Journal of Applied Statistics, 2015
http://dx.doi.org/10.1080/02664763.2015.1049517

Examining the effect of social influence on student performance through network autocorrelation models

Maria Prosperina Vitale[a*], Giovanni C. Porzio[b] and Patrick Doreian[c,d]

[a]*Department of Economics and Statistics, University of Salerno, Via Giovanni Paolo II, 132, Fisciano 84084, Italy;* [b]*Department of Economics and Law, University of Cassino and Southern Lazio, Via S. Angelo, località Folcara, Cassino I-03043, Italy;* [c]*Faculty of Social Sciences, University of Ljubljana, Kardeljeva ploščad 5 SI-1000 Ljubljana, Slovenia;* [d]*Department of Sociology, University of Pittsburgh, 2G03 Forbes Quad, Pittsburgh, PA 15260, USA*

(*Received 31 January 2014; accepted 6 May 2015*)

The paper investigates the link between student relations and their performances at university. A social influence mechanism is hypothesized as individuals adjusting their own behaviors to those of others with whom they are connected. This contribution explores the effect of peers on a real network formed by a cohort of students enrolled at a graduate level in an Italian University. Specifically, by adopting a network effects model, the relation between interpersonal networks and university performance is evaluated assuming that student performance is related to the performance of the other students belonging to the same group. By controlling for individual covariates, the network results show informal contacts, based on mutual interests and goals, are related to performance, while formal groups formed temporarily by the instructor have no such effect.

Keywords: interpersonal network; network autocorrelation; network effects model; social influence; student academic performance

1. Introduction

Studying the factors associated with student performance at the university level is of interest for many institutions. This motivated a wide literature, considerable research over the years and the foundation of dedicated journals. The main idea leading this research area is that understanding the keys of the academic performance could address proper strategies for increasing student potential achievement.

Within this large research field, Mills [35] pointed out that student performance is characterized as a complex process implying the interplay of individual and institutional factors.

*Corresponding author. Email: mvitale@unisa.it

Amongst them, recent works investigated the relationships with background variables [9], family support [11], academic and social integration [40], student learning patterns [44], student learning approach and time spent studying [26], self-perception of ability [12], teacher influence [22], course scheduling [15] and course resources [24].

An additional considered factor is the presence of the so-called peer effect (for a review, see [46]). That is, student achievement is related not only to his/her own individual characteristics but also to the way he/she interacts with peers, an approach largely investigated at lower levels of education from many perspectives (see e.g. [32,39] or the more recent [13]). We can refer also to the 'social influence' (or contagion) mechanism in which the social relations among individuals provide a basis for the alteration of actor behaviors in response to another actor in the network in which they are embedded [28,34].

It is worth noting, though, that far fewer contributions investigated the link between student relations and their performance at the university level. This is probably due to the complex task of defining groups amongst university students: while they arise quite naturally in lower level schools (students are embedded in classes), this typically does not happen at the university level. Some authors defined groups of students using administrative archives (subjects within the same class and sitting at the same exam session [14]). Others collected data through surveys and defined networks descending from the observed ties among students: for instance, Celant [10] defined a network considering if students study together in view of passing the same exams; Poldin [38] constructed networks based on friendship and study helpers ties.

However, in accordance with the economic literature investigating networks at lower levels of education (e.g. [7,41]), all these scholars adopted linear-in-means models to analyze peer effects on university student outcomes. That is, they studied how the individual outcome is affected by his/her own characteristics and by the mean values of the characteristics and outcomes of the group to which s/he belongs. This assumption requires that the whole network can be divided into a set of separate components. Put differently, all the network actors are linked to each other within subgroups, forming complete sub-networks therein, but each member of a group has no connection to others outside it.

This assumption does not fit well within a university context. A partition into disjoint components in which students have no connections with people outside the group, and have relational ties only within the group, is at odds with a reality where networks are typically not closed and complete systems. We also note that assigning group mean values to individuals is likely to mask information on how network ties work. In brief, linear-in-mean models are not really well suited for the study of the presence of social influence (or peer effects) on student performance at the university level, and suggest other statistical models be adopted.

From network studies, the most widely used statistical models to deal with social influence mechanisms are the network autocorrelation models [16,18,20]. They represent a family of regression-like models which can be used to capture and/or control for 'social endogeneity', in the presence of non-exogenous covariates resulting from the interaction between individuals [8]. In analogy with the (linear) spatial autoregressive/moving average models (SARMA) applied in geographical setting [2], network effects models (autoregressive AR component) and network disturbances models (moving-average MA component) were defined to incorporate both individual characteristics as well as interaction group effects as covariates in the regression setting.

The main advantage provided by this class of models is that they allow discarding the linear-in-mean assumptions discussed above. In particular, they are able to take into account all interpersonal ties within and between groups and so avoid the assumption of the diagonal blocks structure of the adjacency matrix under investigation.

Hence, this contribution considers network effects models to discover the presence of social influence mechanisms for explaining student performance at the university level. To this end, by using survey data collected on observed networks formed by a cohort of students enrolled at a graduate level in an Italian University, the relation between interpersonal networks and university performance is evaluated. We assume that student performance is related to the performance of the other students belonging to the same group. In more detail, the effect on the performance of both formal and informal links among students while controlling for other individual covariates is considered. Studying in groups established by the instructor and exchange of learning information are considered as formal relations. In contrast, friendship, personal support and advice plus studying in groups outside of classes are informal relations.

Finally, as a note of caution, it should be mentioned that our analysis cannot eliminate the simultaneity or reverse causality issue [31] in these kind of regression models.[1] Nor can it remove the empirical problem of student self-selection into networks (it cannot be assumed that student groups are formed randomly). That is, it cannot be distinguished whether students reported good performance because of their involvement in networks during their academic career or if their good performance gave them greater opportunity to be part of groups on the basis of similarity of academic goals and efforts. Two effects of selection (i.e. actors select others on the basis of similar behaviors) and influence (i.e. tied actors tend to influence each other) seem to appear in explaining social phenomena where the endogeneity issue still remains a challenging task [6]. Given this, as with most of the research conducted in this area, results should be treated as association between networks and student performance rather than as a study of causal effects. Yet, it still remains a matter of interest to describe how the effect of social interactions is related to student attainment at university.

The paper is organized as follows. Section 2 introduces the case study with emphasis on measuring performance. Section 3 reports a briefly description on the network effects model adopted to study social influence mechanism and provides a summary of model specification and the main results. Section 4 concludes with a discussion and final remarks.

2. The effect of interpersonal relations on performance: a study on a cohort of Italian students

As a result of the two main European higher education policy reform processes (the Bologna Process started in 1999 and the Lisbon Strategy in 2000), the university system in Italy has been subject to numerous changes. This, in turn, yielded a host of studies devoted to the Italian system over the last decades, covering different topics and combining case studies and methodological issues (see for instance [1,4,21], among others, and the contributions in [3]).

Within this debate, we investigate if and how student performance is related to the performance of other students with whom they interact during their academic career, that is, a social influence mechanism is hypothesized. Our main assumption is that, among others, performance is related to different networks in which individuals interact because of a need for exchanging learning information and of studying together, or because of their friendship or personal support and advice relations. More specifically, we presume that not all the networks in which the individuals are embedded during university life have the same effect on performance: informal contacts requiring and/or implying deeper relations (such as friendship or personal support) have a stronger relation with student performance. This hypothesis corresponds to the fact that informal relationships last longer, and are based on mutual interests and goals and cut across courses. On the other hand, formal groups are formed temporarily and by diktat, and generally disbanded after a course has finished.

To this end, the survey data we collected are used in an attempt to evaluate student relations as a factor related to their academic achievement. The reference population is a cohort of 81 students enrolled for the first time in the academic year 2008–2009 at a graduate two-year track in an Italian university.

Table 1. Main characteristics of variables selected for the case study.

Variable	Type	Average (*St.Dev.*)
Gender	Categorical (dummy)	91.9% female
Enrollment age	Continuous	25.4 (*4.6*)
Living near the university location	Categorical (dummy)	63.0% yes
Father years education	Discrete	10.6 (*3.4*)
Secondary school type	Categorical (dummy)	65.0% Lyceum
Secondary school grade	Continuous	79.7 (*10.8*)
Work during university	Categorical (dummy)	52.0% yes
Undergraduate level		
Average grade at exams	Continuous	26.3 (*1.7*)
Dublin descriptor 1	Discrete 1–10	7.2 (*1.4*)
Dublin descriptor 2	Discrete 1–10	7.3 (*1.2*)
Dublin descriptor 3	Discrete 1–10	7.7 (*1.2*)
Dublin descriptor 4	Discrete 1–10	7.5 (*1.6*)
Dublin descriptor 5	Discrete 1–10	8.1 (*1.4*)
Graduate level		
Average grade at exams	Continuous	28.4 (*1.3*)
Dublin descriptor 1	Discrete 1–10	7.8 (*1.2*)
Dublin descriptor 2	Discrete 1–10	8.0 (*1.1*)
Dublin descriptor 3	Discrete 1–10	8.1 (*1.4*)
Dublin descriptor 4	Discrete 1–10	8.1 (*1.2*)
Dublin descriptor 5	Discrete 1–10	8.5 (*1.3*)

2.1 *Data collection and measurement issues*

Our student data were collected through a survey exploiting a web platform as well as face-to-face interviews in May–October 2010. The gathered data were validated and integrated by using an administrative archive updated to December 2012.

The survey instrument was designed to collect information about student socio-demographic characteristics, prior educational attainments and their university careers. Some variables considered in the present study are provided in Table 1. Furthermore, network information about different kind of interpersonal relations were collected (as described below).

A key issue in the definition of the items to be included in the questionnaire was related to how to measure the students' success during their university career. Generally, most of the literature measures academic performance in terms of some objective indicators (e.g. the grade point average score (GPA) or some alternative formulations [1,25], the percentage of exams passed in a given time, the number of credits achieved, the final grade, etc.). Probably because of their availability in administrative databases, these measures are typically used as proxies of student performances. However, they might not be enough to describe the unobservable and not measurable nature of academic achievement [43].

For this reason, we decided to measure the complexity of the success of a learning process[2] by combining an objective performance indicator with a group of learning outcomes[3] as they are registered by students themselves at the end of a learning experience. This is consistent with a literature investigating how self-reported student evaluation could be considered a proper measure of student skills, under the hypothesis that skills acquired by students during their learning experience is part of their academic achievement [30].

Specifically, in our study the average grade at exams at the end of student academic career was considered as an objective performance indicator, whereas student self-reported learning experiences were defined according to the five core competences known as the Dublin Descriptors.[4] These descriptors consider the perceived quality of the learning activities developed during the field of study after completion of the higher education track as awarded to students by identifying five core competences to describe student learning experience (i.e. *knowledge and understanding, applying knowledge and understanding, making judgements, communication* and *learning skills*). For the sake of simplicity, a general item matching each descriptor was adopted. Students were asked to indicate the degree to which (on a 10-point scale, 1 = Low, 10 = High) they (i) have demonstrated knowledge and understanding in their field of study [Descriptor 1]; (ii) can apply their knowledge and understanding in occupational contexts [Descriptor 2]; (iii) have the ability to make autonomous judgements on well-defined problems [Descriptor 3]; (iv) can communicate about their understanding, skills and activities, with peers, supervisors and clients [Descriptor 4] and (v) have the learning skills to undertake further studies with some autonomy [Descriptor 5]. In many ways, these are more important indicators of attainment than simple summaries such as grades which also have a subjective component coming from instructors.

Social interactions among students during their enrollment at university were gathered using a whole-network study design [33] to measure the different network ties of this bounded cohort of students. For whole-network data collection, a roster list of the population was furnished to simplify the reporting task by reminding of the eligible students within each network.

We collected *one-mode* network data [45] for multiple types of links at a single time point. Students were asked to nominate their contacts for formal relations (exchange of learning information, classmate, and belonging to a working group established by the instructor) as well as informal contacts (studying in groups out of classes, friendship, personal support and advice, enrolled in the same on-line community, attending on-campus student associations, and spending spare time in campus activities). The network questionnaire items are shown in the Appendix.

2.2 *A first glance at the data*

Sixty-six students, out of 81, participated in the survey (an 81% response rate). The students not participating in this study did not substantially differ in terms of socio-demographic characteristics and prior scholastic attainment. Furthermore, closer inspection revealed that they were barely involved in their university studies or with other students and many (around 90%) of them had not yet graduated four years after enrollment. It is reasonable to exclude such students given their minimal involvement in the academic and social aspects of graduate study. In addition, four participants not yet graduated at the end of 2012 were excluded from our analysis as well. Finally, a very small amount of missing data were imputed using simple linear regression models. [5]

According to the literature, the variables selected for our study are related to the individual characteristics (sex, age at enrollment, living in a town near the university location), family background (education level and job position of parents), prior scholastic attainment at secondary school (type – lyceum/not lyceum – and final grade) and at the undergraduate level (average grade at exams, motivation, work during university studies, etc.). The main features of some variables are presented in Table 1. About 92% of respondents were female, the average age at enrollment was 25.8. They lived in 63% of cases in a town near the university and they were involved in a temporary employment position in half of the cases. About 65% of students have attended a lyceum at secondary school and the final grade was around 80 on average (on a scale ranging from 60 to 100). Father's education years were, on average, around 11 (about the midpoint in the high school track). Performance in terms of average grade at exams (scale: 18 = pass; 30 = maximum) and of the five Dublin's Descriptors showed high values both at undergraduate and graduate tracks. However, a general increase in student attainment is likely when comparing these two higher education levels.

Five student networks are considered in this study: exchange of learning information (EI), belonging to working group established by the instructor (WG), studying in group out of classes (SG), friendship (FR), and personal support or advice (AD). The other recorded networks were not considered here mainly because they were very sparse, in such cases 'there is little point in

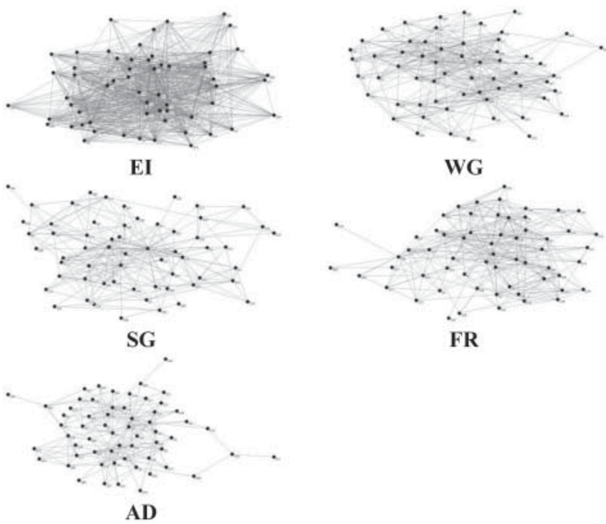


Figure 1. Graph representations of the five student relations. Their labels are exchange of learning information (EI), belonging to working groups established by instructors (WG), studying in groups out of classes (SG), friendship (FR) and personal support and advice (AD).

estimating autocorrelation models' [17, p. 40]. Two students out of 62 included in the study were isolates in at least one of the five observed social relations. Hence, their data were also discarded because their self-exclusion meant that they were not part of the social influence process being studied.

The student relational data are included in five binary adjacency matrices with cell values of 1 if student *i* was linked to student *j* and 0 otherwise. These matrices were symmetrized assuming reciprocated links among subjects for all existent connections. Studying in groups out of classes and belonging to working groups are symmetric relations by definition. As for the exchange of learning information, at face value, it is directed according to the direction of the information flows. However, an alternative conceptualization is to think of the relations as the context within which information can flow. It could even be thought of as a minimum level of trust in the sense of feeling safe enough to ask for information and comfortable enough with providing information. For friendship, personal support and advice, we consider that the relationship is close enough for this kind of support to be requested and given regardless of the direction of the flow.

Figure 1 provides a visualization of the five social relations we considered, while Table 2 provides a summary of some characteristics of the observed networks. As expected, the highest

Table 2. Network characteristics.

Network	Density	Average degree (*St.Dev*)	Degree centralization (%)
EI	0.453	26.733 (*10.542*)	54.822
WG	0.221	13.033 (*7.718*)	33.255
SG	0.132	7.767 (*5.934*)	42.490
FR	0.180	10.600 (*5.386*)	34.015
AD	0.120	7.067 (*4.686*)	22.677

Notes: Their labels are exchange of learning information (EI), belonging to working groups established by instructors (WG), studying in groups out of classes (SG), friendship (FR) and personal support and advice (AD).

Table 3. Dyadic QAP correlations among the adjacency matrices.

	EI	WG	SG	FR	AD
EI	1.00				
WG	0.46	1.00			
SG	0.39	0.56	1.00		
FR	0.38	0.53	0.58	1.00	
AD	0.37	0.57	0.67	0.63	1.00

Notes: Their labels are exchange of learning information (EI), belonging to working groups established by instructors (WG), studying in groups out of classes (SG), friendship (FR) and personal support and advice (AD).

density and average degree occur for the exchange of learning information: most of the students shared information with some others during their study at university. In contrast, studying in groups, providing personal support and advice have lower values: asking for support or studying together are more selective processes.

Finally, we investigated the extent to which multiple networks can present different characteristics of student relations. The correlations for the five adjacency matrices (Table 3) show that all relationships are positively associated.[6] All correlations are significant at or beyond the 5% level.

3. Network models for social influence

From network studies, the most widely used statistical models to deal with social influence mechanisms are the network autocorrelation models (NAMs) [16,18,20]. They have considered a 'workhorse for modeling network influences on individual behavior' [23] and still represent an active area of research [19,23,36,47]. NAMs deal with the presence of interdependent individual units embedded within social structures. In this case, standard linear regression models cannot be adopted because this interdependence violates the assumption of independence between error terms and response variable required to obtain unbiased coefficient estimates.

The two models within this class for dealing with social influence mechanisms are the network effects model and the network disturbances model [16]. Formally, let $\mathbf{y}$ be a $(n \times 1)$ n-vector of values of a dependent (endogenous) variable for n individuals making up a network, let $\mathbf{X}$ represents the $(n \times p)$ matrix of values for the n individuals on p covariates (including an unit vector for the intercept term) and let $\mathbf{W}$ be the $(n \times n)$ network weight matrix whose elements, w_{ij}, measure the influence actor j has on actor i.

The network effects model is defined as

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \epsilon, \tag{1}$$

where $\boldsymbol{\beta}$ is a $(p \times 1)$ vector of regression parameters, ρ is the network autocorrelation parameter referred as the strength of social influence in a network and error terms ϵ are assumed to be normally distributed with zero means and equal variances, $\epsilon \sim (0, \sigma^2 I)$.

The network disturbances model is defined as

$$\mathbf{y} = \mathbf{X}\beta + \epsilon, \tag{2}$$

$$\boldsymbol{\epsilon} = \rho\mathbf{W}\boldsymbol{\epsilon} + \boldsymbol{\nu}, \tag{3}$$

where ρ is the network autocorrelation parameter, and ν is a vector of random perturbations, $\nu \sim (0, \sigma^2 I)$. That is, the first models interdependencies between actors through the inclusion of an autocorrelation parameter in the dependent term, while the second includes interdependencies in the disturbance term.

For our purposes, the network effects model was adopted. It allows individual outcome (i.e. performance) to be directly associated with neighbors' levels of outcome. Furthermore, the estimate of the network autocorrelation parameter ρ provides information on the strength of the network effect.

An additional key element in these models is the specification of the network weight matrix **W**, conceptualizing the interdependencies among individuals. Leenders [29] reported several alternative ways for defining weights in **W**, starting from a scaled adjacency matrix measuring direct connection among units (called cohesion) or deriving (dis)similarity measures for subjects' relational profiles (using structural equivalence). According to a common practice [2], in our study directly scaled adjacency matrices were adopted, where rows are normalized to sum to 1, so that w_{ij} is a measure of the relative influence of subject j on i, with values $0 \leq w_{ij} \leq 1$, with diagonal terms $w_{ii} = 0$.

3.1 *Model specification and results*

Of the two network model options described above, we consider the network effects model (Equation (1)) because of our interest in the direct effects of actors on one another with respect to their academic performance. To incorporate in the model the performance in terms of objective as well as subjective indicators, individual scores for graduate (*GradPerf*), and undergraduate (*UnderGradPerf*) performance were derived by using confirmatory factor analysis.[7] The average

Table 4. Estimated network effects models with all controlling variables (full model).

	EI	WG	SG	FR	AD
Const	– 0.282	– 0.253	– 0.286	0.018	– 0.309
	0.912	*0.923*	*0.901*	*0.903*	*0.872*
Network effect ($\hat{\rho}$)	– 0.476	0.050	0.321	0.399*	0.377**
	0.483	*0.271*	*0.213*	*0.218*	*0.154*
UnderGradPerf	0.332***	0.343***	0.333***	0.298***	0.288***
	0.069	*0.069*	*0.068*	*0.071*	*0.069*
EnrAge	– 0.002	– 0.005	– 0.003	– 0.008	– 0.002
	0.019	*0.019*	*0.019*	*0.019*	*0.018*
Residence	0.231	0.211	0.133	0.209	0.193
	0.180	*0.182*	*0.185*	*0.176*	*0.172*
Lyceum	0.174	0.181	0.167	0.143	0.191
	0.182	*0.184*	*0.180*	*0.179*	*0.174*
HSGrade	0.002	0.002	0.002	– 0.001	0.002
	0.008	*0.008*	*0.008*	*0.008*	*0.008*
FathEdu	0.001	– 0.001	0.003	0.001	– 0.001
	0.025	*0.026*	*0.025*	*0.025*	*0.024*
Work	– 0.070	– 0.048	– 0.035	– 0.054	– 0.080
	0.182	*0.182*	*0.178*	*0.176*	*0.173*
R^2	0.363	0.358	0.357	0.352	0.345
AIC	138.5	139.5	137.4	136.5	134.1
BIC	159.5	160.4	158.3	157.4	155.0

Notes: Estimated coefficients, their standard errors (in italics), R^2, AIC and BIC for the estimated models. Response is the individual graduate performance score. The parameter ρ measures the magnitude of the network effect of each relation. Labels are: Age at Enrollment (EnrAge), Living in a town near university (Residence), Years Education of Father (FathEdu), Secondary school type (Lyceum/Not Lyceum), Secondary school grade (HSGrade), Work during university (Work), individual scores for undergraduate performance (UndGradPerf).

*$p < .10$.

**$p < .05$.

***$p < .01$.

grade at exams at the end of each student academic career was combined with self-reporting learning experiences defined according to the five core competences of the Dublin Descriptors.

Furthermore, according to the literature on student performance, the presence of social influence mechanisms was studied while controlling for the effect of the standard individual characteristics: age at enrollment (EnrAge), living in a town near the university location (Residence), education years of father (FathEdu) and having a job during university studies (Work). Prior scholastic attainment was controlled by using type of secondary school (Lyceum/Not Lyceum) and final grade at diploma (HSGrade), and performance in the undergraduate track (UndGradPerf). Gender was dropped because of the very high percentage of females among these students.

By considering separately social influence mechanisms implied in the defined networks, one-regime network effects models were estimated for each of the different examined network. Here, we report results for the following full model, where all controlling variables were included, (Equation (4)) and reduced model, controlling for a subset of all the involved variables, (Equation (5)):

$$\begin{aligned}\text{GradPerf} = \alpha + \rho_i \mathbf{W_i}\text{GradPerf} + \beta_1\text{UnderGradPerf} + \beta_2\text{EnrAge} + \beta_3\text{Residence}\\ + \beta_4\text{FathEdu} + \beta_5\text{Lyceum} + \beta_6\text{HSGrade} + \beta_7\text{Work} + \zeta,\end{aligned} \tag{4}$$

$$\text{GradPerf} = \alpha + \rho_i \mathbf{W_i}\text{GradPerf} + \beta_1\text{UnderGradPerf} + \zeta, \tag{5}$$

where $\mathbf{W}_i$ $(i = 1, \dots, 5)$ represents the five network weight matrices entering the model separately (described in the following by the acronyms IE, WG, SG, FR, AD), and ρ_i parameter shows the strength of social influence on student performance for each considered network, while taking into account the effect of the individual covariates.

Tables 4 and 5 show the results for the full and reduced estimated network effects models, respectively. In general, there was a significant effect of performance at undergraduate level in all models (a quite natural effect) while the other socio-demographic characteristics are not significant (consistent with what has been observed in other studies in the literature). In detail, when all controlling variables were included (Table 4), the results showed no significance[8] for all the involved parameters, except for the parameter associated to undergraduate performance and the two autocorrelation parameters related to friendship (FR) and support and advice (AD)

Table 5. Estimated network effects models with undergraduate performance (UndGradPerf) (reduced model).

	EI	WG	SG	FR	AD
Const	0.042	– 0.013	– 0.016	– 0.036	– 0.020
	0.099	*0.095*	*0.084*	*0.086*	*0.083*
Network effect ($\hat{\rho}$)	– 0.400	0.092	0.384*	0.416*	0.311**
	0.476	*0.269*	*0.200*	*0.214*	*0.068*
UnderGradPerf	0.357***	0.362***	0.345***	0.314***	0.374***
	0.068	*0.069*	*0.066*	*0.070*	*0.157*
R^2	0.327	0.322	0.326	0.313	0.308
AIC	129.9	130.5	127.3	127.2	125.4
BIC	138.3	138.9	135.6	135.6	133.8

Notes: Estimated coefficients, their standard errors (in italics), R^2, AIC and BIC for the estimated models. Response is the individual graduate performance score. The parameter ρ measures the magnitude of the network effect of each relation.

*$p < .10$.

**$p < .05$.

***$p < .01$.

network effects. The reduced models (Table 5) by controlling for only the undergraduate performance showed some interesting network results: all informal contacts (SG, FR, and AD) between students were related to their academic performance. The effect of social influence mechanisms thus differs according to the kind of relation taken into account. On the one hand, exchange of learning information, and belonging to working groups established by instructors are not relevant in explaining performance. On the other hand, studying in groups out of classes ($\hat{\rho}_{\mathrm{SG}} = 0.384$), friendship relation ($\hat{\rho}_{\mathrm{FR}} = 0.416$) and personal support and advice ($\hat{\rho}_{\mathrm{AD}} = 0.311$) are positively associated with student success at university.

4. Discussion and final remarks

This contribution represents a novel approach in terms of method by adopting a network autocorrelation model to analyze how relations among students are related to their academic attainment at the graduate level by controlling for individual covariates.

One-regime network effects models were estimated by considering individual scores for student performance, measured by combining an objective indicator (average grade at exams) with a set of subjective student self-perception indicators of their learning process based on Dublin's Descriptors. Individual scores were obtained through a confirmatory factor analysis, according to the procedure in [42] for a latent regression model.

A key issue for these models is the definition of a proper weight matrix to effectively capture the connections among students. We started from the idea that links among students in a learning environment can improve their performance. We used direct connections described by row-normalized dichotomous adjacency matrices. This implies that the performance of a student is related to the average performance of her/his neighbors. Of course, there is room for deeper analyses of the way social influence mechanism among peers works within university settings. This could include a way to take into account that probably students with a *worse* performance can improve if they are connected with better students – but students with *better* performances do not perform less well if they are connected with less good students. In addition, the combined use of multiple autocorrelation network effects in a single model [19,47] of the same group of subjects who are embedded in different networks as a further development of this study can be considered.

Our results suggest that belonging to social networks is positively related to student performance – especially if these networks are created by the students rather than be imposed by the instructors. More specifically, by controlling for the effect of individual characteristics, prior scholastic attainments at secondary school and at undergraduate level, some relations turned out to be significant, while others do not. In particular, informal communications (such as friendship, personal support and advice) are significantly related with graduate student success at university, whereas merely exchanging information and working in groups were not.

Given this is a study limited to a graduate track in a single university and in the presence of the endogeneity issue, caution is merited. Its generalizability is very limited and reverse causality cannot be ruled out. However, would our findings be confirmed by further analysis in different university contexts and with endogeneity controlled, one practical implication follows. It would be beneficial if higher educational institutes encourage social activities between students as potentially effective strategies for learning. In particular, universities can address measures to improve students' interactions so as to facilitate the integration in academic and social life, and thereby contributing in their academic success.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. In analogy with economic literature, Dow [19] discussed an alternative estimation solution to deal with the presence of the endogenous term in network autocorrelation models framework based on instrumental variables and 2SLS estimation method. However, the choice of instruments is a task that can be hardly faced in studying peer effects [37], and it goes beyond the scope of this contribution.
2. Worthy of note is the recent 2012 feasibility study on the measure of student performance carried out by the OECD 'Assessment of Higher Education Learning Outcomes (AHELO)' project. For details visit OECD AHELO project website: www.oecd.org/edu/ahelo.
3. 'Statements called intended learning outcomes, commonly shortened to learning outcomes, are used to express what it is expected that students should be able to do at the end of the learning period' [27, p. 3].
4. The Dublin descriptors comprise 'generic statements of typical expectations of achievements and abilities associated with qualifications that represent the end of each Bologna cycle' [5, p. 65].
5. A first model was estimated to predict the graduate average grade of five students from their undergraduate grade. The model was also used to obtain values for five missing undergraduate grades (an inverse regression problem). The same procedure was adopted for the Dublin Descriptors, where each graduate level descriptor was regressed on the corresponding undergraduate one. In such a case, a total of six missing values were imputed using 10 variables.
6. We computed the Pearson correlation for all pairs of the five networks considered, and assessed the frequency of random measures as large as actually observed by using the dyadic QAP-correlation tool implemented in the UCINET software.
7. Several methods are available to obtain individual scores for latent variables, generally providing different results for each unit. We adopted the revised blockwise factor score regression procedure [42], that is based on the estimation of individual scores separately for the dependent (Bartlett scores) and independent (Regression scores) latent variables. It produces consistent estimators for all parameters in the case of a latent regression model.
8. This is probably due to the large number of parameters (if compared with the small sample size) that in turn provided large standard errors.

References

[1] G. Adelfio, G. Boscaino, and V. Capursi. *A new indicator for higher education student performance*, Higher Educ. 68 (2014), pp. 653–668.
[2] L. Anselin, *Spatial Econometrics: Methods and Models*, Kluwer Academic, Dordrecht, 1988.
[3] M. Attanasio and V. Capursi, eds., *Statistical Methods for the Evaluation of University Systems*, Physica, Heidelberg, 2011.
[4] S. Bianconcini and S. Cagnone, *A general multivariate latent growth model with applications to student achievement*, J. Educ. Behav. Stat. 37 (2012), pp. 339–364.
[5] Bologna Working Group, *A framework for qualifications of the European Higher Education Area*. Bologna Working Group Report on Qualifications Frameworks (Copenhagen, Danish Ministry of Science, Technology and Innovation), 2005.
[6] S.P. Borgatti, A. Mehra, D.J. Brass, and G. Labianca, *Network analysis in the social sciences*, Science323 (2009), pp. 892–895.
[7] Y. Bramoullé, H. Djebbari, and B. Fortin, *Identification of peer effects through social networks*, J. Econom. 150 (2009), pp. 41–55.
[8] C.T. Butts, *Social network analysis: A methodological introduction*, Asian J. Social Psychol. 11 (2008), pp. 13–41.
[9] M. Byrne and B. Flood, *Examining the relationships among background variables and academic performance of first year accounting students at an Irish University*, J. Account. Educ. 26 (2008), pp. 202–212.
[10] S. Celant, *The analysis of students' academic achievement: the evaluation of peer effects through relational links*, Qual. Quant. 47 (2012), pp. 615–631.
[11] W. Cheng, W. Ickes, and L. Verhofstadt, *How is family support related to students' GPA scores? A longitudinal study*, Higher Educ. 64 (2012), pp. 399–420.
[12] A. Chevalier, S. Gibbons, A. Thorpe, M. Snell, and S. Hoskins, *Students' academic self-perception*, Econom. Educ. Rev. 28 (2009), pp. 716–727.
[13] D. Contini, *Immigrant background peer effects in Italian schools*, Social Sci. Res. 42 (2013), pp. 1122–1142.
[14] M. De Paola and V. Scoppa, *Peer group effects on the academic performance of Italian students*, Appl. Econ. 42 (2010), pp. 2203–2215.
[15] A.K. Dills and R. Hernández-Julián, *Course scheduling and academic performance*, Econ. Educ. Rev. 27 (2008), pp. 646–654.
[16] P. Doreian, *Linear models with spatially distributed data: spatial disturbances or spatial effects?*, Sociol. Methods Res. 9 (1980), pp. 29–60.

[17] P. Doreian, *When the data points are not independent*, in *Developments in Data Analysis*, A. Ferligoj and A. Kramberger, eds., Metodoloski zvezki, 1996, pp. 27–46.
[18] P. Doreian, K. Teuter, and C.-H Wang, *Network autocorrelation models: some Monte Carlo results*, Sociol. Methods Res. 13 (1984), pp. 155–200.
[19] M.M. Dow, *Galton's problem as multiple network autocorrelation effects cultural trait transmission and ecological constraint*, Cross-cult. Res. 41 (2007), pp. 336–363.
[20] M.M. Dow, M.L. Burton, and D.R. White, *Network autocorrelation: a simulation study of a foundational problem in regression and survey research*, Social Netw. 4 (1982), pp. 169–200.
[21] M. Enea and M. Attanasio, *An association model for bivariate data with application to the analysis of university students' success*, J. Appl. Stat. (2015), pp. 1–12 (ahead-of-print).
[22] G. Foster, *Teacher effects on student attrition and performance in mass-market tertiary education*, Higher Educ. 60 (2010), pp. 301–319.
[23] K. Fujimoto, C.-P. Chou, and T.W. Valente, *The network autocorrelation model using two-mode data: Affiliation exposure and potential bias in the autocorrelation parameter*, Social Netw. 33 (2011), pp. 231–243.
[24] G. Huon, B. Spehar, P. Adam, and W. Rifkin, *Resource use and academic performance among first year psychology students*, Higher Educ. 53 (2007), pp. 1–27.
[25] V.E. Johnson, *An alternative to traditional GPA for evaluating student performance*, Statist. Sci. 12 (1997), pp. 251–269.
[26] D. Kember, Q.W. Jamieson, M. Pomfret, and E.T.T. Wong, *Learning approaches, study time and academic performance*, Higher Educ. 29 (1995), pp. 329–343.
[27] D. Kennedy, *Writing and Using Learning Outcomes: A Practical Guide*, University College Cork, 2007.
[28] R.Th.A.J. Leenders, *Longitudinal behavior of network structure and actor attributes: modeling interdependence of contagion and selection*, in *Evolution of Social Networks*, P. Doreian and F. Stokman, eds., Routledge, The Netherlands, 1997, pp. 165–184.
[29] R.Th.A.J. Leenders, *Modeling social influence through network autocorrelation: Constructing the weight matrix*, Social Netw. 24 (2002), pp. 21–47.
[30] A. Lizzio, K. Wilson, and R. Simons, *University students' perceptions of the learning environment and academic outcomes: implications for theory and practice*, Studies Higher Educ. 27 (2002), pp. 27–52.
[31] C.F. Manski, *Identification of endogenous social effects: The reflection problem*, Rev. Econ. Stud. 60 (1993), pp. 531–542.
[32] S. Maroulis and L.M. Gomez, *Does 'Connectedness' matter? Evidence from a social network analysis within a small-school reform*, Teachers College Record 110 (2008), pp. 1901–1929.
[33] P. Marsden, *Survey methods for network data*, in *The SAGE Handbook of Social Network Analysis*, J. Scott and P.J. Carrington, eds., Sage, London, 2011, pp. 370–388.
[34] P.V. Marsden and N.E. Friedkin, *Network studies of social influence*, Sociol. Methods Res. 22 (1993), pp. 127–151.
[35] C. Mills, J. Heyworth, L. Rosenwax, S. Carr, and M. Rosenberg, *Factors associated with the academic success of first year health science students*, Adv. Health Sci. Educ. 14 (2009), pp. 205–217.
[36] M.S. Mizruchi and E.J. Neuman, *The effect of density on the level of bias in the network autocorrelation model*, Social Netw. 30 (2008), pp. 190–200.
[37] A.J. O'Malley, *The analysis of social network data: an exciting frontier for statisticians*, Stat. Med. 32 (2013), pp. 539–555.
[38] O. Poldin, D. Valeeva, and M. Yudkevich, *How social ties affect peer group effects: Case of university students*, Higher School of Economics Research Paper No. WP BPR 15, 2013.
[39] M.G. Ramírez Ortiz, J.R. Caballero Hoyos, and M.G. Ramírez López, *The social networks of academic performance in a student context of poverty in Mexico*, Social Netw. 26 (2004), pp. 175–188.
[40] B. Rienties, S. Beausaert, T. Grohnert, S. Niemantsverdriet, and P. Kommers, *Understanding academic performance of international students: the role of ethnicity, academic and social integration*, Higher Educ. 63 (2012), pp. 685–700.
[41] B. Sacerdote, *Peer effects in education: How might they work, how big are they and how much do we know thus far?* in *Handbook of the Economics of Education*, E.A. Hanushek, S.J. Machin and L. Woessmann, eds., Vol. 3, Elsevier, St. Louis, MO, 2011, pp. 249–277.
[42] A. Skrondal and P. Laake, *Regression among factor scores*, Psychometrika 66 (2001), pp. 563–575.
[43] K.C. Soh, *Grade point average: what's wrong and what's the alternative?*, J. Higher Educ. Policy Manag. 33 (2010), pp. 27–36.
[44] J.D. Vermunt, *Relations between student learning patterns and personal and contextual factors and academic performance*, Higher Educ. 49 (2005), pp. 205–234.
[45] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge, 1994.

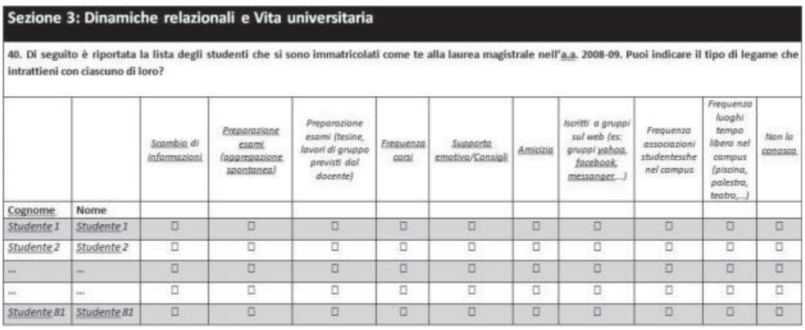
[46] E.A. Weihua, *Models and methods to identify peer effects*, in *The SAGE Handbook of Social Network Analysis*, J. Scott and P.J. Carrington, eds., Sage, London, 2011, pp. 514–532.

[47] B. Zhang, A.C. Thomas, P. Doreian, D. Krackhardt, and R. Krishnan, *Contrasting multiple social network autocorrelations for binary outcomes, with applications to technology adoption*, ACM Trans. Manag. Inf. Syst. 3 (2013), pp. 1–21.

Appendix

The whole questionnaire was implemented in Italian. The network items is partially depicted in Figure A1. The question says: ‘A list of the students enrolled in your Master program in the academic year 2008–2009 is reported below. Please select the kind of relationship you have with each of them’.

Then, for each row, a student name appears and the interviewed person must check a box corresponding to the different kind of relations. They are: Exchange of learning information; Studying together out of classes; Working in groups established by instructors; Classmate; Personal support and advice; Friendship; Enrolled in the same on-line community (e.g. yahoo groups, facebook, messanger, . . .); Attending on-campus student associations; Spending spare time in campus activities (swimming pool, gym, theatre, . . .); I do not know him/her.



Sezione 3: Dinamiche relazionali e Vita universitaria

40. Di seguito è riportata la lista degli studenti che si sono immatricolati come te alla laurea magistrale nell’a.a. 2008-09. Puoi indicare il tipo di legame che intrattieni con ciascuno di loro?

		Scambio di informazioni	Preparazione esami (aggregazione spontanea)	Preparazione esami (tesine, lavori di gruppo previsti dal docente)	Frequenza corsi	Supporto emotivo/Consigli	Amicizia	Iscritti a gruppi sul web (es: gruppi yahoo, facebook, messanger,...)	Frequenza associazioni studentesche nel campus	Frequenza luoghi tempo libero nel campus (piscina, palestra, teatro,...)	Non lo conosco
Cognome	Nome										
Studente 1	Studente 1	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Studente 2	Studente 2	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
...	...	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
...	...	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Studente 81	Studente 81	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Figure A1. An image of the questionnaire network items (in Italian).