School Performance and Network Effects among Classmates

Emma Zavarrone

Dept. of Consumer, Behavior and Communication “G. Fabris”,
Iulm University, Milan, Italy
emma.zavarrone@iulm.it

Agnese Vitali

Carlo F. Dondena Centre for Research on Social Dynamics,
Bocconi University, Milan, Italy
agnese.vitali@unibocconi.it

Draft paper prepared for the
Annual Meeting of the Population Association of America, San Francisco 2012
Please do not cite or quote without the authors’ permission

Abstract

This paper contributes to the literature on the influence of network structure on performance of university students. The paper compares the marginal effect on performance of five different aspects of communication among classmates, represented by five network structures: study, friendship, exchange of information, exchange of lecture notes, and trust networks. Using a set of different weight matrix specifications based on adjacency, outdegree, resistance, reciprocity and homophily criteria, we show that members of a communication network tend to have similar performances. Results, however, are not always robust to different weight matrix specifications. We show that conclusions about the network effect on students’ performance might change dramatically when either different neighbouring structures or different weight specifications are considered. We conclude that in the study of social interaction processes the network structure need to be translated into a meaningful and theory-guided weight matrix specification.

Keywords: Social Networks, Performance, Spatial Weight Matrix

Introduction

Students' ability and motivation have long been considered the main drivers of school performance. Contextual and institutional characteristics have been shown to play a role in determining students' performance. Characteristics of the neighborhood of residence, the school, the classroom and the teacher all have an effect on students' performance. Similarly, socio-economic characteristics of the family of origin, e.g. parental resources and educational achievement, show an association with students' school performance.

A number of studies in Social Network Analysis, Economics, Sociology and Psychology have recently focused on the influence of social interaction with peers on students' school performance and agree that relational networks among peers play an important role in students' learning process and school performance (Hanushek et al., 2003; Ortiz et al., 2004; Cook et al., 2007; Ding at al., 2007; Eggens et al., 2008; Calvó-Armengol et al., 2009; de Klepper et al., 2010).

Students' learning process can be influenced by formal and informal interactions which in turn might take place inside or outside the classroom (Hommes et al., 2012). However, it is not clear which form of social interaction among classmates is mainly associated with the students' academic performance. Defining, measuring and disentangling informal relations among classmates is in fact not straightforward as peers within the same classroom might interact in a number of different ways. While the most widely studied type of relation refers to friendship ties, there are other types of interactions which might matter for the learning process. Trust, collaboration, altruism, but also envy, competition and opportunism are different types of informal interactions which might be associated with academic performance.

Social network analysis is definitely the most appropriate tool for modeling interactions among classmates, as it allows to model social interaction as a network whose nodes are the students in a given classroom and ties refer to their web of relations. On the other hand, spatial econometric techniques offer the opportunity to test whether social interactions affect students' performance. Spatial models in fact allow testing whether students who belong to the same network tend to have similar school performance or not. In spatial models the social interaction process is modeled by the means of a *spatial weight matrix* which incorporates the neighbouring structure of the network (i.e. who is neighbour to whom). Since in the literature on Social Network Analysis there is no consensus on how to model the social interaction process –i.e. generally there are several different relational networks which might be associated with any agents' outcome–, and since the weight matrix represents the interaction process, it follows that there is no consensus on how to build the weight matrix either.

Following Leenders (2002), we acknowledge the need to translate the network structure into a meaningful and theory-guided choice of weight matrix. Using a set of different weight matrices we show that, while in some cases results are robust to different choices of weight matrices, in some other cases conclusions can change according to the chosen weight matrix.

Despite the large literature on the effect of different types of social interaction within the classroom on students' school performance, to our knowledge there are very few attempts to compare directly the effect of different communication networks on performance. In this paper, we hypothesize that, beside friendship relations, classmates can be involved in four additional communication networks which might affect their academic performance. We consider the friendship network, three types of interactions based on collaborative learning, i.e. study relations among peers, exchange of general information about the course contents and exchange of lecture notes, and a network of trust relations.

The contribution of this paper is twofold. First, it adds to the literature on the influence of network relations on students' school performance by comparing the effects of five different network structures: study, friendship, exchange of information, exchange of lecture notes, and trust. Second, the paper provides a general application of Leenders (2002) claim that "at the end of the day, any autocorrelation model is useless when W is not specified with explicit attention and care".

School Performance and Social Interaction

Existing literature usually focuses on one main type of social interaction among classmates, i.e. friendship. Mayer et al. (2008), analyze networks of students on university campuses basing on their friendship ties as measured by Facebook.com and find that the average friends' grade point average (GPA) is strongly associated with own GPA. Other studies confirm that both the size (Baldwin et al., 1997) and density of the social network (Rizzuto et al., 2009) are important predictor of students' academic performance. Similarly, Lubbens et al. (2006) show that students with high peer acceptance (measured on the basis of the number of ties a student receives from his/her classmates) had lower probabilities to retain a grade or to move downward in the track system from one year to the next.

The interest raised by friendship ties in social interaction processes in network analysis is justified by the fact that relations among friends are generally more intimate and long lasting with respect to relations with other peers; hence friendship ties are expected to play the most influential role in all outcomes, in general.

Nonetheless, many authors found that, similarly to friendship relations, other informal communication networks also have a positive effect on students' performance. Jeong et al. (2004), for example, show that collaborative interaction results in an increased "knowledge convergence". Hommes et al. (2012) consider two types of communication networks, giving and receiving information related to the subject studied, and they show that being a central actor in any of the two networks yields a positive effect on learning. Also, Karabenick (2003) shows that help-seeker students have better school performance than help-seeker avoidant students. Baldwin et al. (1997) instead find that centrality in friendship, communication and adversarial networks among MBA students affect their performance; in particular they also consider networks of students identifying in their peers a source of school-related advice or problem. Berten et al. (2011) find a network effect for both best friends and cohesive relations on school deviant behaviours. Finally, Guryan et al. (2008) use data on interactions among classmates collected through handheld computers during math and science lessons to investigate the mechanisms underlying peer influence.

Hypotheses

We expect that informal social interactions among students have an effect on their academic performance, as measured by their grade in Statistics. Beside the commonly studied friendship network, we consider four additional types of interactions among classmates: three types of interactions are based on collaborative learning, i.e. study relations among peers, exchange of general information about the course contents and exchange of lecture notes, and the last one is a network of trust relations among dyads of students. Survey items used to measure the five network are introduced in the following sections.

With the aim of investigating the influence of the five network effects on students' performance, we test the following assumptions:

H1: We test whether students' performance is influenced by the performance of the subgroup of alters with which he/she has a relationship of one of the five kinds and which of these relationships has the higher marginal effect.

H2: We investigate whether, after controlling for individual demographic characteristics, students' performance is influenced by unobservable characteristics common to the networks' structure (i.e., we test whether unobservable factors have an effect on the ego's performance, but also on the performance of his/her alters).

Data

We developed an ad-hoc survey in which students in a given class are asked to detail the structure of five different networks to which they belong (or do not belong). Respondents are master students in the age range 22-23, who attended the Statistics course during the 2009/2010 school year (their first year of a two-year master degree) at Iulm University in Milan, Italy. The final sample is constituted of 41 students.

Students are asked whether they have classmates with which they get together outside the university environment, and if they do, we ask to identify them and consider there classmates as members of the ego's friendship network. Students are then asked whether, in order to prepare the Statistics exam, they studied on their own. If they did not, they are asked to identify the classmates with whom they studied. We consider these peers as members of the ego's study network. In order to identify the exchange of information network and exchange of notes network, we look at classmates whom the ego considers a reliable source of information for what concerns the Statistics course and with which he/she exchanges/compares his/her notes, respectively. Finally, relying on Figure A in the Appendix, the questionnaire asks to identify the peer whom the ego perceives as the central subject among his/her classmates ("You are the person located in the bottom part of this picture. Could you specify, among your classmates, the initials of the person in the upper left of the picture?"). Due to the way the question is asked, the trust network constitutes dyads of relations among classmates. The structure of the five networks we consider is such that relationships do not necessarily need to be reciprocal. Reciprocity will be explicitly considered as one particular criterion for measuring the strength of the interaction through the specification of the weight matrix.

The five communication networks among classmates that we consider are graphically represented in Figure 1.

Our dependent variable, the student' performance, is measured by the grade obtained in the Statistics exam. Grades are expressed in thirtieths (minimum for sufficiency is 18). In addition to the structure of the five networks discussed above, the survey also collects information on the students' field of education during their bachelor studies, whether their university career took place in the same university in which they are surveyed and if this is not the case, in which university they took their bachelor. Further, the students are asked whether, during their university studies, they ever took a Statistics class, which is used as a proxy for students' previous ability.

Table 1 reports descriptive statistics of the variables we employ in our regression models. It is not surprising that male students represent only a minority (34%) at Iulm university (it is well known that a gender difference exists when the field of study is concerned; in particular, women are more often found in humanistic subjects). Almost half (48%) of the students attended their Bachelor in a different university; 56% of the students have received some training in Statistics during their university career (previous ability). Table 1 and Figure 1 give us some information on five networks: presence of cliques (cohesive subgroups) in friendship, study and information networks, average number of ties in these networks equal to 2 in other words each student shares two peers, local centralization not wide.

Figure 1: Structure of five networks of communication among classmates (friendship, study, exchange of information, exchange of lecture notes, trust network)

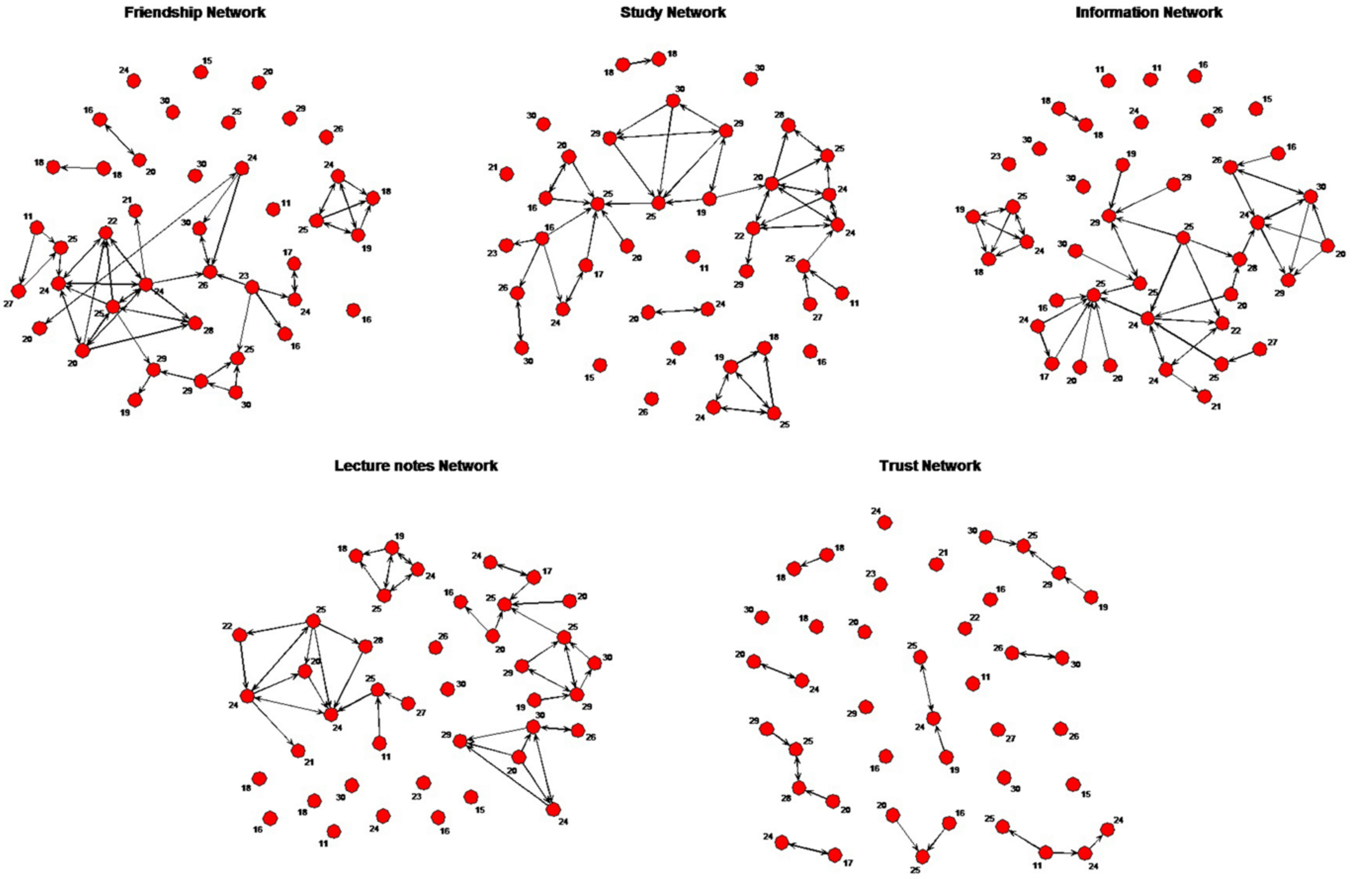


Table 1: Descriptive statistics

Variable	Mean/Pro	Std. Dev.
Final grade in Statistics	22	5.2
Sex (prop. of men)	34%	-
BSc in different university	48%	-
Previous ability	56%	-

Table 2: Network statistics

	Friendship	Study	Notes	Information	Trust
Mean degree	2.289	2.683	2.341	2.293	1.366
Std. Dev. Degree	2.728	2.207	2.1749	2.159	1.089
Density	0.035	0.035	0.029	0.029	0.017
Degree Centralization	0.022	0.107	0.069	0.062	0.035

Methods

In order to test the assumptions, we make use of spatial econometrics techniques. *H1* and *H2* are tested using a spatial lag and a spatial error model, respectively (Anselin, 1988). These two modeling approaches differ in the way the spatial influence process operates within the network and they might induce very different patterns of dependence among the observations (Aselin, 2002). Anselin et al. (2008, p. 6) explain that the "spatial lag model is typically considered as the formal specification for the equilibrium outcome of a spatial or social interaction process, in which the value of the dependent variable for one agent is jointly determined with that of the neighboring agents". The spatial error model instead "does not require a theoretical model of spatial/social integration, but, instead, is a special case of a nonspherical error covariance matrix" (Anselin et al., 2008, p. 8) and such that dependence in the dependent variable is the result of a spatial clustering of unobserved independent variables omitted from the model. A model specification including both a spatial lag and a spatial error term is possible and it is advisable when the assumptions underlying the two models are thought of being satisfied.

The spatial lag model includes a *spatial lag* of the dependent variable as an independent variable which allows the ego's school performance to depend on the performance of his alters. The spatial lag model can be formalized as follows:

$$y = \rho W y + \boldsymbol{X}\beta + \varepsilon, \quad \varepsilon \sim N(0,\sigma_\varepsilon^2 I).$$

The coefficient ρ is called the *network effect* (Doreian, 1989; Leenders, 2002), *spatial lag* (Cliff et al., 1973) or *spatial autocorrelation coefficient* (Anselin, 1988). It measures the spatial autocorrelation in the dependent variable, i.e. the correlation between the ego's performance and the performance of his/her alters. If the spatial autocorrelation coefficient is positive, there is evidence of spatial autocorrelation in students' performance i.e., there is confirmation of *H1*.

The spatial error model, instead, includes a spatially autoregressive error term which captures commonalities in unobserved factors among members of the same network and can be written as follows:

$$y = \boldsymbol{X}\beta + \varepsilon,$$

$$\varepsilon = \rho W \varepsilon + v, \qquad v \sim N(0,\sigma_v^2 I)$$

In this setting, ρ measures the spatial autocorrelation in the error term, i.e. the network disturbance (Doreian, 1989; Leenders, 2002). If ρ is positively significant we interpret that there are common unobserved factors influencing all members of the same network, therefore confirming *H2*.

In general, the conceptual framework of reference should guide the researcher in the choice between the spatial lag or error model, i.e. the two main models which are able to incorporate spatial dependence. There are also diagnostic tests which tests the validity of one model over the other, the most used is the Lagrange Multiplier Statistic.

In both the spatial lag and spatial error models the influence process within the network operates through a pre-defined, user-specified weight matrix (*W*). This matrix "selects" neighbours and indicates how important each neighbour is, assigning a weight to each tie (i.e. relation) between each couple of nodes (i.e. students) in the network. Two students are considered to be "neighbours" if they belong to the same network, in which case the relative entry w_{ij} in the *W* matrix is non-null. Weights, however, can be assigned in a variety of different ways. We discuss this issue in detail in the next section.

Other model assumptions require the spatial autoregressive coefficient to be bounded in absolute value (i.e. $|\rho|<1$), ε_i to be independently and identically normally distributed with zero mean and variance to be estimated. The models are estimated via Maximum Likelihood using the spdep library in R (Bivand, 2010).

As independent variables (***X***) we consider gender (equal to one if the student is male and 0 if female), a dummy equal to one if the student took his/her bachelor degree in a different university with respect to the university in which he/she is studying for the master degree, and a dummy equal to one if the student ever took a Statistics exam in his academic career.

We assume that students who ever took at least one Statistics class in their curriculum will obtain a higher grade in the Statistics exam (our dependent variable) hence we consider past Statistics exams as a proxy for previous ability.

The selection of the weight matrix

The weight matrix W represents the influence process within the network. This matrix informs on two distinct features of the networks: the neighbouring structure (which peers are “neighbours” to the ego) and the strength of the neighbouring relation, operationalized through a weight specification (what weight is assigned to each neighbouring relation). Identifying these two network features is of paramount importance for the study of social influence processes.

The neighbouring structure can be thought of as being described through an adjacency matrix, A, whose entries, a_{ij}, equal one if student i is linked to student j (i.e. if they belong to the same network), and zero if no relation exists between the i and j.

The strength of the neighbouring relation, instead, depends on the specification of the weight matrix, i.e. on how its entries w_{ij} are specified. Many possible specifications are available. For instance, each pair of students belonging to a given network can be given the same weight, or some relations might be evaluated more important than others and thus assigned more weight. Also, weights can be binary, row-standardized, column-standardized, or assigned in different user-specified ways. For an extensive review on the selection of the neighbouring and weighting schemes see Anselin (2002), Leenders (2002) and Chi et al. (2008).

Given the abundance of different ways in which the researcher can specify a weight matrix, there is no consensus on which is the optimal choice in the literature on Social Network Analysis. Leenders (2002) acknowledges the importance of specifying appropriate weight matrices in the study of social influence models.

In the remainder of the paper we will provide an empirical application of how important it is to pay attention to the specification of the weight matrix. We show that conclusions are not always robust to different weight matrix specifications. We also show that attention needs to be paid to the choice of the network structure underlying the social interaction process under study.

Accordingly, for each of the five network structures that we consider (friendship, study, information contents, lecture notes and trust networks), we employ different neighbouring structures and define different weight specifications. By comparing the magnitude of the estimated spatial autocorrelation coefficients, we can draw conclusion on which network membership is more important for students in terms of school performance.

The weighting schemes we chose to model classmates’ influence on student performance are the five criteria summarized below. In total we employ seven different weight structures, and five different network structures, for a total of 35 different models estimated. In all cases it is assumed that a student cannot be its own neighbour i.e., w_{ii}=0.

1. Cohesion Criterion

Cohesion is the most widely used measurement for modeling interaction processes. According to this criterion, the ego is assumed to be influenced by adjacent agents (Marsden et al., 1996). The idea is that students who belongs to the same network show a more similar academic performance than students who belong to different networks. Ties need not be reciprocal in the sense that student

i might report a tie with student j but the contrary needs not be true, hence the resulting weight matrix needs not be symmetric. We will model cohesion according to three alternative specifications: simple adjacency, adjacency with row normalization and resistance.

1.1 Cohesion based on the Adjacency Criterion

The first weight structure is the most widely used specification, based on the adjacency criterion according to which each entry of the spatial weight matrix W, w_{ij}, equals to the relative entry of the adjacency matrix, a_{ij}. The same weight is assigned to each member of a given network, irrespectively of the network's size. It follows that, according to this weight scheme, every existing tie between students i and j will be assigned weight equal to one, while if students i and j are not tied, they will be assigned weight zero. Formally, the general entry of the weight matrix will be specified as follows:

$$w_{ij} = a_{ij} = \begin{cases} 1 & \text{if student } i \text{ is linked to student } j \\ 0 & \text{otherwise.} \end{cases}$$

1.2. Cohesion with Row Normalization of Adjacency Criterion (Outdegree)

If the adjacency matrix is row standardized, the a_{ij} entry of the adjacency matrix is divided by the row total, $a_{.i}$. and in this case, weights become proportional to the number of people belonging to the given network, hence, the difference with respect to the previous weight scheme, lies in the fact that now the size of the network is taken into account. In other words, the same weight is attached to every outgoing tie of student i and this weight is proportional to the student's outdegree, i.e. to the average number of agents with whom student i self-reports to interact. This type of network models the influence that student i receives from his or her alters and such influence decreases as the size of the network increases, i.e. as the number of reported ties increases. Weights can be operationalized as follows:

$$w_{ij} = \frac{a_{ij}}{a_{i.}}.$$

1.3. Cohesion with Resistance Criterion

The cohesion criterion with resistance is a modification of the cohesion criterion with outdegree which take into account the ego's "resistance" to influences of his/her alters (French, 1956). The own influence is operationalized by adding a one to the denominator of the weights expressed by the outdegree criterion as follows:

$$w_{ij} = \frac{a_{ij}}{a_{i.} + 1}.$$

If a student reports to have, say, two friends, the weight assigned to each of his two friends will be 1 according to the adjacency criterion, it will be 1/2 according to the outdegree criterion and it will equal to 1/(2+1)=1/3 according to the resistance criterion.

2. Reciprocity Criterion

As opposed to the cohesion criterion where ties can be unilateral, according to the reciprocity criterion two peers are neighbours if their relation is reciprocal. Because of the existence of a mutual versus a unilateral tie, connections based on the reciprocity criterion are considered to be the appropriate measure of friendship networks (Lubbers et al., 2006). For example, in the friendship network, student *i* and *j* will be considered neighbours according to the reciprocity criterion if *i* identify *j* as a member of his/her friend network and the same is true for *j* with respct to *i* (Stockman, 2004). The general entry of the weight matrix will therefore be defined as follows:

$$w_{ij} = \begin{cases} 1 \;\; \text{if student } i \text{ is linked to student } j \text{ and student } j \text{ is linked to student } i \\ 0 \text{ otherwise.} \end{cases}$$

3. Homophily Criterion

According to the homophily criterion, two peers are neighbours if they are similar with respect to some specified characteristics (Stockman, 2004). Similarity among peers may result from the choice of networking with more similar peers (selection) or it may originate because peers influence each other (influence). A study on selection and influence effects on similarity in friendship network can be found, e.g., in de Klepper et al. (2010) and Kandel (1978). We try three different characteristics, basing on the three independent variables we use in our models, namely gender (two students are neighbours if they are both female, or both male), university of bachelor degree (two students are neighbours if they both studied in the same university in which they are enrolled for their master or if both took their master in another university, irrespectively of which university) and ability in Statistics (two students are neighbours if they both studied Statics before or if they both did not):

$$w_{ij} = \begin{cases} 1 \;\; \text{if students } i \text{ and } j \text{ share the same characteristics} \\ 0 \text{ otherwise.} \end{cases}$$

Results

Table 3 presents results of the 35 spatial lag regression models, each being a combination of a weight structure (adiacency, outdegree, resistance, reciprocity and homophily with respect to sex, same bachelor and previous ability), and a network structure (friendship, study, exchange of information, exchange of lecture notes, trust network). On the rows of Table 3 there are the five network structures while on the columns there are the seven weighting schemes, differing in terms of the specification of the weight matrix according to the criteria described in the previous section.

From Table 3 it is evident that the estimated spatial autocorrelation coefficient is positive and significantly different from zero in most of the models. The estimated spatial autocorrelation coefficient ρ has to be interpreted in terms of network effect on individual school performance. A positive ρ indicates that there exists a positive correlation between the ego's performance and the performance of the classmates with whom he/she has a relation of one of the five kinds, irrespectively of the weighting scheme. We conclude that that there exists a network effect on performance among classmates. In other words, students' performance is influenced by the performance of the subgroup of alters with which he/she interacts according to any of the five relationship, hence we find evidence in support of *H1*.

However, the magnitude of the spatial autocorrelation coefficient changes considerably according to the five network structures (i.e. along the row of Table 3) as well as according to the specification of the weight matrix chosen for modeling the network interaction process (i.e. along the columns of Table 3). The comparison between the estimated spatial autoregressive coefficients ρ can be better appreciated in Figure 2, where the five network structures are placed on the x-axis, the magnitude of the ρ coefficients on the y-axis, and the bars represent each of the seven weighting schemes used.

Moving along the columns of Table 3, it emerges that the specification of the weight matrix which yields the highest estimates for the spatial autocorrelation coefficient across the five different network structures is the resistance criterion. On the basis of the adjacency criterion, ρ results statistically significant in all the five network structures considered. However, its magnitude changes a great deal across the networks structures, varying from 9% in the friend network to 22% in the trust network. When the adjacency criterion is modified to take into account the number of outgoing ties (outdegree), the estimated ρ result higher, ranging from 29% in the trust network to 48% in the exchange of information network. Estimated coefficients increase even further when the resistance criterion is used to construct the weight matrix, from 53% in the trust network to 73% in the exchange of information network. When we choose the reciprocity criterion to specify the weighting scheme, ρ results statistically significant only in the study network. For what concerns homophily, we find that there is an autocorrelation of 13% and 24% between the ego's performance and the performance of peers of the same gender who are friends and who interact by exchanging information with respect to the course content, respectively. Being friend, studying and trusting classmates with whom the ego shares the same bachelor experience yields a correlation with the ego's performance equal to 10%, 13% and 23%, respectively. Finally, sharing previous ability in Statistics does not result to influence the students' school performance as the spatial autoregressive coefficient is not statistically significant in any of the five network structures considered. The network based on exchange of information regarding the course content exerts the most powerful effect on students' performance according to the outdegree, resistance and homophily with respect to gender criteria. According to homophily with respect to bachelor and previous ability and to the adjacency criterion, instead, it is the trust network which exerts the most powerful effect on students' performance. Finally, according to reciprocity, the highest effect on school performance is found in the study network.

Moving along the rows of Table 3 we see that in the case of the friendship network, the estimated spatial autocorrelation coefficient ρ passes from 9% in the model with the weight matrix specified according to the adjacency criterion to 54% when the resistance criterion is chosen, while it is not statistically significant according to reciprocity and homophily with respect to previous ability. Seemingly, in the case of the information network, ρ passes from 16% to 73% when the weight matrix is specified according to the adjacency and resistance criterion, respectively, while it is not statistically significant according to reciprocity and homophily with respect to previous ability and bachelor. Similar variations in the estimated ρ are observed also when the network structure is modelled on the basis of trust and on the exchange of lecture notes. In the case of the study network ρ is robustly estimated at 13% according to the adjacency, reciprocity and homophily with respect to bachelor criteria, while its estimate is much higher according to the outdegree (43%) and resistance (66%) criteria, and it is not statistically significant according to the remaining criteria.

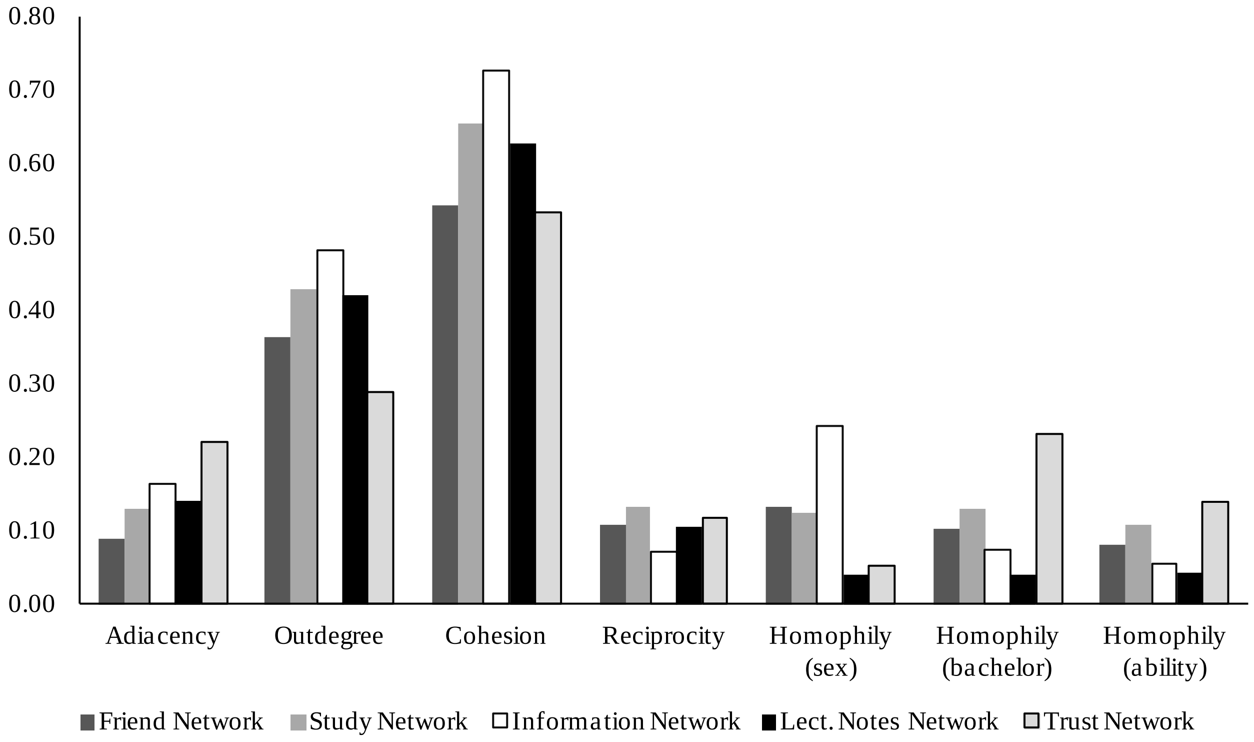
The spatial error model is always rejected with respect to the spatial lag model according to the Lagrange Multiplier Test. Further, when we fit a spatial error model, we do not find evidence of any significant spatial dependence among members of the same network (results not shown). Hence, we do not find evidence of *H2*. In other words, the spatial error model is always rejected in favour of spatial lag model.

Table 3: Results from the spatial lag regression model. Estimated spatial autoregressive coefficients (ρ) from spatial lag models specified according to five different network structures (rows) and seven different weighting schemes (columns). Models are estimated controlling for gender, student has enrolled from a different university and a proxy for previous Statistics ability

	Cohesion									**Reciprocity**			**Homophily**								
	Adiacency			**Outdegree**			**Resistance**						**Gender**			**Bachelor**			**Prev. ability**		
	ρ		AIC	P		AIC	ρ		AIC	ρ		AIC	ρ		AIC	ρ		AIC	ρ		AIC
Friend Network	0.09	**	328.30	0.36	***	325.70	0.54	***	326.11	0.11		330.39	0.13	**	329.81	0.10	*	330.59	0.08		331.77
Study Network	0.13	***	326.19	0.43	***	320.20	0.66	***	321.22	0.13	*	330.01	0.12		331.25	0.13	*	330.34	0.11		331.28
Information Network	0.16	***	325.55	0.48	***	316.98	0.73	***	320.18	0.07		332.21	0.24	**	329.06	0.08		331.56	0.05		332.10
Lecture Notes Network	0.14	***	326.17	0.42	***	321.87	0.63	***	323.38	0.11		331.52	0.04		332.36	0.04		332.36	0.04		332.33
Trust Network	0.22	**	329.34	0.29	***	327.62	0.53	**	328.21	0.12		332.10	0.05		332.55	0.23	*	330.02	0.14		331.82

p-values: ***<0.01; **<0.05; *<0.1, AIC: Akaike Information Criterion.

Figure 1: Comparison of estimated spatial autocorrelation coefficient (y-axis) from spatial lag models specified according to seven different weighting schemes (x-axis) and five network structures (bars)



Conclusions

This paper provided an empirical application of Leenders (2002) acknowledgment of how important the specification of the weight matrix is for drawing conclusions on a social interaction process.

Using data on communication within a classroom of university students, we showed how conclusions about the network effect on students' performance might change dramatically with the choice of both different neighbouring structures and different weight specifications. We considered five different dimensions of communication, namely, getting out together outside the university environment (friend network), studying together, exchanging information with respect to the course content, exchanging lecture notes, and trusting someone. Further, we considered a set of seven different, equally theoretically-grounded weight matrices: three weighting schemes based on the cohesion criterion, i.e. adiacency, outdegree, and resistance, one based on the reciprocity criterion, and three based on homophily with respect to sex, same bachelor and previous ability.

First, we showed that members of any of the five communication networks tend to have similar performances. Therefore we find confirmation of the fact that informal relations within the classroom play a role in influencing students' academic performance.

Second, we showed that while in some cases results are robust to different specifications of the weight matrix, in the majority of cases parameter estimates –hence conclusions– based on autocorrelation models can change according to the chosen specification of the weight matrix. It follows that if the interest is on the magnitude of the estimated ρ coefficient estimated by the spatial econometric model, then ultra care need to be paid in the choice of the weight matrix which is appropriate for the process studied.

The double choice of how to specify the neighbouring structure and weighting scheme to be used in the construction of the weight matrix remains however arbitrary and subjective, and therefore criticisable in any spatial model. For this reason, the network structure needs to be translated into a meaningful and theory-guided choice of weight matrix. In other words, it is advisable to specify the weight matrix on the basis of formal theoretical assumptions regarding the model of social interaction under study.

Acknowledgements

Paper prepared within the project *Consequences of Demographic Change* (CODEC), funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) ERC Grant agreement n° 201194.

References

Anselin, L. (2002). Under the hood. Issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, 27(3): 247–267.

Anselin, L. (1988). *Spatial econometrics: methods and models*. Dordrecht: Kluwer.

Baldwin, T.T., Bedell, M.D. and Johnson, J.L. (1997).The Social Fabric of a Team-Based M.B.A. Program: Network Effects on Student Satisfaction and Performance. *The Academy of Management Journal*, 40(6): 1369–1397.

Berten, H. and Van Rossem, R. (2011). Mechanisms of Peer Influence Among Adolescents: Cohesion Versus Structural Equivalence. *Sociological Perspectives*, 54(2): 183–204.

Bivand, R., 2010. Package spdep in R.

Butts, C. T., 2008. Social Network Analysis with sna. *Journal Of Statistical Software*, 24 (6).

Calvó-Armengol, A., Patacchini, E. and Zenou, Y. (2009). Peer effects and social networks in education. *Review of Economic Studies*, 76(4): 1239–1267.

Chi, G. and Zhu, J. (2008). Spatial Regression Models for Demographic Analysis. *Population Research and Policy Review*, 27(1): 17–42.

Cliff, A.D. and Ord, J.K. (1973). *Spatial Autocorrelation*. London: Pion Limited.

Cook, T.D., Deng, Y., and Morgano, E. (2007). Friendship influences during early adolescence: The special role of friends' grade point average. *Journal of Research on Adolescence*, 17(2): 325–356.

De Klepper, M., Sleebos, E., van de Bunt, G. and Agneessens, F. (2010). Similarity in friendship networks: Selection or influence? The effect of constraining contexts and non-visible individual attributes. *Social Networks*, 32(1): 82–90.

Ding,W. and Steven, F.L. (2007). Do Peers Affect Student Achievement in China's Secondary Schools? *Review of Economic and Statistics*, 89(2): 300–312.

Doreian, P. (1989). Models of network effects on social actors. In: Freeman, L.C., White, D.R., Romney, K. (Eds.), *Research Methods in Social Analysis*. Fairfax: George Mason University Press.

Eggens, L., van der Werf, M.P.C. and Bosker, R.J. (2008). The influence of personal networks and social support on study attainment of students in university education. *Higher Education*, 55(5): 553–573.

French, J.R.P. (1956). A formal theory of social power. *Psychological Review*, 63(3): 181–194.

Guryan, J., Jacob, B., Klopfer, E. and Groff, J. (2008). Using technology to explore social networks and mechanisms underlying peer effects in classrooms. *Developmental Psychology*, 44(2): 355–364.

Hanushek, E.A., Kain, J.F., Markman, J.M. and Rivkin, S.G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5): 527–544.

Hommes, J., Rienties, B., de Grave, W., Bos, G., Schuwirth, L. and Scherpbier, A. (2012). Visualising the invisible: a network approach to reveal the informal social side of student learning. *Advances in Health Sciences Education*, 1–15.

Jeong, H.W.N. and Chi, M.T. H. (2007). Knowledge convergence and collaborative learning. *Instructional Science*, 35, 287–315.

Kandel, D.B. (1978). Homophily, selection and socializationin adolescent friendships. *The American Journal of Sociology*, 84(2), 427–36.

Karabenick, S.A. (2003). Seeking help in large college classes: A person-centered approach, *Contemporary Educational Psychology*, 28(1): 37–58.

Kelejian, H.H. and Prucha, I.R. (1998). A Generalized Spatial Two Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbance. *Journal of Real Estate Finance and Economics*, 17(1): 99–121.

Leenders, R.Th.A.J. (2002). Modeling Social Influence through Network Autocorrelation: Constructing the Weight Matrix. *Social Networks*, 24: 21–47.

Lubbers et al. (2006). The impact of peer relations on academic progress in junior high

Marsden, P.V. and Friedkin, N.E. (1993). Network Studies of Social Influence *Sociological Methods & Research*, 22(1): 127–151.

Mayer, A. and Puller, S.L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics*, 92(1–2): 329–347.

Ortiz, M.G.R., Hoyos, J.R.C. and Lopez, M.G.R. (2004). The social networks of academic performance in a student context of poverty in Mexico. *Social Networks*, 26(2): 175–188.

Rizzuto, T.E., LeDoux, J. and Hatala, J.P. (2009). It's not just what you know, it's who you know: Testing a model of the relative importance of social networks to academic performance. *Social Psychology of Education*, 12(2): 175–189.

Appendix

Figure A: Questionnaire figure for identifying members of the trust network.