = [(\tilde{\phi}_2G^*Y + X^*\tilde{\delta}_2), GY, X^*]$, where $\tilde{\phi}_2, \tilde{\delta}_2$ are $\sqrt{n}$-consistent preliminary estimates. The 2SLS estimator considered is $\hat{\vartheta}_{1,tsls} = (\tilde{Z}_1^{*\prime}P_K\tilde{Z}_1^*)^{-1}\tilde{Z}_1^{*\prime}P_KY$.

Under the null hypothesis, it follows by a similar argument as in Liu and Lee (2010) that if $K/n \to 0$ then $\sqrt{n}(\hat{\vartheta}_{1,tsls} - \vartheta_{10} - b_{1,tsls}) \xrightarrow{d} N(0, \sigma_1^2\bar{H}_1^{-1})$, where $\bar{H}_1 = \lim_{n\to\infty}\frac{1}{n}f_1'f_1$ and $b_{1,tsls} = \sigma_1^2(\tilde{Z}_1^{*\prime}P_K\tilde{Z}_1^*)^{-1}[\phi_{20}\text{tr}(P_KG^*M_{10}), \text{tr}(P_KGM_{10}), 0_{1\times 2m}]'$. The term $b_{1,tsls}$ is a bias due to the presence of many IVs. We can adjust for the many-IV bias by considering the C2SLS estimator $\hat{\vartheta}_{1,ctsls} = \hat{\vartheta}_{1,tsls} - \tilde{b}_{1,tsls}$, where $\tilde{b}_{1,tsls}$ is a consistent estimator of $b_{1,tsls}$. If $K/n \to 0$

then $\sqrt{n}(\hat{\vartheta}_{1,ctsls} - \vartheta_{10}) \xrightarrow{d} N(0, \sigma_1^2 \bar{H}_1^{-1})$.

B.2.2 GMM estimation of the augmented model

The GMM estimator uses both linear moment conditions $Q_K'\epsilon_1(\vartheta_1) = 0$ and quadratic ones $[U_1\epsilon_1(\vartheta_1), U_2\epsilon_1(\vartheta_1)]'\epsilon_1(\vartheta_1) = 0$, where $U_1 = (JG^*M_{10}J)^t$, $U_2 = (JGM_{10}J)^t$, and $\epsilon_1(\vartheta_1) = J(Y - \tilde{Z}_1^*\vartheta_1)$. The vector of linear and quadratic empirical moment functions for the GMM estimation is given by $g_1(\vartheta_1) = [Q_K, U_1\epsilon_1(\vartheta_1), U_2\epsilon_1(\vartheta_1)]'\epsilon_1(\vartheta_1)$. By the generalized Schwarz inequality, the optimal GMM estimator is given by $\hat{\vartheta}_{1,gmm} = \arg\min g_1'(\vartheta_1)\Omega^{-1}g_1(\vartheta_1)$, where

$$\Omega = \begin{pmatrix} \sigma_1^2 Q_K' Q_K & \mu_3 Q_K'\omega \\ \mu_3\omega' Q_K & (\mu_4 - 3\sigma_1^4)\omega'\omega + \sigma_1^4\Upsilon \end{pmatrix},$$

$\omega = [\mathrm{vec}_D(U_1), \mathrm{vec}_D(U_2)]$ and $\Upsilon = \frac{1}{2}[\mathrm{vec}(U_1^s), \mathrm{vec}(U_2^s)]'[\mathrm{vec}(U_1^s), \mathrm{vec}(U_2^s)]$.

Let $B_1^{-1} = (\mu_4 - 3\sigma_1^4)\omega'\omega + \sigma_1^4\Upsilon - \frac{\mu_3^2}{\sigma_1^2}\omega' P_K\omega$,

$$D_1 = -\sigma_1^2 \begin{pmatrix} \phi_{20}\mathrm{tr}(U_1^s G^* M_{10}) & \mathrm{tr}(U_1^s G M_{10}) & 0_{1\times 2m} \\ \phi_{20}\mathrm{tr}(U_2^s G^* M_{10}) & \mathrm{tr}(U_2^s G M_{10}) & 0_{1\times 2m} \end{pmatrix},$$

$\bar{D}_1 = D_1 - \frac{\mu_3}{\sigma_1^2}\omega' f_1$, and $\check{D}_1 = D_1 - \frac{\mu_3}{\sigma_1^2}\omega' P_K \tilde{Z}_1^*$. Under the null hypothesis, if $K^{3/2}/n \to 0$, the optimal GMM estimator[28] has the asymptotic distribution

$$\sqrt{n}(\hat{\vartheta}_{1,gmm} - \vartheta_{10} - b_{1,gmm}) \xrightarrow{d} N(0, (\sigma_1^{-2}\bar{H}_1 + \lim_{n\to\infty}\frac{1}{n}\bar{D}_1' B_1 \bar{D}_1)^{-1}), \tag{25}$$

where $b_{1,gmm} = (\sigma_1^{-2}\tilde{Z}_1^{*\prime} P_K \tilde{Z}_1^* + \check{D}_1' B_1 \check{D}_1)^{-1}[\phi_{20}\mathrm{tr}(P_K G^* M_{10}), \mathrm{tr}(P_K G M_{10}), 0_{1\times 2m}]' = O(K/n)$.

With a consistently estimated leading order bias $\tilde{b}_{1,gmm}$, it follows by a similar argument as in Liu and Lee (2010) that, if $K^{3/2}/n \to 0$, the CGMM estimator $\hat{\vartheta}_{1,cgmm} = \hat{\vartheta}_{1,gmm} - \tilde{b}_{1,gmm}$ has a proper centered asymptotic normal distribution as given in (25).

[28] With consistent preliminary estimates of the unknown parameters in U_1, U_2, Ω, the feasible optimal GMM estimator can be shown to have the same asymptotic distribution given by (25).

Appendix C: Data appendix

Table C.1: Description of Data

Label	**Variable definition**	**Screen data** *(n.obs. 3196 networks 33)*		**Sport data** *(n.obs. 2934 networks 32)*		**GPA data** *(n.obs. 1443 networks 13)*	
		mean	std. dev.	mean	std. dev.	mean	std.dev.
Screen activity index	In the text	1.37	1.23	-	-	-	-
Sport activity index	In the text	-	-	1.53	1.05	-	-
GPA index	In the text	-	-	-	-	3.03	1.10
Individual socio-demographic variables							
Female	Dummy variable taking value one if the respondent is female.	0.53	0.50	0.53	0.50	0.51	0.5
Black	Ethnic group dummies, white is the reference category	0.18	0.38	0.18	0.39	0.14	0.35
Other races	“	0.09	0.28	0.07	0.26	0.21	0.40
Grade	Grade of the student in the current year.	9.59	1.58	9.47	1.58	9.61	1.7
Self esteem (Screen, Sport)	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.	3.90	1.06	3.90	1.06		
Self esteem (Gpa)	Response to the question: "I have a lot of good qualities", coded as 1= strongly agree, 2= agree, 3= neither agree nor disagree, 4= disagree, 5= strongly disagree.					1.72	0.76
Math_sc_A (Screen, Sport)	Mathematics score dummies, including a category capturing missing values. D is the reference category	0.29	0.45	0.31	0.46		
Math_sc_B (Screen, Sport)	“	0.31	0.46	0.32	0.47		
Math_sc_C (Screen, Sport)	“	0.23	0.42	0.22	0.42		
Math_sc_mis (Screen, Sport)	“	0.06	0.24	0.05	0.22		
Teacher troubles	Response to the question: “How often have you had trouble getting along with your teachers?” 0= never, 1= just a few times, 2= about once a week, 3= almost everyday, 4=everyday	0.84	0.9	0.82	0.88	1.17	1.26

School attachment	Response to the question: “I feel like you are part of your school", all coded as 1= strongly agree, 2= agree, 3=neither agree nor disagree, 4= disagree, 5= strongly disagree.	4	0.96	4.03	0.94	2.22	1.12
Family background variables							
Family size	Number of people living in the household	4.45	1.38	4.46	1.37	4.26	1.13
Parental education (Screen, Sport)	Schooling level of the (biological or non-biological) parent who is living with the child, distinguishing between "eighth grade or less", "more than eighth grade, but did not graduate from high school", "high school graduate", "completed a GED", "went to a business, trade, or vocational school after high school", “went to college but did not graduate”, “graduated from college or a university”, ” professional training beyond a four-year college”, coded as 1 to 8. If both parents are in the household the education of the father is considered. It is coded as zero if no parent lives with child or the reported level is “unknown”.	5.03	2.26	5.02	2.27		
Parental education (Gpa)	Schooling level of the (biological or non-biological) 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. If both parents are in the household the education of the father is considered. It is coded as zero if no parent lives with child or the reported level is “unknown”.					2.78	0.97
Parent occupation manager	Parent occupation dummies. Closest description of the job of (biological or non-biological) 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	0.1	0.29	0.10	0.3	0.13	0.34
Parent occupation professional/technical	”	0.17	0.37	0.17	0.37	0.22	0.41
Parent occupation office or sales worker	”	0.11	0.31	0.11	0.31	0.14	0.34
Parent occupation manual	”	0.35	0.48	0.35	0.48	0.38	0.49
Parent occupation military or security	”	0.03	0.17	0.03	0.17	0.04	0.19
Parent occupation farm or fishery	”	0.03	0.17	0.03	0.17	0.01	0.1
Parent occupation other	”	0.15	0.36	0.15	0.36	0.04	0.18
Married Parents	Dummy variable taking value one if the child leaves in a family with both parents who are married	0.73	0.44	0.74	0.44	0.78	0.41

Parental care	Dummy taking value one if both parents care very much about her/him	0.92	0.27	0.92	0.27	0.57	0.50
Residential neighborhood variables							
Neighborhood quality (Screen, Sport)	Interviewer response to the question "How well kept is the building in which the respondent lives", coded as 4= very poorly kept (needs major repairs), 3= poorly kept (needs minor repairs), 2= fairly well kept (needs cosmetic work), 1= very well kept.	1.53	0.8	1.53	0.81	-	-
Neighborhood quality (Gpa)	Response to the question: "I feel safe in my neighborhood" all coded as 1= strongly agree, 2= agree, 3=neither agree nor disagree, 4= disagree, 5= strongly disagree.	-	-	-	-	2.07	1.07
Total number of links	Number of individual social contacts derived from the nomination data (Directed networks)	2.18	2.05	2.14	2.00	2.34	2.13
	Number of individual social contacts derived from the nomination data (Undirected networks)	3.48	2.42	3.43	2.35	3.68	2.45

Notes: *We consider networks with network size between 50 and 150 individuals. The reported summary statistics are for the samples which are constructed using directed network, unless differently specified. Screen and Sport data are derived from the Add Health in-home questionnaire. GPA data are instead derived from the Add-Health in-school questionnaire. Some variables have been slightly redefined. Differences in sample sizes are mainly due to missing values in variables.*

Table 1: Testing local aggregate versus local average models for sport activities
Increasing sets of controls- biased corrected optimal GMM
Directed networks
Hybrid model (19)

	(1)	(2)	(3)	(4)
Aggregate (φ_1)	**0.0216*** **	**0.0217*** **	**0.0215*** **	**0.0213*** **
	(0.0065)	**(0.0066)**	**(0.0065)**	**(0.0065)**
Average (φ_2)	**0.0148**	**0.0143**	**0.0136**	**0.0129**
	(0.0266)	**(0.0266)**	**(0.0266)**	**(0.0266)**
Female	-0.5511***	-0.5523***	-0.5511***	-0.5502***
	(0.0379)	(0.0380)	(0.0379)	(0.0379)
Grade	-0.0638***	-0.0624***	-0.0653***	-0.0651***
	(0.0174)	(0.0174)	(0.0174)	(0.0175)
Math_sc_A	0.0881	0.0866	0.0251	0.0300
	(0.0684)	(0.0685)	(0.0700)	(0.0700)
Math_sc_B	0.0646	0.0613	0.0335	0.0372
	(0.0669)	(0.0670)	(0.0672)	(0.0673)
Math_sc_C	0.0574	0.0547	0.0460	0.0497
	(0.0694)	(0.0695)	(0.0693)	(0.0695)
Math_sc_mis	-0.0073	-0.0055	-0.0226	-0.0125
	(0.1031)	(0.1031)	(0.1030)	(0.1031)
Black	-0.0679	-0.0724	-0.0798	-0.0836
	(0.0775)	(0.0782)	(0.0781)	(0.0782)
Other races	-0.1722*	-0.1775**	-0.1738*	-0.1699*
	(0.0905)	(0.0906)	(0.0904)	(0.0908)
Parental education	0.0066	0.0066	0.0059	0.0061
	(0.0081)	(0.0081)	(0.0081)	(0.0081)
School attachment	0.1350***	0.1351***	0.1277***	0.1284***
	(0.0204)	(0.0204)	(0.0206)	(0.0206)
Teacher troubles	-0.0003	0.0004	0.0037	0.0059
	(0.0220)	(0.0220)	(0.0220)	(0.0220)
Neighborhood quality		-0.0046	0.0010	0.0042
		(0.0244)	(0.0244)	(0.0247)
Family size		0.0258*	0.0261*	0.0276**
		(0.0142)	(0.0141)	(0.0141)
Married parents		-0.0223	-0.0315	-0.0244
		(0.0458)	(0.0459)	(0.0475)
Self esteem			0.0762***	0.0724***
			(0.0187)	(0.0188)
Parental care			-0.0362	-0.0336
			(0.0697)	(0.0698)
Network fixed effects	yes	yes	Yes	yes
Parental occupation dummies	no	no	No	yes

J test	**Model (12)** Null Hp. (α_1=0)	No local aver.	*t statistic*	**0.483**
	Model (16) Null Hp. (α_2=0)	No local aggr.	*t statistic*	**3.252**

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 2: Testing local aggregate versus local average models for screen activities
Increasing sets of controls- biased corrected optimal GMM
Directed networks
Hybrid model (19)

	(1)	(2)	(3)	(4)
Aggregate (φ_1)	**-0.0022**	**-0.0022**	**-0.0022**	**-0.0026**
	(0.0089)	**(0.0089)**	**(0.0089)**	**(0.0089)**
Average (φ_2)	**-0.0118**	**-0.0126**	**-0.0123**	**-0.0111**
	(0.0278)	**(0.0279)**	**(0.0279)**	**(0.0279)**
Female	-0.1702***	-0.1685***	-0.1683***	-0.1676***
	(0.0445)	(0.0446)	(0.0446)	(0.0446)
Grade	-0.0891***	-0.0897***	-0.0892***	-0.0897***
	(0.0204)	(0.0205)	(0.0205)	(0.0205)
Math_sc_A	0.0064	0.0050	0.0069	0.0041
	(0.0784)	(0.0785)	(0.0808)	(0.0808)
Math_sc_B	0.1012	0.1012	0.1034	0.0970
	(0.0767)	(0.0767)	(0.0773)	(0.0774)
Math_sc_C	0.1352**	0.1350**	0.1348**	0.1316**
	(0.0794)	(0.0795)	(0.0796)	(0.0797)
Math_sc_mis	0.1554	0.1512	0.1507	0.1545
	(0.1123)	(0.1124)	(0.1126)	(0.1126)
Black	0.2973***	0.3038***	0.3037***	0.3038***
	(0.0885)	(0.0894)	(0.0895)	(0.0895)
Other races	0.0678	0.0739	0.0750	0.0940
	(0.0992)	(0.0995)	(0.0995)	(0.0999)
Parental education	-0.0041	-0.0040	-0.0040	-0.0049
	(0.0095)	(0.0095)	(0.0095)	(0.0096)
School attachment	-0.0307	-0.0307	-0.0321	-0.0309
	(0.0232)	(0.0233)	(0.0236)	(0.0236)
Teacher troubles	0.0556***	0.0557***	0.0568***	0.0568***
	(0.0250)	(0.0250)	(0.0251)	(0.0252)
Neighborhood quality		0.0053	0.0058	0.0077
		(0.0286)	(0.0286)	(0.0289)
Family size		-0.0168	-0.0166	-0.0158
		(0.0164)	(0.0164)	(0.0164)
Married parents		0.0313	0.0284	-0.0016
		(0.0535)	(0.0538)	(0.0555)
Self esteem			-0.0026	-0.0017
			(0.0220)	(0.0220)
Parental care			0.0544	0.0560
			(0.0816)	(0.0816)
Network fixed effects	yes	Yes	yes	yes
Parental occupation dummies	no	No	no	yes

	Model (12)			
J test	Null Hp. (α_1=0)	No local aver.	*t statistic*	**0.368**
	Model (16)			
	Null Hp. (α_2=0)	No local aggr.	*t statistic*	**0.312**

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 3: Testing local aggregate versus local average models for education (GPA)
Increasing sets of controls- biased corrected optimal GMM
Directed networks
Hybrid model (19)

	(1)	(2)	(3)	(4)
Aggregate (φ_1)	**0.0153*** **	**0.0150*** **	**0.0149*** **	**0.0146*** **
	(0.0046)	**(0.0045)**	**(0.0045)**	**(0.0045)**
Average (φ_2)	**0.2279*** **	**0.2258*** **	**0.2216*** **	**0.2177*** **
	(0.0340)	**(0.0340)**	**(0.0340)**	**(0.0339)**
Female	0.2066***	0.2062***	0.2321***	0.2320***
	(0.0561)	(0.0562)	(0.0565)	(0.0566)
Grade	0.0487*	0.0497*	0.0517*	0.0563**
	(0.0271)	(0.0271)	(0.0270)	(0.0270)
Black	-0.2466**	-0.2006	-0.2374*	-0.2585**
	(0.1249)	(0.1258)	(0.1254)	(0.1255)
Other races	0.0114	0.0098	0.0103	0.0167
	(0.0792)	(0.0793)	(0.0789)	(0.0790)
Parental education	0.0691***	0.0605***	0.0636***	0.0578***
	(0.0275)	(0.0276)	(0.0275)	(0.0279)
School attachment	-0.1237***	-0.1152***	-0.0877***	-0.0912***
	(0.0236)	(0.0243)	(0.0251)	(0.0252)
Teacher troubles	-0.0850***	-0.0841***	-0.0804***	-0.0814***
	(0.0212)	(0.0212)	(0.0211)	(0.0211)
Neighborhood quality		-0.0168	0.0065	0.0106
		(0.0258)	(0.0263)	(0.0263)
Family size		-0.0060	-0.0069	-0.0048
		(0.0248)	(0.0247)	(0.0247)
Married parents		0.2078***	0.0994	0.1162
		(0.0687)	(0.0832)	(0.0864)
Self esteem			-0.1213***	-0.1237***
			(0.0371)	(0.0371)
Parental care			0.1514***	0.1481***
			(0.0684)	(0.0684)
Network fixed effects	yes	Yes	yes	yes
Parental occupation dummies	no	No	no	yes

	Model (12)			
J test	Null Hp. (α_1=0)	No local aver.	*t statistic*	**6.411**
	Model (16)			
	Null Hp. (α_2=0)	No local aggr.	*t statistic*	**3.235**

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 4: Testing local aggregate versus local average models for sport activities
Hybrid model (19) only - all controls - alternative estimators-
Directed networks

	2SLS-1	**2SLS-2**	**C2SLS**	**GMM-1**	**GMM-2**	**CGMM**
Aggregate (φ_1)	**0.0208*** **	**0.0243*** **	**0.0262*** **	**0.0223*** **	**0.0223*** **	**0.0213*** **
	(0.0077)	**(0.0074)**	**(0.0077)**	**(0.0066)**	**(0.0065)**	**(0.0065)**
Average (φ_2)	**0.0182**	**-0.2539**	**-0.4910**	**0.0209**	**0.0160**	**0.0129**
	(0.3125)	**(0.1670)**	**(0.1745)**	**(0.0267)**	**(0.0266)**	**(0.0266)**
Female	-0.5507***	-0.5496***	-0.5487***	-0.5503***	-0.5502***	-0.5502***
	(0.0380)	(0.0387)	(0.0404)	(0.0379)	(0.0379)	(0.0379)
Grade	-0.0657***	-0.0715***	-0.0767***	-0.0650***	-0.0648***	-0.0651***
	(0.0187)	(0.0181)	(0.0189)	(0.0175)	(0.0175)	(0.0175)
Math_sc_A	0.0295	0.0252	0.0219	0.0297	0.0296	0.0300
	(0.0702)	(0.0714)	(0.0746)	(0.0700)	(0.0700)	(0.0700)
Math_sc_B	0.0364	0.0432	0.0496	0.0363	0.0367	0.0372
	(0.0678)	(0.0687)	(0.0717)	(0.0673)	(0.0673)	(0.0673)
Math_sc_C	0.0493	0.0462	0.0437	0.0497	0.0496	0.0497
	(0.0696)	(0.0708)	(0.0739)	(0.0695)	(0.0695)	(0.0695)
Math_sc_mis	-0.0129	-0.0306	-0.0458	-0.0117	-0.0125	-0.0125
	(0.1050)	(0.1056)	(0.1102)	(0.1031)	(0.1031)	(0.1031)
Black	-0.0839	-0.1004	-0.1149	-0.0825	-0.0833	-0.0836
	(0.0804)	(0.0802)	(0.0838)	(0.0782)	(0.0782)	(0.0782)
Other races	-0.1717*	-0.1612*	-0.1517	-0.1708*	-0.1703	-0.1699*
	(0.0917)	(0.0927)	(0.0969)	(0.0908)	(0.0908)	(0.0908)
Parental education	0.0060	0.0071	0.0081	0.0060	0.0060	0.0061
	(0.0082)	(0.0083)	(0.0087)	(0.0081)	(0.0081)	(0.0081)
School attachment	0.1278***	0.1309***	0.1339***	0.1279***	0.1281***	0.1284***
	(0.0210)	(0.0211)	(0.0220)	(0.0206)	(0.0206)	(0.0206)
Teacher troubles	0.0045	0.0118	0.0182	0.0053	0.0057	0.0059
	(0.0236)	(0.0228)	(0.0239)	(0.0220)	(0.0220)	(0.0220)
Neighborhood quality	0.0052	-0.0020	-0.0087	0.0049	0.0048	0.0042
	(0.0261)	(0.0255)	(0.0266)	(0.0247)	(0.0247)	(0.0247)
Family size	0.0275*	0.0276*	0.0277*	0.0276**	0.0276**	0.0276**
	(0.0142)	(0.0144)	(0.0151)	(0.0141)	(0.0141)	(0.0141)
Married parents	-0.0238	-0.0178	-0.0128	-0.0246	-0.0243	-0.0244
	(0.0480)	(0.0485)	(0.0507)	(0.0475)	(0.0475)	(0.0475)
Self esteem	0.0723***	0.0731***	0.0738***	0.0723***	0.0724***	0.0724***
	(0.0188)	(0.0191)	(0.0200)	(0.0188)	(0.0188)	(0.0188)
Parental care	-0.0360	-0.0291	-0.0229	-0.0346	-0.0339	-0.0336
	(0.0702)	(0.0712)	(0.0743)	(0.0698)	(0.0698)	(0.0698)
Network fixed effects	yes	yes	yes	yes	yes	yes
Parental occupation dummies	yes	yes	yes	yes	yes	yes
1st Stage F statistic	**0.883**	**1.369**				
OIR test			***p-value***	**0.887**		

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 5: Testing local aggregate versus local average models for screen activities
Hybrid model (19) only - all controls - alternative estimators-
Directed networks

	2SLS-1	**2SLS-2**	**C2SLS**	**GMM-1**	**GMM-2**	**CGMM**
Aggregate (φ_1)	**-0.0021**	**-0.0017**	**-0.0030**	**-0.0008**	**-0.0006**	**-0.0026**
	(0.0103)	**(0.0098)**	**(0.0102)**	**(0.0091)**	**(0.0089)**	**(0.0089)**
Average (φ_2)	**0.3749**	**-0.0319**	**-0.4726**	**-0.0116**	**-0.0106**	**-0.0111**
	(0.3345)	**(0.1527)**	**(0.1597)**	**(0.0280)**	**(0.0278)**	**(0.0279)**
Female	-0.1499***	-0.1683***	-0.1882***	-0.1677***	-0.1676***	-0.1676***
	(0.0488)	(0.0451)	(0.0472)	(0.0446)	(0.0446)	(0.0446)
Grade	-0.0594*	-0.0910***	-0.1255***	-0.0894***	-0.0892***	-0.0897***
	(0.0335)	(0.0235)	(0.0246)	(0.0205)	(0.0205)	(0.0205)
Math_sc_A	-0.0102	0.0044	0.0209	0.0035	0.0034	0.0041
	(0.0850)	(0.0810)	(0.0847)	(0.0808)	(0.0808)	(0.0808)
Math_sc_B	0.0672	0.0978	0.1317	0.0962	0.0962	0.0970
	(0.0844)	(0.0782)	(0.0817)	(0.0774)	(0.0774)	(0.0774)
Math_sc_C	0.1267	0.1310*	0.1360	0.1310	0.1311	0.1316*
	(0.0830)	(0.0797)	(0.0834)	(0.0797)	(0.0797)	(0.0797)
Math_sc_mis	0.1137	0.1573	0.2049	0.1546	0.1541	0.1545
	(0.1225)	(0.1137)	(0.1189)	(0.1126)	(0.1126)	(0.1126)
Black	0.3261***	0.3027***	0.2771***	0.3033***	0.3040***	0.3038***
	(0.0951)	(0.0899)	(0.0940)	(0.0895)	(0.0895)	(0.0895)
Other races	0.0883	0.0939	0.1003	0.0934	0.0934	0.0940
	(0.1041)	(0.0999)	(0.1045)	(0.0999)	(0.0999)	(0.0999)
Parental education	-0.0016	-0.0051	-0.0090	-0.0049	-0.0049	-0.0049
	(0.0104)	(0.0097)	(0.0101)	(0.0096)	(0.0096)	(0.0096)
School attachment	-0.0374	-0.0308	-0.0232	-0.0314	-0.0314	-0.0309
	(0.0251)	(0.0237)	(0.0248)	(0.0236)	(0.0236)	(0.0236)
Teacher troubles	0.0578***	0.0567***	0.0556***	0.0566***	0.0567***	0.0568***
	(0.0262)	(0.0252)	(0.0263)	(0.0252)	(0.0252)	(0.0252)
Neighborhood quality	-0.0143	0.0089	0.0337	0.0081	0.0080	0.0077
	(0.0356)	(0.0302)	(0.0315)	(0.0289)	(0.0289)	(0.0289)
Family size	-0.0070	-0.0164	-0.0267	-0.0157	-0.0157	-0.0158
	(0.0188)	(0.0168)	(0.0176)	(0.0164)	(0.0164)	(0.0164)
Married parents	-0.0452	0.0009	0.0511	-0.0018	-0.0018	-0.0016
	(0.0690)	(0.0580)	(0.0607)	(0.0555)	(0.0555)	(0.0555)
Self esteem	0.0017	-0.0019	-0.0058	-0.0018	-0.0017	-0.0017
	(0.0231)	(0.0221)	(0.0231)	(0.0220)	(0.0220)	(0.0220)
Parental care	0.0937	0.0546	0.0118	0.0564	0.0564	0.0560
	(0.0908)	(0.0828)	(0.0866)	(0.0816)	(0.0816)	(0.0816)
Network fixed effects	yes	yes	yes	yes	yes	yes
Parental occupation dummies	yes	yes	yes	yes	yes	yes
1st Stage F statistic	**0.870**	**1.641**				
OIR test			***p-value***	**0.698**		

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 6: Testing local aggregate versus local average models for education (GPA)
Hybrid model (19) only - all controls - alternative estimators-
Directed networks

	2SLS-1	2SLS-2	C2SLS	GMM-1	GMM-2	CGMM
Aggregate (φ_1)	**0.0059**	**0.0143*****	**0.0230*****	**0.0147*****	**0.0148*****	**0.0146*****
	(0.0068)	**(0.0052)**	**(0.0058)**	**(0.0045)**	**(0.0045)**	**(0.0045)**
Average (φ_2)	**0.6952****	**0.1401**	**-0.4520****	**0.2447*****	**0.2270*****	**0.2177*****
	(0.3201)	**(0.1523)**	**(0.1706)**	**(0.0340)**	**(0.0338)**	**(0.0339)**
Female	0.2070***	0.2434***	0.2820***	0.2326***	0.2317***	0.2320***
	(0.0624)	(0.0576)	(0.0646)	(0.0565)	(0.0566)	(0.0566)
Grade	0.1156***	0.0446	-0.0311	0.0588**	0.0574**	0.0563**
	(0.0494)	(0.0331)	(0.0370)	(0.0269)	(0.0269)	(0.0270)
Black	-0.2725***	-0.2534***	-0.2332***	-0.2556***	-0.2585***	-0.2585***
	(0.1310)	(0.1261)	(0.1413)	(0.1254)	(0.1255)	(0.1255)
Other races	-0.0081	0.0230	0.0560	0.0156	0.0163	0.0167
	(0.0841)	(0.0797)	(0.0893)	(0.0789)	(0.0790)	(0.0790)
Parental education	0.0323	0.0593**	0.0882***	0.0547**	0.0571**	0.0578**
	(0.0329)	(0.0289)	(0.0324)	(0.0279)	(0.0279)	(0.0279)
School attachment	-0.0998***	-0.0893***	-0.0782***	-0.0915***	-0.0911***	-0.0912***
	(0.0269)	(0.0255)	(0.0285)	(0.0252)	(0.0252)	(0.0252)
Teacher troubles	-0.0643***	-0.0835***	-0.1040***	-0.0801***	-0.0811***	-0.0814***
	(0.0245)	(0.0218)	(0.0244)	(0.0211)	(0.0211)	(0.0211)
Neighborhood quality	0.0128	0.0109	0.0089	0.0107	0.0107	0.0106
	(0.0274)	(0.0264)	(0.0296)	(0.0263)	(0.0263)	(0.0263)
Family size	-0.0083	-0.0040	0.0005	-0.0051	-0.0049	-0.0048
	(0.0259)	(0.0249)	(0.0279)	(0.0247)	(0.0247)	(0.0247)
Married parents	0.1571	0.1120	0.0638	0.1205	0.1169	0.1162
	(0.0935)	(0.0876)	(0.0981)	(0.0863)	(0.0864)	(0.0864)
Self esteem	-0.1031***	-0.1253***	-0.1491***	-0.1216***	-0.1232***	-0.1237***
	(0.0407)	(0.0377)	(0.0423)	(0.0371)	(0.0371)	(0.0371)
Parental care	0.1059	0.1615***	0.2209***	0.1451**	0.1471**	0.1481**
	(0.0780)	(0.0702)	(0.0787)	(0.0683)	(0.0684)	(0.0684)
Network fixed effects	yes	yes	yes	yes	yes	yes
Parental occupation dummies	yes	yes	yes	yes	yes	yes
1st Stage F statistic	**1.127**	**2.886**				
OIR test			***p-value***	**0.540**		

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 7: Testing local aggregate versus local average models for sport activities
Increasing sets of controls- biased corrected optimal GMM
Undirected networks
Hybrid model (19)

	(1)	(2)	(3)	(4)
Aggregate (φ_1)	**0.0153*** **	**0.0152*** **	**0.0148*** **	**0.0143*** **
	(0.0052)	**(0.0052)**	**(0.0052)**	**(0.0052)**
Average (φ_2)	**0.0019**	**0.0024**	**0.0025**	**0.0026**
	(0.0252)	**(0.0252)**	**(0.0252)**	**(0.0251)**
Female	-0.5486***	-0.5499***	-0.5492***	-0.5484***
	(0.0378)	(0.0379)	(0.0378)	(0.0378)
Grade	-0.0617***	-0.0615***	-0.0617***	-0.0610***
	(0.0256)	(0.0256)	(0.0256)	(0.0256)
Math_sc_A	0.0730	0.0715	0.0118	0.0158
	(0.0685)	(0.0686)	(0.0700)	(0.0701)
Math_sc_B	0.0598	0.0559	0.0291	0.0320
	(0.0669)	(0.0670)	(0.0672)	(0.0673)
Math_sc_C	0.0497	0.0471	0.0385	0.0413
	(0.0695)	(0.0695)	(0.0694)	(0.0695)
Math_sc_mis	-0.0119	-0.0099	-0.0272	-0.0184
	(0.1028)	(0.1028)	(0.1027)	(0.1028)
Black	-0.0198	-0.0202	-0.0283	-0.0382
	(0.0881)	(0.0884)	(0.0883)	(0.0885)
Other races	-0.1541*	-0.1599*	-0.1562*	-0.1529*
	(0.0916)	(0.0917)	(0.0915)	(0.0920)
Parental education	0.0078	0.0077	0.0070	0.0070
	(0.0081)	(0.0081)	(0.0081)	(0.0082)
School attachment	0.1304***	0.1302***	0.1236***	0.1245***
	(0.0204)	(0.0205)	(0.0206)	(0.0207)
Teacher troubles	0.0033	0.0042	0.0070	0.0095
	(0.0220)	(0.0220)	(0.0220)	(0.0220)
Neighborhood quality		-0.0105	-0.0053	-0.0016
		(0.0243)	(0.0243)	(0.0246)
Family size		0.0270*	0.0273*	0.0288**
		(0.0142)	(0.0141)	(0.0142)
Married parents		-0.0268	-0.0354	-0.0261
		(0.0458)	(0.0459)	(0.0475)
Self esteem			0.0749	0.0714
			(0.0187)	(0.0188)
Parental care			-0.0394	-0.0374
			(0.0697)	(0.0698)
Network fixed effects	yes	yes	yes	yes
Parental occupation dummies	no	no	no	yes

J test	**Model (12)** Null Hp. (α_1=0)	No local aver.	*t statistic*	**0.105**
	Model (16) Null Hp. (α_2=0)	No local aggr.	*t statistic*	**2.746**

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 8: Testing local aggregate versus local average models for screen activities
Increasing sets of controls- biased corrected optimal GMM
Undirected networks
Hybrid model (19)

	(1)	(2)	(3)	(4)
Aggregate (φ_1)	**-0.0003**	**-0.0003**	**-0.0004**	**-0.0006**
	(0.0070)	**(0.0070)**	**(0.0070)**	**(0.0070)**
Average (φ_2)	**-0.0062**	**-0.0062**	**-0.0061**	**-0.0054**
	(0.0271)	**(0.0271)**	**(0.0271)**	**(0.0271)**
Female	-0.1678***	-0.1660***	-0.1655***	-0.1656***
	(0.0444)	(0.0445)	(0.0445)	(0.0445)
Grade	-0.0559*	-0.0560*	-0.0553*	-0.0579*
	(0.0300)	(0.0300)	(0.0300)	(0.0301)
Math_sc_A	0.0054	0.0037	0.0052	0.0024
	(0.0784)	(0.0785)	(0.0807)	(0.0808)
Math_sc_B	0.0954	0.0954	0.0975	0.0914
	(0.0765)	(0.0766)	(0.0772)	(0.0773)
Math_sc_C	0.1333*	0.1329*	0.1327*	0.1296
	(0.0794)	(0.0795)	(0.0796)	(0.0797)
Math_sc_mis	0.1653	0.1612	0.1608	0.1651
	(0.1119)	(0.1120)	(0.1121)	(0.1122)
Black	0.2179***	0.2224***	0.2225***	0.2272***
	(0.0975)	(0.0981)	(0.0982)	(0.0983)
Other races	0.0591	0.0660	0.0676	0.0875
	(0.1009)	(0.1012)	(0.1012)	(0.1016)
Parental education	-0.0040	-0.0039	-0.0039	-0.0049
	(0.0095)	(0.0095)	(0.0095)	(0.0096)
School attachment	-0.0337	-0.0338	-0.0352	-0.0339
	(0.0233)	(0.0234)	(0.0236)	(0.0236)
Teacher troubles	0.0570***	0.0569***	0.0582***	0.0580***
	(0.0250)	(0.0251)	(0.0252)	(0.0252)
Neighborhood quality		0.0054	0.0060	0.0085
		(0.0284)	(0.0285)	(0.0287)
Family size		-0.0188	-0.0186	-0.0178
		(0.0164)	(0.0164)	(0.0164)
Married parents		0.0324	0.0293	0.0007
		(0.0533)	(0.0536)	(0.0554)
Self esteem			-0.0023	-0.0016
			(0.0219)	(0.0220)
Parental care			0.0578	0.0579
			(0.0816)	(0.0816)
Network fixed effects	yes	yes	yes	yes
Parental occupation dummies	no	no	no	yes

J test	**Model (12)** Null Hp. (α_1=0)	No local aver.	*t statistic*	**0.185**
	Model (16) Null Hp. (α_2=0)	No local aggr.	*t statistic*	**-0.086**

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 9: Testing local aggregate versus local average models for education (GPA)
Increasing sets of controls- biased corrected optimal GMM
Undirected networks
Hybrid model (19)

	(1)	**(2)**	**(3)**	**(4)**
Aggregate (φ_1)	**0.0150***** **(0.0035)**	**0.0146***** **(0.0035)**	**0.0145***** **(0.0035)**	**0.0144***** **(0.0035)**
Average (φ_2)	**0.2167***** **(0.0315)**	**0.2173***** **(0.0315)**	**0.2144***** **(0.0315)**	**0.2135***** **(0.0314)**
Female	0.2198*** (0.0549)	0.2201*** (0.0549)	0.2409*** (0.0552)	0.2416*** (0.0553)
Grade	0.0493 (0.0464)	0.0508 (0.0463)	0.0458 (0.0461)	0.0468 (0.0462)
Black	-0.2092 (0.1326)	-0.1685 (0.1332)	-0.2011 (0.1328)	-0.2310* (0.1333)
Other races	0.0270 (0.0782)	0.0290 (0.0784)	0.0295 (0.0780)	0.0331 (0.0781)
Parental education	0.0625*** (0.0269)	0.0544** (0.0271)	0.0575** (0.0269)	0.0521* (0.0274)
School attachment	-0.1116*** (0.0234)	-0.1040*** (0.0240)	-0.0783*** (0.0248)	-0.0807*** (0.0249)
Teacher troubles	-0.0781*** (0.0210)	-0.0775*** (0.0209)	-0.0735*** (0.0209)	-0.0749*** (0.0209)
Neighborhood quality		-0.0162 (0.0253)	0.0062 (0.0258)	0.0102 (0.0259)
Family size		-0.0103 (0.0244)	-0.0112 (0.0243)	-0.0102 (0.0244)
Married parents		0.1900*** (0.0676)	0.0822 (0.0814)	0.0982 (0.0846)
Self esteem			-0.1116*** (0.0365)	-0.1145*** (0.0366)
Parental care			0.1542*** (0.0669)	0.1533*** (0.0670)
Network fixed effects	yes	yes	yes	yes
Parental occupation dummies	no	no	no	yes
	Model (12)			
J test	Null Hp. (α_1=0)	No local aver.	*t statistic*	**6.796**
	Model (16)			
	Null Hp. (α_2=0)	No local aggr.	*t statistic*	**4.103**

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 10: Testing local aggregate versus local average models for sport activities
Hybrid model (19) only - all controls - alternative estimators-
Undirected networks

	2SLS-1	2SLS-2	C2SLS	GMM-1	GMM-2	CGMM
Aggregate (φ_1)	**0.0122****	**0.0134****	**0.0123****	**0.0160*****	**0.0161*****	**0.0143*****
	(0.0059)	**(0.0056)**	**(0.0068)**	**(0.0052)**	**(0.0052)**	**(0.0052)**
Average (φ_2)	**0.4262**	**0.1160**	**-0.9655**	**0.0151**	**0.0029**	**0.0026**
	(0.2701)	**(0.1739)**	**(0.2109)**	**(0.0252)**	**(0.0251)**	**(0.0251)**
Female	-0.5488***	-0.5490***	-0.5492***	-0.5474***	-0.5486***	-0.5484***
	(0.0393)	(0.0379)	(0.0460)	(0.0378)	(0.0378)	(0.0378)
Grade	-0.0772***	-0.0643***	-0.0212	-0.0601***	-0.0604***	-0.0610***
	(0.0288)	(0.0266)	(0.0323)	(0.0256)	(0.0256)	(0.0256)
Math_sc_A	-0.0126	0.0050	0.0704	0.0153	0.0144	0.0158
	(0.0745)	(0.0710)	(0.0861)	(0.0701)	(0.0701)	(0.0701)
Math_sc_B	0.0005	0.0217	0.0987	0.0318	0.0310	0.0320
	(0.0724)	(0.0685)	(0.0831)	(0.0673)	(0.0673)	(0.0673)
Math_sc_C	0.0105	0.0316	0.1072	0.0408	0.0406	0.0413
	(0.0746)	(0.0707)	(0.0857)	(0.0695)	(0.0695)	(0.0695)
Math_sc_mis	0.0103	-0.0137	-0.0965	-0.0167	-0.0187	-0.0184
	(0.1088)	(0.1039)	(0.1260)	(0.1028)	(0.1028)	(0.1028)
Black	-0.0489	-0.0425	-0.0197	-0.0388	-0.0384	-0.0382
	(0.0921)	(0.0887)	(0.1076)	(0.0885)	(0.0885)	(0.0885)
Other races	-0.2093**	-0.1696*	-0.0271	-0.1536*	-0.1543*	-0.1529*
	(0.1018)	(0.0949)	(0.1150)	(0.0920)	(0.0920)	(0.0920)
Parental education	0.0039	0.0060	0.0134	0.0069	0.0070	0.0070
	(0.0087)	(0.0083)	(0.0100)	(0.0082)	(0.0082)	(0.0082)
School attachment	0.1138***	0.1223***	0.1536***	0.1229***	0.1238***	0.1245***
	(0.0228)	(0.0213)	(0.0258)	(0.0207)	(0.0207)	(0.0207)
Teacher troubles	0.0007	0.0056	0.0234	0.0101	0.0093	0.0095
	(0.0233)	(0.0222)	(0.0270)	(0.0220)	(0.0220)	(0.0220)
Neighborhood quality	0.0103	0.0030	-0.0252	-0.0005	-0.0007	-0.0016
	(0.0264)	(0.0250)	(0.0303)	(0.0246)	(0.0246)	(0.0246)
Family size	0.0244	0.0272*	0.0368**	0.0287**	0.0288**	0.0288**
	(0.0149)	(0.0143)	(0.0173)	(0.0142)	(0.0142)	(0.0142)
Married parents	-0.0263	-0.0251	-0.0223	-0.0258	-0.0257	-0.0261
	(0.0494)	(0.0476)	(0.0578)	(0.0475)	(0.0475)	(0.0475)
Self esteem	0.0633***	0.0690***	0.0898***	0.0712***	0.0711***	0.0714***
	(0.0202)	(0.0191)	(0.0232)	(0.0188)	(0.0188)	(0.0188)
Parental care	-0.0403	-0.0388	-0.0327	-0.0367	-0.0378	-0.0374
	(0.0726)	(0.0700)	(0.0848)	(0.0698)	(0.0698)	(0.0698)
Network fixed effects	yes	yes	yes	yes	yes	yes
Parental occupation dummies	yes	yes	yes	yes	yes	yes
1st Stage F statistic	**1.334**	**1.261**				
OIR test			***p-value***	**0.234**		

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 11: Testing local aggregate versus local average models for screen activities
Hybrid model (19) only - all controls - alternative estimators-
Undirected networks

	2SLS-1	2SLS-2	C2SLS	GMM-1	GMM-2	CGMM
Aggregate (φ_1)	**-0.0011**	**0.0006**	**-0.0030**	**0.0006**	**0.0024**	**-0.0006**
	(0.0080)	**(0.0078)**	**(0.0086)**	**(0.0071)**	**(0.0070)**	**(0.0070)**
Average (φ_1)	**0.1680**	**-0.1723**	**-0.7149**	**0.0017**	**-0.0094**	**-0.0054**
	(0.4455)	**(0.1886)**	**(0.2080)**	**(0.0273)**	**(0.0271)**	**(0.0271)**
Female	-0.1650***	-0.1651***	-0.1644***	-0.1657***	-0.1660***	-0.1656***
	(0.0449)	(0.0448)	(0.0494)	(0.0446)	(0.0446)	(0.0445)
Grade	-0.0561*	-0.0589*	-0.0656**	-0.0573*	-0.0569*	-0.0579*
	(0.0307)	(0.0303)	(0.0334)	(0.0301)	(0.0301)	(0.0301)
Math_sc_A	-0.0017	0.0026	0.0124	0.0016	0.0012	0.0024
	(0.0818)	(0.0812)	(0.0896)	(0.0808)	(0.0808)	(0.0808)
Math_sc_B	0.0728	0.1073	0.1660*	0.0896	0.0902	0.0914
	(0.0905)	(0.0800)	(0.0882)	(0.0773)	(0.0773)	(0.0773)
Math_sc_C	0.1160	0.1395*	0.1793**	0.1283	0.1287	0.1296
	(0.0863)	(0.0812)	(0.0895)	(0.0797)	(0.0797)	(0.0797)
Math_sc_mis	0.1374	0.1912	0.2805**	0.1633	0.1641	0.1651
	(0.1337)	(0.1166)	(0.1286)	(0.1122)	(0.1122)	(0.1122)
Black	0.2282***	0.2262***	0.2254***	0.2268***	0.2261***	0.2272***
	(0.0992)	(0.0988)	(0.1090)	(0.0983)	(0.0983)	(0.0983)
Other races	0.0829	0.0891	0.1038	0.0863	0.0854	0.0875
	(0.1030)	(0.1022)	(0.1127)	(0.1017)	(0.1016)	(0.1016)
Parental education	-0.0044	-0.0055	-0.0073	-0.0049	-0.0049	-0.0049
	(0.0098)	(0.0096)	(0.0106)	(0.0096)	(0.0096)	(0.0096)
School attachment	-0.0379	-0.0297	-0.0143	-0.0345	-0.0347	-0.0339
	(0.0264)	(0.0242)	(0.0267)	(0.0236)	(0.0236)	(0.0236)
Teacher troubles	0.0568***	0.0585***	0.0624***	0.0577***	0.0575***	0.0580***
	(0.0256)	(0.0253)	(0.0279)	(0.0252)	(0.0252)	(0.0252)
Neighborhood quality	-0.0003	0.0165	0.0422	0.0086	0.0092	0.0085
	(0.0362)	(0.0303)	(0.0334)	(0.0287)	(0.0287)	(0.0287)
Family size	-0.0136	-0.0221	-0.0362***	-0.0175	-0.0176	-0.0178
	(0.0200)	(0.0171)	(0.0189)	(0.0164)	(0.0164)	(0.0164)
Married parents	-0.0099	0.0121	0.0474	0.0004	0.0008	0.0007
	(0.0628)	(0.0569)	(0.0628)	(0.0554)	(0.0554)	(0.0554)
Self esteem	0.0001	-0.0030	-0.0078	-0.0017	-0.0017	-0.0016
	(0.0226)	(0.0222)	(0.0245)	(0.0220)	(0.0220)	(0.0220)
Parental care	0.0730	0.0427	-0.0062	0.0586	0.0579	0.0579
	(0.0914)	(0.0836)	(0.0922)	(0.0816)	(0.0815)	(0.0815)
Network fixed effects	yes	yes	yes	yes	yes	yes
Parental occupation dummies	yes	yes	yes	yes	yes	yes
1st Stage F statistic	**0.473**	**1.093**				
OIR test			***p-value***	**0.804**		

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.

Table 12: Testing local aggregate versus local average models for education (GPA)
Hybrid model (19) only - all controls - alternative estimators-
Undirected networks

	2SLS-1	2SLS-2	C2SLS	GMM-1	GMM-2	CGMM
Aggregate (φ_1)	**0.0107*** **	**0.0104*** **	**0.0209*** **	**0.0145*** **	**0.0149*** **	**0.0144*** **
	(0.0044)	**(0.0040)**	**(0.0049)**	**(0.0035)**	**(0.0035)**	**(0.0035)**
Average (φ_2)	**0.6099*** **	**0.6449*** **	**-0.6802*** **	**0.2480*** **	**0.2250*** **	**0.2135*** **
	(0.2790)	**(0.1941)**	**(0.2376)**	**(0.0311)**	**(0.0312)**	**(0.0314)**
Female	0.2217***	0.2197***	0.2966***	0.2412***	0.2409***	0.2416***
	(0.0581)	(0.0572)	(0.0700)	(0.0552)	(0.0553)	(0.0553)
Grade	0.0603	0.0615	0.0167*	0.0471*	0.0467*	0.0468*
	(0.0476)	(0.0473)	(0.0579)	(0.0461)	(0.0462)	(0.0462)
Black	-0.2136	-0.2120	-0.2796	-0.2297	-0.2293	-0.2310
	(0.1353)	(0.1355)	(0.1659)	(0.1330)	(0.1332)	(0.1333)
Other races	0.0576	0.0600	-0.0351	0.0344	0.0344	0.0331
	(0.0813)	(0.0804)	(0.0984)	(0.0779)	(0.0781)	(0.0781)
Parental education	0.0295	0.0276	0.1026***	0.0495*	0.0510*	0.0521*
	(0.0316)	(0.0297)	(0.0364)	(0.0273)	(0.0273)	(0.0274)
School attachment	-0.0852***	-0.0857***	-0.0686***	-0.0800***	-0.0802***	-0.0807***
	(0.0255)	(0.0254)	(0.0311)	(0.0249)	(0.0249)	(0.0249)
Teacher troubles	-0.0586***	-0.0572***	-0.1111***	-0.0735***	-0.0743***	-0.0749***
	(0.0239)	(0.0225)	(0.0276)	(0.0208)	(0.0208)	(0.0209)
Neighborhood quality	0.0132	0.0135	0.0024	0.0103	0.0105	0.0102
	(0.0262)	(0.0263)	(0.0322)	(0.0258)	(0.0258)	(0.0259)
Family size	-0.0156	-0.0161	0.0033	-0.0106	-0.0104	-0.0102
	(0.0249)	(0.0249)	(0.0304)	(0.0243)	(0.0243)	(0.0244)
Married parents	0.1114	0.1127	0.0635	0.0979	0.0986	0.0982
	(0.0860)	(0.0861)	(0.1053)	(0.0844)	(0.0845)	(0.0846)
Self esteem	-0.0942***	-0.0924***	-0.1613***	-0.1123***	-0.1137***	-0.1145***
	(0.0396)	(0.0384)	(0.0470)	(0.0365)	(0.0365)	(0.0366)
Parental care	0.1268***	0.1241***	0.2299***	0.1528***	0.1522***	0.1533***
	(0.0712)	(0.0697)	(0.0853)	(0.0669)	(0.0670)	(0.0670)
Network fixed effects	yes	yes	yes	yes	yes	Yes
Parental occupation dummies	yes	yes	yes	yes	yes	Yes
1st Stage F statistic	**1.381**	**1.730**				
OIR test			***p-value***	**0.862**		

Note: Coefficients marked with one (two) [three] stars are significant at 10 (5) [1] percent level.