De Giorgi, Giacomo; Pellizzari, Michele; Redaelli, Silvia

Working Paper

Be as careful of the books you read as of the company you keep: evidence on peer effects in educational choices

IZA Discussion Papers, No. 2833

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: De Giorgi, Giacomo; Pellizzari, Michele; Redaelli, Silvia (2007) : Be as careful of the books you read as of the company you keep: evidence on peer effects in educational choices, IZA Discussion Papers, No. 2833, Institute for the Study of Labor (IZA), Bonn, https://nbn-resolving.de/urn:nbn:de:101:1-2008032811

This Version is available at:
https://hdl.handle.net/10419/33925

DISCUSSION PAPER SERIES

IZA DP No. 2833

Be as Careful of the Books You Read as of the Company You Keep: Evidence on Peer Effects in Educational Choices

Giacomo De Giorgi
Michele Pellizzari
Silvia Redaelli

June 2007

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Be as Careful of the Books You Read as of the Company You Keep: Evidence on Peer Effects in Educational Choices

Giacomo De Giorgi
Stanford University

Michele Pellizzari
IGIER-Bocconi and IZA

Silvia Redaelli
Bocconi University

Discussion Paper No. 2833
June 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 2833
June 2007

ABSTRACT

Be as Careful of the Books You Read as of the Company You Keep: Evidence on Peer Effects in Educational Choices[*]

In this paper we investigate whether peers' behavior influences the choice of college major. Using a unique dataset of students at Bocconi University and exploiting the organization of teaching at this institution, we are able to identify the endogenous effect of peers on such decision through a novel identification strategy which solves the common econometric problems of studies of social interactions. Results show that, indeed, one is more likely to choose a major when many of her peers make the same choice. We estimate that, when it diverts students from majors in which they seem to have a relative ability advantage, this effect leads to lower average grades and graduation mark, a penalty that could cost up to 1,117 USD a year in the labor market.

JEL Classification: J0, I21

Keywords: peer effects, education, social interaction, reflection

Corresponding author:

Giacomo De Giorgi
Department of Economics
Stanford University
579 Serra Mall
94305-6072 Stanford, CA
USA
E-mail: degiorgi@stanford.edu

[*] We thank Ran Abramitzki, Liran Einav, Eliana La Ferrara, Enrico Moretti, Steve Pischke, Luigi Pistaferri and Imran Rasul for their comments. We also thanks seminar participants at IZA/CEPR Summer Symposium, Stanford University, Bocconi University, the London School of Economics, ISER, the Universities of Bologna, Modena and Verona. We are grateful to Stefano Gagliarducci, Pietro Garibaldi, Francesco Giavazzi, Andrea Ichino and Enrico Rettore for providing the initial data for this project. We also thank Giacomo Carrai, Mariele Chirulli, Cherubino Profeta, Alessandra Startari and the administration of Bocconi University for allowing access to their archives and helping with the extraction of additional data. Enrica Greggio, in particular, has been a constant source of precious information and help. Special thanks also to Amelia Spinelli for suggesting the title of this work. The usual disclaimer applies.

1 Introduction

The importance of peers in shaping individual and social behavior has been widely recognized in both the economic and the sociological literature. Numerous studies have produced empirical evidence showing the existence of relevant peer effects in many areas, from schooling performances to criminal behavior, from productivity to financial decisions (Katz and Case, 1991; Hoxby, 2000; Sacerdote, 2001; Duflo and Saez, 2002; Ammermueller and Pischke, 2006).

However the identification of social interactions remains very problematic because of two well-known issues: endogeneity - due either to peers self-selection or to common group (correlated) effects - and reflection - a particular case of simultaneity (Manski, 1993; Brock and Durlauf, 2001; Moffitt, 2001 and Soetevent, 2006).

The contribution of this paper is twofold. First, on the methodological side, we develop a new strategy for the identification of endogenous peer effects.[1] In a policy perspective, this is the crucial parameter for interventions that directly influence one's peers outcomes. A leading example is the immunization program implemented in Kenya analyzed in Miguel and Kremer (2004) or the PROGRESA program in Mexico (Todd and Wolpin, 2006). Moreover, in a general equilibrium framework with social interactions, endogenous effects are paramount for the ex-ante evaluation of any policy intervention.

Second, we estimate the role of peer effects on students' choices of college major. Further, we are able to assign a wage value to choosing an academic major according to one's peers, thus shedding some light on the mechanism that generates social interactions.

Our study is based on a newly constructed set of administrative data of undergraduate students from Bocconi University.

The particular structure of the degree programs offered by this institution allows to define peer groups that vary at the level of the single individual and as such are not subject to the usual simultaneity (reflection) problem. Moreover, we identify a natural set of exogenous instruments to control for correlated (within group) effects, which have been recognized as important determinants of group outcomes.[2]

[1] The endongenous peer effect is usually defined as the impact of the average peers' outcomes on individual outcomes. See Section 4 for details.

[2] A similar approach could in principle be adopted in a number of other contexts, i.e. whenever units of analysis are linked directly to some other units (the peers) but only indirectly (through peers) they are further connected to

At Bocconi University, students initially enroll in a common program and only at the end of the third semester (i.e. after 1 and 1/2 years) do they choose whether to specialize in one of two majors: business or economics. During these first three semesters all students take nine compulsory courses and attend lectures in randomly assigned classes. Since the number of available lecturers varies for each course, the assignment of students to classes is repeated for each course (see Section 2).

This setting allows us to define peer groups using information on class assignment. In other words, we assume that student i interacts with the other students attending the same lecture (in the same classroom with the same lecturer) in any of the common compulsory courses. The repeated process of random assignment to classes generates peer groups that vary at the individual level: student i's peers study with i but also with other students who are not necessarily members of i's peer group. As the peer groups of i and i''s peers do not coincide, we are able to solve the reflection problem (see Section 4).

Moreover, since the allocation into the classes is random, endogeneity in peer group formation is excluded by construction. Having peer groups that vary at the individual level also guarantees the presence of excluded classmates, i.e. students who did not attend classes with i but did attend some courses with some of i's peers. The exogenous characteristics of excluded peers represent a natural set of instruments to overcome potential endogeneity generated by common (correlated) group effects.[3]

Further, our data also contain a very rich set of observable proxies for those variables that are commonly believed to induce self-selection (i.e. ability, motivation, preferences etc.).

The combination of the particularly rich dataset, the repeated randomization and the peculiar construction of the peer groups allows us to solve the two key econometric problems of this literature: reflection and endogeneity, be it induced by sorting or correlated effects.

Our econometric methodology differs from the existing literature that tries to recover peer effects using either laboratory experiments (Falk and Ichino, 2006), natural experiments (Sacerdote, 2001; Zimmerman, 2003), quasi-experimental designs (Hoxby, 2000), or fixed effects (Hanushek *et al.*,

others (the peers of peers). In the network literature (Calvó at al., 2004) this corresponds to the existence of links of degree 2. For example, groups of this type may arise when members of a football team are also members of other social groups (baseball, study group, etc...) and the two groups do not perfectly overlap.

[3]The usual suspects for group shocks in the education framework are teachers's effects or classmates' disruptive behaviors.

2003). The repeated randomization process exploited here distinguishes our approach from most previous studies (Sacerdote, 2001; Zimmerman, 2003) where the randomized assignment - when there is one - is typically done only once and for all.[4] Laschever (2005) is, to our knowledge, the first application of a multiple group framework.

The spirit of our identification strategy is also similar to Bayer *et al.* (2004), a study of criminal behavior that exploits the length of the individual's sentence to weight each peer's characteristic by the time spent together in the same correctional institution. Two other papers - developed independently at the same time as ours - are very close to our approach: Bramoullé *et al.,* 2006 and Calvó *et al.,* 2006. The former provides the theoretical identification conditions for endogenous peer effects in a network framework, while the latter focuses on the position of a player in a particular network in influencing performance.

With our approach we are able to identify the causal effect of peers' choice of major (economics vs. business) on one's own decision. The only paper that looks at this particular outcome is Sacerdote (2001), which does not find any significant influence of peers. In that paper, however, peer groups were defined on the basis of a single random assignment to rooms in campus dorms. Therefore, only the potential bias from endogenous sorting is excluded, while groups are fixed across peers and as such cannot (directly) account for reflection and possible common shocks. In fact, consistently with Sacerdote's results, our estimates are statistically significant only when we control for the potential bias from correlated effect.

Our results show that, indeed, one is more likely to choose a major when many of her peers make the same choice. We, then, look at whether students who specialize in a major following the choices of their peers and against their revealed relative ability (measured as the ratio of one's average grade in economics and business courses during the first three semesters) perform better (in terms of average grades in the last three semesters, final graduation mark and time to graduation) than similar students who chose the major according to their revealed ability and against the majority

[4]However, the first insights to our identification strategy were somewhat implicit in the work of Manski (1993) and particularly in Moffitt (2001). Manski (1993) suggests the possibility of extending the model of interactions to multiple groups in Footnote 1(i) of page 534. Moffitt (2001) suggests the use of a partial population experiment, generating an exclusion restriction, along the same line of reasoning as our approach, which is also based on exclusion restrictions and can be seen as a peculiar partial population experiment. The original approach proposed by Moffit (2001) is taken in Bobonis and Finan (2005), Lalive and Cattaneo (2005), Cipollone and Rosolia (2006), and Cooley (2006).

of their peers.[5] Our findings indicate that, indeed, there is a negative effect of following one's peers when revealed ability would suggest a different choice. We, then, try to assign a monetary value to this effect by looking at the wage cost of such a lower academic performance. We estimate that cost to be as high as 1,117 USD a year.

We can think of at least three mechanisms that are potentially important in generating the type of social interaction we see. First, peer pressure (Mas and Moretti, 2006), being it monitoring or imitation, might be substantial in leading a student towards a particular major choice. Second, there might simply be a utility gain to studying with friends. Third, peers may facilitate the acquisition of information (or constitute a reference group in the formation of expectations) on university life and job opportunities associated with a particular major.

Although our research design is not best suited to distinguish among these alternatives, it seems plausible to rule out the information mechanism. In fact, better informed students should make "better" choices and this is at odds with our findings in terms of average grades and graduation mark. Other papers (Ichino and Maggi, 2000) have devoted more attention to the analysis of the specific mechanism that generate social interactions but typically without being able to separately identify endogenous and exogenous peer effects.

The paper is organized as follows: Section 2 describes the institutional structure of Bocconi University, the available data and the details of the allocation of students into classes; Section 3 presents our approach for the construction of the peer groups; Section 4 discusses the identification strategy and the results of the analysis of the choice of major. In Section 5 we provide a number of robustness checks. Section 6 discusses the effects of the decision modes on average GPA, graduation mark and time to graduation. Finally, Section 7 concludes.

2 Data and institutional details

The analysis in this paper is based on administrative data from Bocconi University, an Italian private institution of higher education that specializes in business and economics. The data provide detailed information on the university curricula of all students enrolled at Bocconi since 1989.

[5] Final graduation mark is in the particular University the sum of GPA plus additional points rewarded to a compulsory dissertation.

Until the academic year 1999/2000, the most popular degree offered by Bocconi was called CLEA/CLEP. Students in this degree would first take a series of nine common exams during the first three semesters and would then choose whether to specialize in business (CLEA) or economics (CLEP) (See Figure 1). The nine common compulsory courses are listed in Table 1 and can be classified by subject areas according to the department responsible for the teaching: business, economics, quantitative subjects and law.

[FIGURE 1 and TABLE 1]

In the academic year 1999/2000 Bocconi introduced a major reform of its structure (the so-called "Bocconi 2000" plan), abandoning this initial common track and forcing students to choose between economics and business upon entering the University. The information on the random allocation of students to classes has unfortunately been lost for the earlier cohorts of students and is reliable only starting with the academic year 1998/1999. This forces us to use only one cohort of students, i.e. students enrolled in the old CLEA/CLEP program in the academic year 1998/1999.

At that time, Bocconi offered four other degree programs: one in "Economic and Social Sciences" (DES), one in "Economics of Financial Market Institutions" (CLEFIN), one in "Management of the Public Administration and International Institutions" (CLAPI) and one in "Law and Business Administration" (CLELI).[6] These degree programs differ both in their curricula and in the number of students admitted in each academic year.[7] In September 1998, a total of 2,580 students were admitted and 2,055 of them eventually enrolled at Bocconi.[8]

In their application package, perspective students had to rank the five programs according to their preferences. Admission was based on a standardized entry test combined with high school performance. Applicants were then ranked according to these results and, starting from the top of the ranking, students were assigned to their preferred programs depending on availability. Specifically, a student was assigned to her first choice if there were still places available in that program,

[6] Created in 1970, CLEA (Degree in Business Administration) and CLEP (Degree in Economics) are the oldest degrees offered at Bocconi University. Four years later, they were joined by DES, a more quantitative and academic version of the CLEP. All the other degrees (CLEFIN, CLAPI and CLELI) were introduced in 1990.

[7] Enrolment ceilings and admission tests were introduced in 1984.

[8] We are excluding students transferring from other universities and students from abroad who were given reserved places.

otherwise, if all places in her first choice had already been taken by students higher up in the ranking, the candidate was assigned to her second choice and so on.

It is important to notice that in this mechanism the student's stated preferences across the five programs do not influence the probability of being admitted and thus excludes any strategic behavior in the reporting of preferences. This allows us to use this information to construct an indicator of ex-ante preferences. In particular, we consider students who indicated the DES degree - the more academic oriented version of CLEP - as a first or a second choice as "determined" to do economics since the beginning of their studies.[9] Similarly, students who indicated DES as a last choice are coded as "determined" to specialize in business.

Admitted candidates who decided not to register freed places for students further down in the ranking. However, only a few students (48 out of the 753 rejected candidates) who had been initially rejected took up a place freed by others, possibly because at the time of making these decisions most people had already obtained admission to another university and started to make arrangements for the registration and the accommodation.[10]

Eventually, the admission procedure in September 1998 led to 1,385 students (against a ceiling of 1,600) enrolled in the common CLEA/CLEP track, followed by CLELI (239 against a ceiling of 350), CLEFIN (208 against a ceiling of 230) CLAPI and DES (respectively with 132 and 91 against ceilings of 200 each). Once enrolled, CLEA/CLEP students were not allowed to switch to any of the other degrees, while students enrolled in the CLELI, CLEFIN, CLAPI and DES programs could move to CLEA/CLEP only after the first academic year.

In this paper we will focus exclusively on students enrolled in the CLEA/CLEP common track. Excluding a few missing values on our variables of interest and those students who did not complete the courses of the first 3 semesters, our working sample consists of the 1,141 observations described in Table 2. All of these students have complete information about their courses in the initial three semesters. A few of them (slightly less than 10%) have not graduated, either because they dropped out, changed university or are still enrolled and trying to graduate.

[TABLE 2]

[9] These are students who either had CLEA/CLEP as a first choice and DES as a second or DES as first and CLEA/CLEP as second and did not get a place in the DES.

[10] Note also that candidates in the lower tail of the distribution of the admission test were not offered any of these residual places.

After the first 3 semesters of common courses, each student originally enrolled in CLEA/CLEP had to choose whether to specialize in business (CLEA) or in economics (CLEP). Table 3 reports some descriptive statistics on the ability and performance of these two groups of students.

[TABLE 3]

Considering all the common exams in the first three semesters, the 145 students choosing CLEP score on average almost 2 grade-points above CLEA students (exams are graded on a scale 0 to 30 with pass equal to 18). This difference is even higher when the exams are disaggregated by field. As expected, CLEP students perform relatively better in economics, while the difference is considerably smaller for the average grade in business exams, suggesting - as we will see more formally later on - that students choose their field of specialization according to their relative abilities or interest. Furthermore, the difference in the average grade of the exams of the quantitative courses is also very large, reflecting the nature of the CLEP program that was considerably more quantitative than CLEA.

2.1 Lecturing classes

The number of classes created for each of the nine common exams depends on the number of available lecturers.[11] Moreover, the capacity of the available classrooms at Bocconi varied considerably and the number of students in each class had to be determined accordingly.

Students were randomly allocated to classes for each course. The decision to adopt a random allocation algorithm was dictated by the need to avoid congestion in the classrooms resulting from students wanting to attend lectures with their friends or with the "best" teachers.

Towards the end of each term, students had to enroll in courses of the following term either at the administration desk or through some computers located in the university buildings.[12] Moreover,

[11] The terms *class* and *lecture* often have different meanings in different countries and sometimes also in different schools within the same country. In most British universities, for example, lecture indicates a teaching session where an instructor - typically a full faculty member - presents the main material of the course. Classes are instead practical sessions where a teacher assistant solves problem sets and applied exercises with the students. At Bocconi there was no such distinction, meaning that the same randomly allocated groups were kept for both regular lectures and applied classes. Hence, in the remaining of the paper we are going to use the two terms interchangeably.

[12] Enrolment in the courses of the first term of the first year was automatic. Students were also free to choose

students who failed to pass an exam during the academic year in which they had attended the corresponding course, were required to re-register and were also assigned randomly to a new class (together with other students).[13] For these reasons, the total number of students enrolled in each course (the sum over all the classes) may vary slightly across courses.

At the time of enrollment, the algorithm would randomly assign the student to a class for each course and communicate the allocated class number.[14] The algorithm was designed to fill all classrooms at the same rate in order to obtain a final distribution with an adequate number of students in each room. By no means could the students interfere with the algorithm. For example, there was no guarantee that two students enrolling in the same course one right after the other would be placed in the same teaching class (and, in fact, despite the many that attempted to do so, this instance was extremely rare).

In principle, students were required to attend lectures in their assigned classes but enforcement varied substantially over time, becoming stricter in more recent years. Actually, the evolution of enforcement practices is closely related to the availability of the information on lecturing classes: as the enforcement of the allocations was made more and more stringent, lecturing classes were also recorded on various official documents and thus maintained in the administration's archives.

The mere fact that lecturing classes have been carefully recorded for the 1998/1999 cohort is an indication that the system was effectively enforced.[15] Students were forced to attend their classes by various methods. First, lecturers were supposed to circulate attendance sheets at the beginning of the class for students to sign their presence. Obviously, with a large number of students in each class (the average class size was 202 students), this method could be easily circumvented by those who wanted to attend a different class by, for example, having some friends signing for them. Mid-terms were also important in encouraging students to attend their assigned classes. In fact, while the final exams were identical for all students regardless of their classes, mid-terms were organized directly by the lecturers. Therefore, if a student wanted to take the mid-term (which were not compulsory but highly recommended and very popular among the students) she'd better attend her assigned

whether they wanted to postpone some of the courses (e.g. take a course of the second semester in the third and so on) provided they satisfied the pre-requisites for each exam (e.g. statistics could only be taken after having passed math).

[13] There are normally up to 7 exam sessions per year for each of the 9 common courses during the academic year.

[14] This was just a particular number or letter by which it would be easy to look for venues and communication concerning a particular class on the University notice board system.

[15] There are less than 2% of missing values.

class as the exam was prepared and marked by the same lecturer.

[TABLE 4]

Table 4 describes the average characteristics of the lecturing classes for each course. The number of classes ranges from 4 (private and public law) to 10 (mathematics, management and accounting) and the average number of enrolled students varies accordingly. The other variables in Table 4 are derived from students' questionnaires. At the end of each course, during a regular lecture time, students were distributed a standardized anonymous questionnaire to collect their opinions about numerous aspects of the teaching (quality of the lectures, logistics, etc.). A detailed description of the data available at the level of the single class is provided in Table A.1 in the appendix.

The number of completed questionnaires is a one-off measure of attendance, as it should correspond to the number of students present in class on the day the questionnaire was distributed. Attendance is also self-reported by the students in the questionnaire, where they have to indicate the fraction of lectures they attended for that course. These figures indicate that attendance was typically very high, with students being present at over 80% of the lectures for economics, management and quantitative courses.

Only law subjects have very low attendance levels. At that time Bocconi did not have a law department and relied exclusively on external professors (from other universities). For this reason the number of law classes that could be created was relatively small (4) and their size was consequently extremely high; the administration was well aware of low attendance for these courses.

3 Peer group definition

Our definition of peer groups is based on students attending the same classes and it is meant to capture the network in which students interact academically and socially. The underlying assumption is that these interactions are fostered by class attendance so that the relevant set of peers for each student overlaps (at least partly) with classmates.[16]

[16] If two students were to attend only one course together they would sit in the same class for six hours a week (three two hours classes) for one semester.

While considering classes is standard in the literature on peer effects in high-school, in our case effective attendance as well as the size of the lecturing classes cast doubts on the possibility of capturing relevant peer interactions by looking at assigned classmates (see Table 4). We address this problem by excluding the two law courses from our definition of peer groups and also by weighting peers by the number of common courses attended together.[17]

Formally, individual i's peer group (G_i) includes all individuals j who were assigned to the same class as individual i for at least one of the 7 courses that we consider (all 9 common exams minus the 2 law subjects). Furthermore, each of the $j \in G_i$ is given an importance weight, $\omega_{ij}(0,1]$, according to the number of common courses taken together with i, i.e. $\omega_{ij} = 1$ if j attends all 7 courses in the same class as i, $\omega_{ij} = 1/7$ if j attends only 1 course with i.[18]

As a further robustness check, we conduct our analysis also using groups formed on the basis of a stricter definition of peers, namely students who have attended at least 4 of the 7 common courses together.[19] This restricted definition is particularly interesting because it leads to groups sizes that are comparable to other papers in the literature, particularly those that have looked at high-school classes.

[TABLE 5]

The first two columns of Table 5 report some characteristics of these groups. In column 1 the groups are constructed considering 7 courses while in column 2 we consider peers only students who have attended at least 4 courses in the same classes. The mean raw group size is approximately 674 students in the first case and goes down to 18 in the second. On average students in these groups are assigned to the same classes for 1.6 and 4.2 courses respectively, which implies that, when peers are weighted by the number of courses taken together, the size of the groups goes down to 151 with 7 courses and 10.7 with our restricted definition.

[17] Public and private law are the courses with the largest average class size as well as the lowest average class attendance, both self-reported and measured by the ratio of collected questionnaires over the number of officially enrolled students.

[18] The weights adopted in the core of the paper are linear in the number of courses attended together. We have experimented with many other specifications and the results are robust to the weighting scheme, see Section 5.

[19] This is essentially a variation in the weighting scheme that assigns zero weight to peers who have attended less than 4 courses together. We choose the threshold of 4 courses because it is the highest that guarantees a non-empty peer group for all students (i.e. there are some students who have never taken more than 4 courses with others).

[TABLE 6]

Table 6 provides evidence to support our definition of peers by showing that, after the initial 3 semesters, students who have attended lectures in the same random classes also show similar academic patters. In the upper panel of Table 6, for example, we contrast the incidence of peers and non-peers that choose the same sub-major (i.e. field). In fact, within each of the main majors - economics and business - students can specialize in different fields, like marketing or accounting within business and finance or theory within economics. The students in our sample could choose among 8 sub-majors within the economics area and 16 sub-majors within the business area. Using our most comprehensive definition of peers, i.e. students who have attended at least one of the 7 common courses in the same random class (column 1 in Table 6), on average slightly more than 9.6% of peers choose the same major. This compares to a marginally lower incidence of students making similar choices among the non-peers (i.e. students who have never taken any class together). As we restrict our definition of peers to students who have attended more and more courses in the same classes (columns 2 and 3), the difference between peers and non-peers increases and becomes statistically significant. Only with the stricter definition (column 4) this difference becomes smaller and insignificant again.

The lower panel of Table 6 analyzes graduation sessions. In the period covered by our data students could graduate in several different sessions throughout the year (almost one session per month). During these sessions, which lasted one or two days, students present their final dissertation to a commission which decides their final mark (based on both the dissertation and their GPA). Students can freely choose when to graduate, a decision that is usually affected both by how quickly they complete their coursework and by how much time they spend on their dissertation.[20] For the average student in our sample approximately 12.5% of the non-peers graduate in the same session. This number goes up to 13.4% for peers in our widest definition (column 1) and increases steadily as the definition becomes more stringent (columns 2 to 4). The differences are always strongly significant.

The evidence in Table 6 shows that randomly assigned peers eventually follow similar academic patterns, suggesting that they actually interact with each other. Moreover, the stronger effects

[20] Late graduation has always been one of the most serious problems of the Italian university system. See Garibaldi *et al.* (2007).

that emerge for peers that have attended more and more courses together supports our weighting scheme, which should indeed emphasize the most intense interactions. In section 5 we perform additional robustness checks by modifying our definitions of peer groups.

4 Peer effects in major choices

The identification of endogenous social effects has been the topic of several papers (Manski, 1993; Brock and Durlauf, 2001 and Moffitt, 2001 to cite just a few) and it rests on two distinct dimensions: endogeneity and reflection. Endogeneity may arise for at least two reasons: first, people usually choose endogenously their peers and, second, the unobserved shocks that affect the group as a whole (teacher effects are the usual suspect in studies of education) may also generate endogeneity. As a consequence, when detecting a significant correlation between individual and group outcomes, one cannot say whether this result is due to true peer effects or simply to endogenous group formation (along some unobservable characteristics) and/or common correlated effects.

The second problem - reflection - arises because in a peer group everyone's behavior affects the others and, as in a mirror reflection, we cannot know if one's action is the cause or the effect of peers' influence. This is essentially a problem of simultaneity.

Let us start with a discussion of how we address reflection. This problem has been commonly described by using a simple linear in means model:

$$y_i = \alpha + \beta E(y|G_i) + \boldsymbol{\gamma} E(\mathbf{x}|G_i) + \boldsymbol{\delta}\mathbf{x}_i + u_i \tag{1}$$

In our framework, y_i is the chosen major (i.e. economics or business), $\mathbf{x}_i$ is a set of individual traits, $E(\mathbf{x}|G_i)$ contains the averages of the $\mathbf{x}$'s in the peer group of individual i, denoted by G_i. Following the literature, β measures the endogenous effect, $\boldsymbol{\gamma}$ the exogenous effects. For now assume $E(u_i|G_i, \mathbf{x}_i) = 0$, i.e. no correlated effects or self-selection into groups.

In the standard framework, peer-groups are fixed across individuals, i.e. if A and B are both in the peer group of C, it must also be that A and B are in the same group. Put in the wording of equation (1), if i and j are in the same peer-group, then the two groups coincide, i.e. $G_i = G_j$. In this situation, endogenous effects cannot be distinguished from exogenous effects (Manski, 1993). In fact, it is easy to show, by simply averaging equation (1) over group G_i, that $E(y|G_i)$ is a linear

combination of the other regressors:

$$E(y|G_i) = \left(\frac{\alpha}{1-\beta}\right) + \left(\frac{\boldsymbol{\gamma}+\boldsymbol{\delta}}{1-\beta}\right) E(\mathbf{x}|G_i) \tag{2}$$

In our framework peer groups are instead individual specific. Consider the simple case of only three students. Students A and B study together (e.g. they attend 3 courses in the same classes), however, B also studies with C (e.g. they attend some of the remaining 4 courses in the same class, different from A's class). A's peer group, thus, includes only B while B's peer group includes both A and C. This identification can also be seen as a case of triangularization. In the standard simultaneous equation model at least one exogenous variable is excluded from each equation; here, A is excluded from the peer group of C, who is excluded from the peer group of A.

With 7 courses, each divided into 6 to 10 lecturing classes, our data exhibit enough variation to generate peer-groups that vary at the level of the single individual, e.g. every student has a different peer-group. The weighting scheme described in the previous section adds more variation to the individual peer groups.

To formally see the advantage of this framework in solving the reflection problem, rewrite equation (2) allowing peer-groups to vary at the level of the single individual:

$$E(y_i|G_i) = \alpha + \beta E[E(y|G_j)|G_i] + \boldsymbol{\gamma} E[E(\mathbf{x}|G_j)|G_i] + \boldsymbol{\delta} E(\mathbf{x}_i|G_i) \tag{3}$$

where j is a generic member of i's peer group. The key to understanding this equation is the fact that j's peer group G_j never coincides with G_i.

This result can also be clarified by the previous example with 3 students: A, B and C where A and B are in the same class for one subject and B and C sit together in another course. This structure implies that $G^A : \{B\}$, $G^B : \{A, C\}$ and $G^C : \{B\}$. Equation (1), then, translates in the following three equations:

$$\begin{aligned}
y_A &= \alpha + \beta y_B + \gamma x_B + \delta x_A + u_A^A \\
y_B &= \alpha + \beta \left(\frac{y_B + y_C}{2}\right) + \gamma \frac{(x_B + x_C)}{2} + \delta x_B + u_B^B \\
y_C &= \alpha + \beta y_B + \gamma x_B + \delta x_C + u_C^C
\end{aligned}$$

Now, consider the corresponding reduced form equations:

$$\begin{aligned}
y_A &= \left(\alpha + \frac{\alpha\beta\left(1+\beta\right)}{1-\beta^2}\right) + \left(\frac{\beta(\gamma+\delta)}{1-\beta^2} + \gamma\right) x_B + \left(\frac{\beta(\gamma+\delta\beta)}{1-\beta^2}\right)\left(\frac{x_A+x_C}{2}\right) + \delta x_A + \eta_A^A \\
y_B &= \left(\frac{\alpha\left(1+\beta\right)}{1-\beta^2}\right) + \left(\frac{\gamma+\delta}{1-\beta^2}\right) x_B + \left(\frac{\gamma+\delta\beta}{1-\beta^2}\right)\left(\frac{x_A+x_C}{2}\right) + \eta_B^B \\
y_C &= \left(\alpha + \frac{\alpha\beta\left(1+\beta\right)}{1-\beta^2}\right) + \left(\frac{\beta(\gamma+\delta)}{1-\beta^2} + \gamma\right) x_B + \left(\frac{\beta(\gamma+\delta\beta)}{1-\beta^2}\right)\left(\frac{x_A+x_C}{2}\right) + \delta x_C + \eta_C^C
\end{aligned}$$

where the new reduced form error terms- η_A^A, η_B^B and η_C^C - are linear combinations of the structural error terms - u_A^A, u_B^B and u_C^C.[21] The example above shows how we achieve identification: we are left with four reduced form parameters and four structural ones. Notice, additionally that in this particular case the last equation is redundant and, in fact, only observations with distinct groups of peers contribute to identification.[22]

Although this particular setting allows to solve reflection, one might still worry about the presence of correlated effect, i.e. common unobservable shocks at the group level which could flaw the previous identification result. Suppose, in fact, that the general error term is of the following form:

$$u_i^g = \mu_i + \theta^g + \varepsilon_i \tag{4}$$

with $g = A, B, C$ and where μ_i is an individual fixed effect, θ^g a group fixed effect (e.g. teacher quality, disruptions), and ϵ_i an i.i.d. random component.[23]

If we were to substitute 4 into 1 we would have to face two problems of endogeneity arising from the individual effect (μ_i) and the group effect (θ^g). In our particular case, the random nature of the peer groups rules out correlation between the individual effect and any endogenous or exogenous effect ($E(y|G_i)$ and $E(x|G_i)$).[24] However, unobservable group shocks could still be present and induce endogeneity, i.e. $Cov\left(E\left(y|G_i\right), \theta^g\right) \neq 0$.[25] Even if our strategy effectively solves reflection,

[21] The meaning of the double indexing - subscript and superscript - will become clear in a few paragraphs.

[22] In fact A and C here have the same peer group, $\{B\}$, although they are not peers to each other.

[23] The double indexing of the previous error terms should clarify the fact that these errors include both an individual specific (μ_i) and a group shock (θ^g).

[24] Additionally, our data include several observable proxies for variables that are generally unobservable to the econometrician (i.e. standardized ability test, high-school grades, type of high-school, preferences, etc.) and we make use of all of them to purge our results from potential residual endogeneity

[25] Note that correlated effects cannot induce endogeneity of the exogenous effect - $Cov\left(E\left(x|G_i\right), \vartheta^g\right) = 0$ - since

the presence of correlated effects may still generate endogeneity of $E(y|G_i)$ and impede identification.

One possible solution is to use instrumental variables. Fortunately, this setting naturally offers valid instruments, namely peers of peers who are not in one's peer group. In fact, the $\mathbf{x}$'s of students who are excluded from i's peer group but included in the group of one or more of i's peers are by construction uncorrelated with the group fixed effect of i and correlated with the mean outcome of i's group through peer effects. In our previous example, x_C would be a valid instrument for y_B in group A. The logical chain is the following: x_C, which is uncorrelated with θ^A, affects y_C and, since C is a peer of B, through endogenous effects y_C also affects y_B. For the same reasoning x_A would be a valid instrument for y_B in group C.[26]

In our data, the group of peers of peers - which we label *excluded peers* for clarity - for a generic student i includes all other students who have never taken any of the 9 common courses in the same lecturing classes of i but have taken some of the 7 courses that we consider with one or more of i's peers. The average raw size of these groups is 252 students, as reported in the third column of Table 5. Notice additionally that the union of the groups of excluded and actual peers never spans the entire sample. The student with the largest groups is linked either directly or indirectly to 1085 students, thus allowing for more than 50 totally excluded peers. On average the sum of the two groups is 927 and notice that we keep the same definition of excluded peers also when using the restricted definition of peers.[27]

To better document the absence of self-selection in our setup, Table 7 reports the correlation coefficients between individual and group averages of some measures of predetermined ability and motivation for various definitions of peers. In column [1] peer groups are constructed considering all common exams excluding the two courses in law, whereas in column [2] the groups are based on the restricted definition of peers, i.e. students who have attended at least 4 courses together.

[TABLE 7]

The numbers in Panel A of Table 7 show that peers are not clustered by any of the attributes

the x's are determined prior to the allocation to the groups.

[26] Bramoullè et al. (2006), a working paper developed parallelly and independently from this work, also discuss this IV methodology but, as far as we know, our paper is the first empirical application of this strategy.

[27] This guarantees that the excluded peers of student i never attended any course in the same class of i. We could eliminate from the excluded peers those who have attended less than 4 courses together with any of i's peers but this would lead to an empty set of excluded peers for many observations.

considered. This result is obviously not very surprising given that our peer-groups are based on random assignment to classes. The same reasoning applies to the results in Panel B which shows the correlations between individual and the average excluded peers' attributes.

To conclude, Panel C reports correlations between the average peers' characteristics and the average characteristics of the excluded peers, which are nothing but a random subsample of each peer group's complementary set. The negative and significant correlations arise mechanically from the fact that any small deviation of peers' attributes from the population averages is counterbalanced by a symmetric opposite deviation in the characteristics of the non-peers and hence also of the excluded peers. This mechanical correlation adds power to our instruments.

4.1 Results

As already mentioned, the CLEA/CLEP program offered only two majors: economics and business. Students had to make their choice after the initial three common semesters and the remaining five terms were clearly differentiated across the two majors.[28]

To estimate the effect of peers on one's decision to specialize in economics versus business, we run both a linear probability model and a probit regression similar to equation (1), where $y_i = 1$ if a student chooses economics and 0 otherwise. $E(y|G_i)$ is the share of peers choosing economics (weighted by the number of exams taken together) and $\mathbf{x}_i$ is a set of controls for individual characteristics that includes a gender dummy, household income (as recorded at the first registration), a dummy for students who reside outside the city of Milan (the site of Bocconi), a set of dummies for the region of origin, a series of controls for academic performance and ability (high-school type and grades, results of the admission test) and an indicator of ex-ante preferences over the two majors (i.e. whether a student was determined to do economics at enrollment, as described in Section 2). Given the randomness and the relatively large size of the peer groups we have very little variation in $E(x|G_i)$ to separately identify the constant and γ in equation (1). Therefore, in the main specification we omit the average predetermined characteristics of the peer group. The results are however robust to controlling for a subset of $x's$ at the group level, the estimated γ in that specific case is never significantly different from zero.[29] Moreover, when working with the smaller groups

[28] Although some elective courses could be picked from any of the two majors, nevertheless such practice was quite uncommon and the number of such options very limited.

[29] Our IV strategy uses the mechanical correlation between the $x's$ of peers and those of the excluded peers, if we

of restricted peers, the (random) sampling variation in $E(x|G_i)$ is larger and allows to separately identify γ and the constant (see Tables A.4 and A.5).

[TABLE 8]

Table 8 reports the results of the estimation of a linear probability model and for our two definitions of peer groups, one based on all 7 common courses - columns 1 to 3 - and one based on the restricted set of peers who have attended at least 4 courses in the same classes - columns 4 to 6. For each of these definitions we estimate the model under three different specifications: simple OLS, IV using the exogenous characteristics of the excluded peers (i.e. the peers of peers who are not in one's peer group) as instruments and IV using the same instruments weighted by the number of courses that each excluded peer has attended with any of the student's peers (see Section 4 for a detailed description of the instruments).

These estimates clearly indicate the presence of significant endogenous peer effects in the choice of major. Considering the first definition of peers in columns 1 to 3, only in the OLS specification the estimated endogenous effect is not significant while the IV results are considerably (5 to 6 times) larger.

For the correct interpretation of these results one should keep in mind that our measure of the endogenous effect weights peers by the number of courses attended in the same classes. Thus, for the average student the effect of one additional average peer - i.e. students with whom she has taken 1.57 courses together (see Table 5) - who chooses economics increases the probability of choosing economics by approximately 0.8-0.9 percentage points (according to our IV estimates).[30] Similarly, the effect for the average student of having one more of her strongest peers - i.e. students with whom she has taken all the 7 common courses together - choosing economics is approximately a 4 percentage-point increase in the probability of choosing economics.[31]

were to control for $E(\mathbf{x}|G_i)$ in the main equation we would lose part of the IV strength.

[30] The average student in our sample has approximately 151 average peers - i.e. peers with whom he/she has attended an average of 1.57 courses - and approximately 13% (i.e. approximately 20) of them choose economics as a major (see Table 5). Hence, if one additional (average) peer chooses economics, $E(y|G_i)$ for the average student increases by 1.57/(7x151)=0.0015, which according to our IV estimates leads to an increase in the probability of choosing economics of 0.0015x5.785=0.0086.

[31] Hence, if one additional strong peer chooses economics, $E(y|G_i)$ for the average student increases by 1.57/151=0.0066, which, according to our IV estimates, leads to an increase in the probability of choosing economics of 0.0066x5.785=0.0383.

It is generally thought that the OLS results over-estimate the actual size of the peer-effect because they cumulate the impact of both exogenous and endogenous effects, further it is often unclear how a group shock biases those results. In our analysis, however, exogenous peer-effects are ruled out by the random nature of the groups and we consequently exclude them from the main specification presented in Table 8.[32] The remaining concern in estimating the endogenous effects in our study rests on the possible presence of correlated effects, i.e. common group shocks.

The most common interpretation of correlated effects typically assumes that the group shock affects all students in the same direction, thus leading to lower dispersion in individual outcomes within groups. In our case, the positive difference between the IV and the OLS estimates suggests, instead, that the correlated group shock leads students in the same group to make more differentiated choices that they would have otherwise made. A possible example with teacher quality would be the following: encountering the most informative of economics professors offers all students a clear picture of what the subject is really about thus allowing them to make their own choice according to their actual preferences and without relying much on their peers.

To support this interpretation, we repeated the analysis focusing exclusively on the subset of students with the most homogeneous groups of peers along one specific dimension, that is the initial preference for economics. We selected only those students with either very many (top 90^{th} percentile) or very few (bottom 10^{th} percentile) peers who were "determined" to do economics since their first enrolment at Bocconi. In these homogenous groups we expect the correlated shock to affect (almost) everyone in the same direction, thus leading to similar individual behaviors. Consistently with this interpretation, in this selected sub-sample the IV estimate are smaller than the OLS (we omit the results for brevity). Thus, the difference between the OLS and the IV coefficients seem to indicate that correlated effects play an important role in this set up. In particular, the relative quality of teachers in the two areas (economics vs. business) may in fact be one of the crucial determinants of students decisions. It might also be that in some classes a particularly disruptive behavior could effectively compromise the thorough understanding of the more formal subjects, creating a similar effect to that described earlier.

Notice additionally that the limited variation in the endogenous variable - $E(y|G_i)$ - exacerbates

[32] As mentioned earlier and documented in Tables A.4 and A.5, the inclusion of exogenous peer effects does not affect our results significantly, especially when we consider the smaller groups of restricted peers that allow for more sampling variation in $E(\mathbf{x}|G_i)$.

the downward bias of the OLS estimate (which is in the order of 5 to 6 times the IV estimates).[33] This feature of the data also helps comparing our results with what is found in Sacerdote (2001), where peers are also randomly assigned but the data do not allow to solve the reflection problem nor the potential endogeneity due to common group shocks. Our OLS estimates are in line with the results in Sacerdote (2001) where no significant effect is found on major choice. However, once we account for possible endogeneity, the effect becomes sizeable and significant.

The last three columns (4 to 6) of Table 8 repeat the same exercise using our restricted definition of peers. The estimated coefficients are now much smaller. This is consistent with the fact that, given the smaller groups, $E(y|G_i)$ now varies a lot more.[34] In fact, the standard deviation of the (weighted) share of peers choosing economics is now equal to 0.09 while it was only to 0.01 if the groups include all peers encountered in the 7 common courses that we consider.

The estimated endogenous effect is still insignificant in the OLS specification and becomes statistically important in the IV estimation. The magnitude of these effects is also in line with previous findings: one additional (average) peer opting for economics raises the probability of making the same choice by approximately 0.7 percentage points.

[TABLE 9]

These results are very robust to changes in the specification of the model. Table 9 reports the same estimates produced under a probit specification and shows that the endogenous effects are now slightly more significant and that the marginal effects computed at the average of the distribution of the right-hand-side variables are of about the same magnitude of the results of the linear models.

These estimates indicate the presence of strong and large endogenous peer effects and are obtained using instrumental variables that appear to be very significant in explaining the endogenous

[33] To clarify this point, consider a simple linear model with just one regressor: $y = x\beta + \varepsilon$, where x is endogenous and a valid instrument z is available. In this simple case, the OLS estimator can be written as: $\widehat{\beta}_{OLS} = \beta + \frac{Cov(x,\varepsilon)}{Var(x)}$. In the particular case of the linear probability model, it is easy to show that, for given $Var(E(x|y))$, the bias is larger the smaller the variance of the endogenous variable.

[34] An argument similar to the one used to explain the difference between OLS and IV clarifies this point. In a simple model with just one regressor $\beta = \frac{Cov(x,y)}{Var(x)}$. However, when y is a dummy variable this becomes: $\beta = \frac{\overline{y}[E(x|y=1)-\overline{x}]}{Var(x)}$, which clearly shows that when $Var(x)$ increases only (or mostly) within groups defined by y the value of β declines.

term.[35] The F-test of excluded instruments, reported at the bottom of Table 8, is always very large. Table A.2 in the appendix shows the complete first-stage regressions for all our IV specifications (note that the first-stage regressions are identical for both the linear and the probit models).

Theoretically we could have used a very large set of instruments (all the exogenous characteristics of the excluded peers), however, in order to maximize efficiency, we have selected a subset of the most powerful ones, i.e. admission test, high school final grade and preferences for economics. All these instruments are also singularly significant in the first-stage regressions at a very strong level, with t-statistics between 4 and 8.

As discussed in Section 4, by construction the exogenous characteristics of the excluded peers influence the outcomes of peers in the opposite direction as one own's characteristics. This explains why the first-stage coefficients of the instruments are negative, while the corresponding individual variables have a positive impact on individual outcomes. The only exception is the high school final grade which is, however, very highly correlated with the admission test result.

5 Robustness

Throughout the paper we relied on a number of more or less stringent assumptions. In this section we present a series of robustness checks to give a sense of whether a particular stand is central to the main results of the paper.

First of all, we repeat our estimates including the (weighted) average of all exogenous characteristics of peers in the set of regressors. We can do this only when the groups are based on the restricted definition of peers who sit at least 4 courses in the same classes otherwise the exogenous peer effects $E(\mathbf{x}|G_i)$ vary too little to be separately identified from the constant. As shown in row 2 of Tables A.4 (for linear models) and A.5 (for probit models), results are very similar to our baseline estimates. Moreover, none of the exogenous effects is significant in these regressions, given that their variation is generated only by random sampling differences.

The definition of peer groups that we have adopted throughout the paper is based on the assumption that students interact in the classroom and that this particular group (classmates) is

[35] Although the analysis focus on a selective institution we have no reasons to believe that endogenous peer effects should be stronger in such setting.

relevant as far as the majoring decision is concerned. It is likely that students who meet regularly (if they share only one subject they meet at least 6 hours a week for a semester) are somehow bound to interact and influence each other.[36] However some students attend more than one subject together and, given that they spend more time in the same venue, on average they should also interact more.

We deal with this particular feature of the class assignment process by assigning a larger weight to students that meet more often. Throughout the paper we presented results based on either a very simple linear weighting or on a more extreme scheme that assigns weight zero to any peer that has been encountered in less than 4 courses. However, we have experimented a number of other schemes and in Tables A.4 and A.5 we present all the results on endogenous peer effects employing two alternative weighting schemes. First (in the third row of the tables), we exclude the courses in quantitative subjects (math and statistics) from our definition of peer groups, thus relying only on peers that attended courses in economics and business (5 courses in total) in the same classes. Additionally, in the fourth rows of Tables A.4 and A.5, we go back to considering all 7 courses but we adopt an exponential weighting scheme, which assigns to each peer j of student i a weight equal to the exponential of the number of courses that i and j have attended in the same classes, minus 1.

The estimated endogenous peer effects are highly comparable with our baseline specifications in Tables 8 and 9. The weighting does not seem to be central for identification. Moreover, the exponential (but also, although to a smaller extent, the restriction to only 5 courses) generates by construction a larger variation in the exogenous group effects - $E(\mathbf{x}|G_i)$ - and in fact it helps the identification of our parameter of interest, as indicated by the significance levels in Tables A.4 and A.5.

We have also performed the following thought experiment: assume that the peer groups we have defined have nothing to do with any type of social interaction and the effects that we estimate are generated by mere sample variation (or anything else). We then construct placebo peer groups by artificially and randomly assigning students to hypothetical classes. We expect to find no significant endogenous peer effect when the groups are formed using this artificial allocation. In fact, in none of the many specifications reported in the fifth rows of Tables A.4 and A.5 is there indication of significant social interactions and the magnitude of the point estimates is much closer to zero then

[36] Three 2-hour lectures per week.

in Tables 8 and 9 (when using the same definition of peers based on all 7 courses).

Our strategy for dealing with group specific shocks relies on the IV approach discussed in Section 4. As an alternative, we can also construct observable proxies for the plausibly most important group shock: teacher quality. From the student evaluation questionnaires we can identify for each of the 7 courses the "best" and the "worst" lecturers as those who received the highest and the lowest average mark on the item named *quality of teaching*.[37] The results reported in Tables A.4 and A.5 (row 6) are obtained from models similar to those in Tables 8 and 9, where we have augmented the set of control variables with a dummy for each of the courses considered, which takes value 1 if the student attended the course in the class of the lecturer who obtained the best students' evaluation. The estimates are again in line with those of Tables 8 and 9, suggesting that either the teacher dummies are not fully capturing the effect of teaching quality or that other group shocks may also be important (e.g. disruptive behavior).

Finally, information from the students' questionnaires (see Table 4 and Table A.1) suggests that in some cases the actual allocation of students into the classes might have not been maintained. Several anecdotes tell that, especially for the most difficult courses, students tended to cluster in the class of the best teacher regardless of their officially assigned class. Our data provide some evidence in this direction. For example, from Table 4 we know that in mathematics class 12[38] the number of questionnaires collected on the day of the course evaluation (253) was almost 60% higher than the number of officially enrolled students (161).

To account for the possibility that students assigned to the same teaching class may actually attend a course in different classes, we adjust our weights by proportionally lowering the importance of peers encountered in courses where there are signals that the official allocation was not effectively maintained. We identify these particular courses by exploiting the following question from the students' questionnaire: *"For your learning, the number of students attending your class has been: insufficient (1), too low (2), ideal (3), too high (4), excessive (5)"*. Tables 4 and A.1 report the average score of this question - which we label *congestion* - across courses and for each single class, respectively.

Courses in which the random allocation is not maintained should be characterized by a large

[37] Students are asked to give a synthetic evaluation on a scale 0 to 10.

[38] For anonymization purposes, this is a randomized version of the true class identifier.

variation across classes in this measure of congestion, i.e. there should be some classes with very many students and others with very few students. We, then, construct course weights by assigning weight 1 to the course with the lowest maximum level of reported congestion across classes (i.e. 2.51 for Management II) and the weights of the other courses are scaled down accordingly (these weights are shown in the last column of Table 4). The peers of a generic student i are then assigned a weight equal to the sum of the course weights corresponding to the courses taken in the same classes as i (normalized to sum to 1 within groups).

The last rows of Tables A.4 and A.5 report the estimated endogenous peer effects under this particular weighting scheme and show that the results are very similar to our baseline, suggesting that, in fact, problems of congestion were limited to a few cases.

6 Are books better than company?

In this section we analyze the relationship between students academic performance in the second half of their degree (i.e. the non-common semesters) and how they chose their major, i.e. based more on their own revealed ability or on their peers' behavior.

To this purpose, we construct two indicators. The first one, f_i, measures the relative fraction of peers who made one's same choice of major. Suppose individual i chooses to specialize in economics, then f_i is computed as the ratio between the (weighted) fraction of i's peers who also chose economics and the fraction of all students in the sample who chose economics. If $f_i > 1$ it means that in i's peer group there is a higher than average incidence of students in economics. Similarly for students who chose business. More formally, f_i is defined as follows:

$$f_i = \begin{cases} \frac{\sum_{j \in G_i} \omega_j ECON_j}{N^{-1} \sum ECON_j} & \text{if} \quad ECON_i = 1 \\ \frac{\sum_{j \in G_i} \omega_j BUSINESS_j}{N^{-1} \sum BUSINESS_j} & \text{if} \quad BUSINESS_i = 1 \end{cases} \tag{5}$$

where $ECON_i$ is a dummy variable equal to 1 if student i chooses economics and zero otherwise (similarly for $BUSINESS_i$).

The second indicator, g_i, is a measure of relative ability. Our data include very detailed information on each exam, including the grade. We consider the nine common exams taken during the first three semesters and group them into areas - economics, business, quantitative and other - as

described in Section 2. Suppose individual i chooses to specialize in economics, then g_i is computed as the ratio between i's average grade in the exams of the economics area over i's average grade in the exams of the business area. Similarly for students who chose business. We normalize also this measure by the relative performance of the full sample of students. Formally, g_i is defined as follows:

$$g_i = \begin{cases} \frac{GPA_i^{ECON}}{GPA_i^{BUSINESS}} \cdot \frac{\sum GPA_j^{BUSINESS}}{\sum GPA_j^{ECON}} \text{if} \quad ECON_i = 1 \\ \frac{GPA_i^{BUSINESS}}{GPA_i^{ECON}} \cdot \frac{\sum GPA_j^{ECON}}{\sum GPA_j^{BUSINESS}} \quad \text{if} \quad BUSINESS_i = 1 \end{cases} \tag{6}$$

where GPA_i^{ECON} is i's average grade in economics' exams and $GPA_i^{BUSINESS}$ is i's average grade in business' exams. If $g_i > 1$ it means that, during the first three semesters and compared to all other students, student i has performed better in the exams of the major she eventually chose as a specialization. Note that in constructing this indicator we only consider the common exams of the first three semesters, namely economics I and II for economics and management I, II and accounting for business. According to these indicators we define four groups of students. The first group, which we label *ability driven*, includes those students who chose the major subject in which they performed (relatively) better during the first three semesters against the (relative) majority of their peers, i.e. $f_i < 1$ and $g_i > 1$. The second group - the *peer driven* - are students who chose as the (relative) majority of their peers and against their (relative) revealed ability, i.e. $f_i > 1$ and $g_i < 1$. The third group - the *coherent* - includes those students who made a choice of major that is coherent with their performance as well as with their peers' behavior, i.e. $f_i > 1$ and $g_i > 1$. Finally, some students - the *incoherent* - chose against both their academic record and their peers, i.e. $f_i < 1$ and $g_i < 1$. Table 10 summarizes these definitions.

[TABLE 10]

As the table shows, students are rather evenly spread across the four groups. The largest (27.17%) is represented by the *ability driven*, i.e. students who choose against the relative majority of their peers and following the signal of their revealed performance. The *coherent*, i.e. students who choose both according to their ability and their peers, are only slightly less numerous (25.12%). *Peer driven* students, i.e. those who follow peers in contrast with the indication of their academic

performance, represent 23.76% of the sample, leaving a sizeable 23.95% of students in the group of the *incoherent*, i.e. those who choose against both peers and revealed ability.

We use these groups to estimate the effect of these "decision modes" on three academic outcomes: average grade in the last two and a half years of the degree (i.e. after the major choice is made), graduation mark and time to graduation. A general specification of the equations that we estimate in this section is the following:

$$y_i = c + \pi_1[\text{peer driven}]_i + \pi_2[\text{coherent}]_i + \pi_3[\text{incoherent}]_i + \boldsymbol{\vartheta}\mathbf{x}_i + u_i. \tag{7}$$

where y is the outcome considered and the other variables are dummies that identify the groups (with the *ability driven* kept as a reference group). The set of controls - $\mathbf{x}_i$ - includes a gender dummy, household income (as recorded at first enrolment), a dummy for students who reside outside the city of Milan, a set of dummies for the region of origin, a series of controls for academic performance and ability (high-school grades and type, average grades in the common exams, a dummy for the specialization and the number of common exams taken on the first available session).

[TABLE 11]

The results of this exercise are presented in Table 11. Columns 2 and 4 extend the specification for average grades in the non-common courses and graduation mark with time to graduation. In column 5, when we look at time to graduation, we replace the average grade in the common courses with the average grade in all courses. Notice that the maintained assumption is that, conditional on the observables, the four categories are independent from the outcome variable.[39]

Although the effect is small in magnitude, there is clear evidence that *peer driven* students on average perform worse than the *ability driven* in terms of both average and final grade, while there seems to be no detectable difference in time to graduation. We estimate a significant negative effect of -0.15 to -0.18 of a grade point on the average grade in non-common exams (exam grades are given on a scale from 0 to 30 with pass equal to 18) and of -0.57 to -0.62 on the final grade (given on a scale 0 to 110 with pass equal to 66).

[39] A basic version of a Conditional Independence Assumption (CIA), where selection is on observables and we can control for all those variables affecting both the decision mode and the outcomes considered.

6.1 Labor market effects

In this section we try to assign a 'price' to the decision of following one's peers in contrast with one's revealed ability (i.e. being a *peer driven* student as opposed to an *ability driven*) in terms of entry wages. The ideal strategy would be one in which entry wages for the same students used for the estimation of equation (7) are regressed on the dummies for the decision modes, controlling for a set of individual characteristics.

Unfortunately, information on wages is only available in a dataset constructed by Bocconi university by interviewing almost all its graduates between one and one and a half years after graduation and these surveys currently cover only those who graduated between 2000 and 2003.[40] Only for about 1/3 of the observations used in the previous sections of this paper it is possible to recover information on labour market outcomes from these surveys and this is obviously a very selected group of early graduates.

For these reasons, we take a different approach and merge academic records with all available surveys of graduates to compute the penalty associated with a lower graduation mark for the whole sample of Bocconi students who graduated between 2000 and 2003. The data on labour market outcomes include information on monthly wages in the first job, the type of occupation and contract and a number of questions on satisfaction with the university.

[TABLE 12]

In Table 12 we report the results of these estimates. In these regressions we are mostly interested in the coefficients on graduation mark but we also control for time to graduation and the entire set of ability measures and individual traits used throughout the paper. Moreover, since wages are recorded in intervals the results in Table 12 are produced with interval regressions.[41]

The results show a sharp discontinuity at the top of the distribution of graduation marks. When this variable is introduced linearly (column 1) the estimated effect is relatively small: a one point increase in the final grade raises monthly wages by a mere 6 euros (8 USD) per month - i.e. about

[40] At the time of the surveys (i.e. approximately 1.5 years after graduation), several male students were on compulsory military service and others (both male and female) could not be reached.

[41] The same results have been produced with alternative econometric specifications (i.e. linear OLS on the mid-points of the intervals, quantile regression, ordered probit) and the magnitude and significance of the estimated effects are extremely robust.

78 euros (100 USD) per year. However, this effect is much bigger for students obtaining full marks (i.e. 110 with or without honors), who earn almost 67 euros (86 USD) per month (871 euros - 1,117 USD - per year) more than students who just fail to get full marks.[42]

Furthermore, if we were to consider a constant life-time loss of those amounts we would get, on average, a net present value loss of (roughly) 2,100 USD.[43] Unfortunately we cannot test whether the penalty of a *peer driven* decision is constant over time since no other information on later wages is available at this time.

7 Conclusions

In this paper we investigate whether peers' behavior has an important and significant effect on the choice of college major using a unique dataset from Bocconi University. The available data and the peculiar structure of the degree allow us to identify the endogenous effect of peers on this decision, circumventing the two crucial identification problems of studies of social interactions: endogeneity and reflection.

Our contribution to the literature is twofold. First, we solve the long-standing identification problems in the estimation of social interactions. Second, we estimate the importance of peers' actions on one's choice of major. Further, we assign a monetary value on choosing a major based more on peers' behavior than on one's revealed ability and find a small negative impact that becomes sizeable at the "full mark" threshold.

There can be many possible mechanisms generating the endogenous peer effects. On the one hand, if students follow the choices of their peers simply because there is a utility gain in studying together, one may interpret the wage effect estimated in the last section as the monetary value of such utility advantage. Alternatively, if it is peer pressure or imitation that generates peer effects (Mas and Moretti, 2006), than this wage loss can effectively be interpreted as the cost of decisions that are not based exclusively on efficiency considerations. Finally, in the introduction we also suggested that peers may represent a source of useful information about some hard-to-see features

[42] These results are broadly consistent with similar estimates produced on a different data source, i.e. the Bank of Italy Survey of Household Incomes and Wealth.

[43] The net present value (NPV) has been computed by assuming a constant interest rate of 5 percent and a life-time of 40 years. For those students at the margin of getting full marks the NPV loss would be a quite large (roughly) 23,500 USD.

of university life and/or major choice (where to find the right material to study, which are the best or the easiest courses, the best teachers, etc.). Our estimates of the labor market effects suggest that this is unlikely to be the mechanism that generates peer effects in our study. Better informed individuals should in principle make better choices but the wage penalty associated with the peer driven students is in contrast with this interpretation, unless peers deliver incorrect information. It should also be noted that any combination of these explanations (and possibly others) may actually be at the origin of the effects that we estimate.

Having convincingly shown the existence of pure endogenous peer effects, as in this paper, understanding the exact mechanism that underlies social interactions is perhaps the next big open question in this branch of the literature.

References

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

[2] Altonji, J.G., Elder, T.E., Taber, C.R., (2005), "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools", Journal of Political Economy, 113 (1): 151-183.

[3] Ammermueller, A. and Pischke J.S., (2006), "Peer Effects in European Primary Schools: Evidecne from PIRLS", IZA Discussion Paper n. 2077.

[4] Bayer, P., Pintoff, R. and Pozen, D., (2004), "Building Criminal Capital Behind Bars: Peer Effects in Juvenile Corrections", mimeo, Duke University

[5] Bayer, P., Ross, S. and Topa, G., (2005), "Place of work and place of residence: Informal hiring networks and labor market outcomes", NBER working paper series no.11019

[6] Bobonis, G. and Finan, F. (2005), "Endogenous Social Interaction Effects in School Participation in Rural Mexico", mimeo, University of California Berkeley.

[7] Bramoullé, Y., Djebbari, H. and Fortin, B., (2006), "Identification of Peer Effects through Social Networks", mimeo, University of Laval.

[8] Brock, W. and Durlauf, S., (2001), "Interaction-based Models", Handbook of Econometrics, vol. 5, J. Heckman and Leamer E. (Eds), Amsterdam: North-Holland.

[9] Brock, W. and Durlauf, S., (2004). Identification of binary choice models with social interactions. Working Paper.

[10] Calvó-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. The American Economic Review 94 (3), 426.

[11] Calvó-Armengol, A., Patacchini, E. and Zenou, Y., (2006), "Peer Effects and Social Networks in Education", WP, Research Institute of Industrial Economics.

[12] Case, A. and Katz, L., (1991), "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths", NBER Working Paper: 3705.

[13] Cipollone, P. and Rosolia, A., (2006), "Social Interactions in High School: Lessons from an Earthquake", American Economic Review, forthcoming.

[14] Cooley, J., (2006), "Desegregation and the Achievement Gap: Do Diverse Peers Help", mimeo, Univesrity of Wisconsin-Madison.

[15] Duflo, E. and Saez, E. (2002), "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," unpublished working paper, MIT and University of California, Berkeley.

[16] Falk, A. and Ichino A., (2006), "Clean Evidence on Peer Effects", Journal of Labor Economics, 24(1): 39-57.

[17] Garibaldi, P., Giavazzi, F., Ichino, A. and Rettore, E., (2007), "College cost and time to complete a degree: Evidence from tuition discontinuities", mimeo.

[18] Graham, B., (2006), Identifying Social Interactions through Conditional Variance Restrictions, mimeo, UC Berkely.

[19] Graham, B. and Hahn, J., (2005), Identification and Estimation of the Linear-in-Means Model of Social Interactions, Economic Letters 88 (1): 1 - 6

[20] Hanushek, E., Kain, J., Markman, J. and Rivkin, S., (2003), "Does Peer Ability Affect Student Achievement?", Journal of Applied Econometrics 18(5): 527-544.

[21] Hoxby, C., (2000), "Peer Effects in the Classroom: Learning from Gender and Race Variation", NBER Working Paper: 7867.

[22] Ichino, A. and Maggi, G., (2000), "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm", Quarterly Journal of Economics, August, 115 (3), 1057-1090.

[23] Kling, J., Liebman, J. and Katz, L., (2005), "Experimental analysis of neighborhood effects", NBER Working Paper No. 11577

[24] Lalive, R. and Cattaneo, A., (2005), "Social Interactions and Schooling Decisions", Mimeo, University of Zurich.

[25] Laschever, R., (2005), "The doughboys network: social interactions and labor market outcomes. of World War I veterans", Unpublished manuscript

[26] Manski, C., (1993), "Identification of Endogenous Social Effects: The Reflection Problem", Review of Economic Studies, 60: 531-542.

[27] Mas, A. and Moretti, E., (2006), "Peers at Work", IZA Discussion Paper n. 2292.

[28] Miguel, E. and Kremer, M. (2004),"Worms: identifying impacts on education and health in the presence of treatment externalities," Econometrica, 72(1): 159-217.

[29] Moffitt, R., (2001), "Policy Interventions, Low-Level Equilibria, and Social Interactions", in Social Dynamics, S. Durlauf and H. P. Young eds., Cambridge: MIT Press.

[30] Sacerdote, B. (2001), "Peer Effects with Random Assignment: Results for Dartmouth Roommates," Quarterly Journal of Economics, 116: 681-704.

[31] Stutzer, A. and R. Lalive (forthcoming). The role of social work in job searching and subjective well-being. Journal of the European Economic Association.

[32] Soetevent, A. (2006), "Empirics of The Identification of Social Interactions: An Evaluation of the Approaches and their Results", Journal of Economic Surveys, 20(2): 193-228.

[33] Todd, P. and Wolpin, K., (2006), "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility", American Economic Review, 96(5):

[34] Zimmerman, D., (2003), "Peer Effects in Higher Education: Evidence from a Natural Experiment, Williams Project on the Economics of Higher Education", Review of Economics and Statistics, 85(1): 9-23.

Figure 1: Degree structure

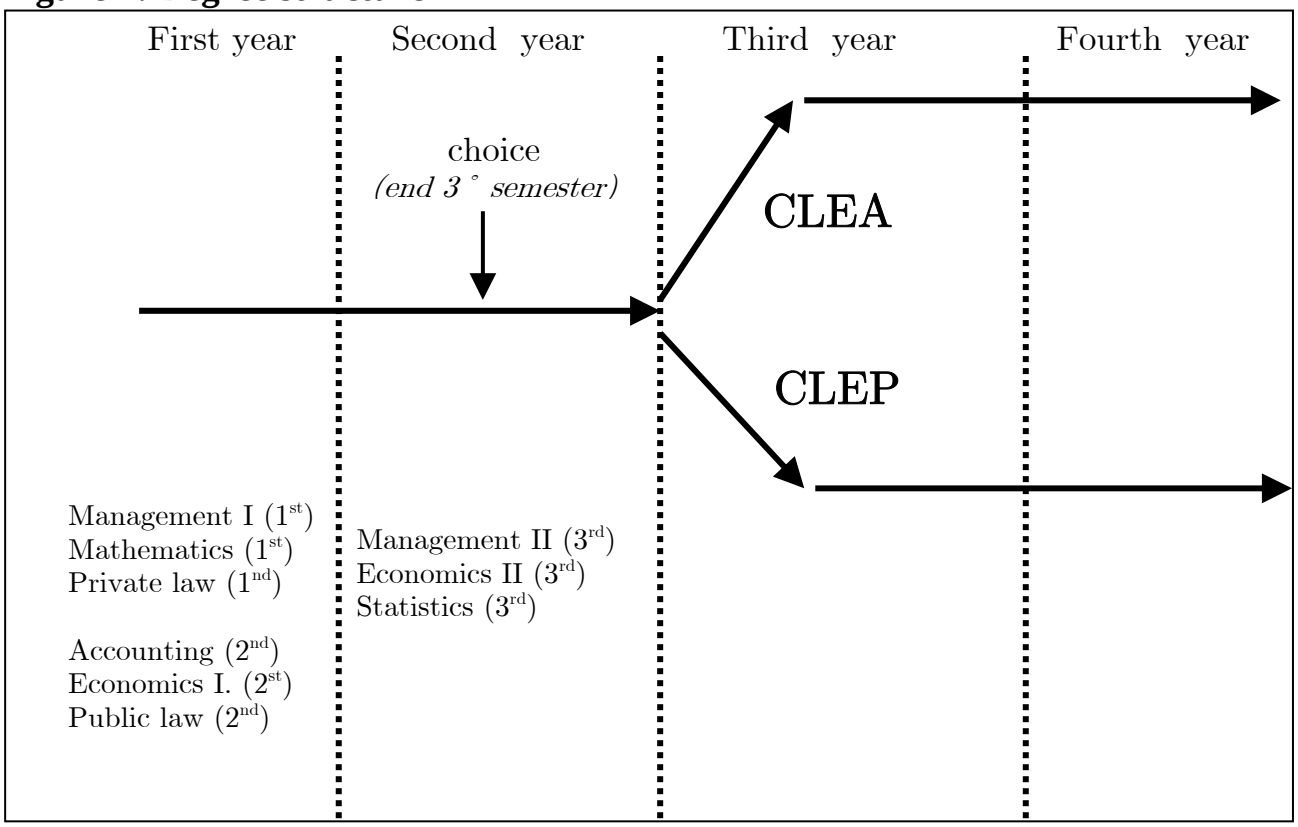

Table 1: Common exams CLEA/CLEP

	Semester	Area
Management I	1st	Business
Mathematics	1st	Quantitative
Private Law	1st	Law
Accounting	2nd	Business
Economics I	2nd	Economics
Public Law	2nd	Law
Economics II	3rd	Economics
Management II	3rd	Business
Statistics	3rd	Quantitative

Table 2: Descriptive statistics:

Variable	Mean	(s.d.)	min	max	Obs.
Individual characteristics					
1=CLEP	0.127	-	0	1	1141
1=female	0.396	-	0	1	1141
(log) household income[1]	7.91	(4.44)	0	11.7	1141
highest income braket[1]	0.227	(0.419)	0	1	1141
1=non-resident[2]	0.633	-	0	1	1141
1=determined economics[3]	0.15	-	0	1	1141
Academic measures					
Graduation mark[4]	102	(7.7)	76	111	1027
time to graduation (in years)[5]	5.34	(0.661)	4	7	1027
av. grade in all exams	26.2	(2.05)	20	30	1141
av. grade in common exams	24.8	(2.29)	19	30.3	1141
av. grade in quatitative common exams	23.7	(3.09)	18	31	1141
av. grade in economics common exams	24.7	(2.94)	18	31	1141
av. grade in business common exams	25.6	(2.49)	18	31	1141
admission test[6]	69.1	(7.42)	43	91	1141
high school final grade[7]	86.3	(11.2)	60	100	1141

Notes:

1. If a student declares that household income falls in the highest income bracket no further information is collected therefore household income is coded to 1 for households in the last bracket and an ad-hoc dummy controls for this group.
2. Resident outside the province of Milan.
3. DES as first or second preferred course in admission test courses' ranking
4. Range 0-111 (pass = 60).
5. Official duration is 4 years.
6. Normalised between 0 and 100.
7. Normalised between 0 and 100 (pass = 60).

Table 3: Characteristics of CLEA/CLEP students.

	Obs.	AVERAGE GRADE COMMON EXAMS				High School final grade	Admission Test final score
		Area Business	Area Economics	Area Quantitative	Total		
Total	1141	25.63	24.69	23.67	24.83	86.3	69.06
CLEP	145	26.82	26.79	25.81	26.52	92.2	72.48
CLEA	996	25.48	24.39	23.35	24.59	85.4	68.57
Difference (CLEP-CLEA)		1.36***	2.40***	2.46***	1.94***	6.79***	3.91***

Table 4: Characteristics of courses and lecturing classes:

	Semester	Number of classes	Characteristics	Average	coeff. of variation	Min	Max	Weight[3]
Management I	I	10	Enrolled students	140.40	0.11	130	169	0.70
			Student questionnaires	80.70	0.17	62	109	
			Average attendance[1] (%)	85.67	0.01	84.08	87.24	
			Congestion[2] (1 to 5)	3.33	0.05	3.16	3.61	
Mathematics	I	10	Enrolled students	140.80	0.12	125	164	0.55
			Student questionnaires	102.80	0.62	28	253	
			Average attendance[1] (%)	83.89	0.02	81.39	86.51	
			Congestion[2] (1 to 5)	3.77	0.14	3.00	4.57	
Private Law	I	4	Enrolled students	351.75	0.47	189	510	0.78
			Student questionnaires	70.00	0.39	38	104	
			Average attendance[1] (%)	79.73	0.06	74.91	83.89	
			Congestion[2] (1 to 5)	3.07	0.04	2.95	3.23	
Accounting	II	10	Enrolled students	142.80	0.33	109	258	0.57
			Student questionnaires	100.30	0.61	54	215	
			Average attendance[1] (%)	84.80	0.01	82.26	86.58	
			Congestion[2] (1 to 5)	3.46	0.14	3.02	4.40	
Economics I	II	6	Enrolled students	216.50	0.43	85	316	0.52
			Student questionnaires	136.83	0.76	24	317	
			Average attendance[1] (%)	84.92	0.01	83.56	86.84	
			Congestion[2] (1 to 5)	3.63	0.20	2.83	4.82	
Public Law	II	4	Enrolled students	351.75	0.42	217	528	0.83
			Student questionnaires	41.00	0.49	15	64	
			Average attendance[1] (%)	82.72	0.03	79.45	85.62	
			Congestion[2] (1 to 5)	2.89	0.06	2.67	3.03	
Economics II	III	6	Enrolled students	222.83	0.45	156	381	0.67
			Student questionnaires	109.17	0.48	19	176	
			Average attendance[1] (%)	83.87	0.02	81.42	86.80	
			Congestion[2] (1 to 5)	2.96	0.16	2.47	3.72	
Management II	III	8	Enrolled students	184.25	0.56	123	382	1.00
			Student questionnaires	80.75	0.32	56	125	
			Average attendance[1] (%)	84.38	0.01	83.38	85.27	
			Congestion[2] (1 to 5)	2.14	0.12	1.76	2.51	
Statistics	III	8	Enrolled students	272.25	0.33	142	404	0.56
			Student questionnaires	140.75	0.42	35	203	
			Average attendance[1] (%)	85.66	0.01	83.31	86.53	
			Congestion[2] (1 to 5)	3.27	0.29	2.09	4.46	

Notes:

1. Self reported by the students.
2. Congestion is defined from students evaluations as the average answer given to the following question: *"For your learning, the number of students attending your class has been: insufficient (1), too low (2), ideal (3), too high (4), excessive (5)"*.

3. Weight A is the ratio between the lowest maximum level of congestion (i.e. 2.51 for Managemetn II) and the maximum level of congestion across the classes of each course.

Table 5: Size of peer groups, various definitions.

		All peers[1] [1]	Restricted peers[2] [2]	Excluded peers[3] [3]
Raw group size	*Mean*	674.47	18.08	252.53
	Std. dev.	(79.10)	(6.77)	(60.96)
Average number of courses taken together	*Mean*	1.57	4.16	0.00
	Std. dev.	(0.06)	(0.11)	(0.00)
Weighted group size	*Mean*	151.07	10.77	--
	Std. dev.	(19.73)	(4.08)	--

1. Students who have been assigned to the same lecturing class at least once over the 7 common courses considered.
2. Students who have been assigned to the same lecturing class in at least 4 of the 7 common courses considered.
3. Students who have never been assigned to the same lecturing class in any of the 9 common courses but who have attended some of the 7 courses considered who at least one peer student.

Table 6: Peers and later academic patters

Definition of peers: number of courses attended in the same class	**at least 1** [1]	**at least 2** [2]	**at least 3** [3]	**at lest 4** [4]
Panel A: Percentage of students who choose the same sub-major				
Peers	9.645	9.685	9.908	9.633
Non-peers		9.603		
Diff.	*0.042*	*0.082**	*0.306****	*0.030*
Panel B: Percentage of students who graduate in the same session				
Peers	13.438	13.890	16.346	22.418
Non-peers		12.523		
Diff.	*0.915****	*1.367****	*3.823****	*9.895****

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Correlation between individual and group level predetermined variables

	All peers [1]	Restricted peers [2]
Panel A: correlation between individual and peer attributes:		
Admission test score	0.0052	0.0236
High school final grade	-0.0325	-0.0701
Determined economics	0.0181	0.0169
Panel B: correlation between individual and excluded peers' attributes:		
Admission test score	-0.0474	-0.0488
High school final grade	-0.0192	-0.0169
Determined economics	-0.0050	-0.0079
Panel C: correlation between peers' and excluded peers' attributes:		
Admission test score	-0.5054***	-0.1088***
High school final grade	-0.4113***	-0.0655***
Determined economics	-0.4931***	-0.0785***

Table 8: Peer effects in the choice of major. Linear probability model

Dependent variable: probability of choosing CLEP	All peers			Restricted peers		
	OLS	**2SLS[1]**	**2SLS[1] weighted**	**OLS**	**2SLS[1]**	**2SLS[1] weighted**
	[1]	[2]	[3]	[4]	[5]	[6]
Fraction Peers choosing CLEP	1.000	5.785*	5.289*	0.150	1.260*	1.242*
	(0.768)	(3.171)	(2.899)	(0.105)	(0.698)	(0.643)
	[0.193]	[0.068]	[0.068]	[0.155]	[0.071]	[0.053]
Individual characteristics						
Admission test[2]	0.005**	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
High school final grade[3]	0.004***	0.004***	0.004***	0.424***	0.410***	0.416***
	(0.001)	(0.001)	(0.001)	(0.102)	(0.106)	(0.106)
1=determined economics	0.095**	0.089**	0.089**	0.098***	0.112***	0.112***
	(0.031)	(0.032)	(0.031)	(0.031)	(0.034)	(0.033)
1=female	-0.011	-0.012	-0.012	-0.012	-0.016	-0.016
	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)	(0.021)
Log household income[4]	0.005	0.006	0.006	0.005	0.006	0.006
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
1=highest income bracket[4]	0.044	0.054	0.053	0.043	0.058	0.057
	(0.064)	(0.065)	(0.064)	(0.065)	(0.069)	(0.069)
1=non resident[5]	0.001	-0.003	-0.002	-0.000	-0.012	-0.012
	(0.025)	(0.026)	(0.026)	(0.025)	(0.028)	(0.027)
High school type dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region of residence dummies	Yes	Yes	Yes	Yes	Yes	Yes
Nr. Obs.	1141	1141	1141	1141	1141	1141
R^2	0.12	0.09	0.09	0.12	0.16	0.16
Shea Partial R^2	--	0.0603	0.0708	--	0.0242	0.0304
1st stage F-test	--	25.46	30.30	--	8.49	12.17

Note:

1. Excluded instruments: averages of admission test, high school final grade, determined to do economics in the group of excluded peers who are not in one's peer group.
2. Normalised between 0 and 100. Average in the sample = 69.10
3. Normalised between 0 and 100 (pass = 60). Average in the sample = 86.3
4. If a student declares that household income falls in the highest income bracket no further information is collected therefore household income is coded to 1 for households in the last bracket and an ad-hoc dummy controls for this group.
5. Resident outside the province of Milan.

Robust standard errors in parentheses: p-values in square brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 9: Peer effects in the choice of major. Probit model

	All peers			Restricted peers		
Dependent variable:	**Probit**	**IVProbit**	**IVProbit weighted**	**Probit**	**IVProbit**	**IVProbit weighted**
Probability of choosing CLEP	[1]	[2]	[3]	[4]	[5]	[6]
Fraction Peers choosing CLEP	5.30	32.747**	30.873**	0.756	6.402***	6.059***
	(4.120)	(14.134)	(13.308)	(0.548)	(2.404)	(2.257)
	[0.198]	[0.021]	[0.020]	[0.167]	[0.008]	[0.007]
marginal effects	0.905	6.154	5.723	0.129	1.351	1.248
Individual characteristics						
Admission test[2]	0.022**	0.023***	0.023***	0.022***	-0.000	-0.000
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
High school final grade[3]	0.028***	0.025***	0.026***	2.748***	0.014	0.011
	(0.006)	(0.006)	(0.006)	(0.637)	(0.697)	(0.673)
1=determined economics	0.437***	0.377**	0.384**	0.455***	-0.013*	0.473***
	(0.127)	(0.132)	(0.131)	(0.127)	(0.123)	(0.124)
1=female	-0.069	-0.068	-0.069	-0.072	0.003	0.004
	(0.110)	(0.106)	(0.106)	(0.111)	(0.102)	(0.103)
Log household income[4]	0.044	0.042	0.042	0.044	-0.001	0.038
	(0.041)	(0.038)	(0.039)	(0.042)	(0.036)	(0.037)
1=highest income bracket[4]	0.418	0.416	0.413	0.423	-0.016	-0.016
	(0.489)	(0.454)	(0.456)	(0.489)	(0.429)	(0.436)
1=non resident[5]	0.044	0.019	0.021	0.042	0.008	0.008
	(0.161)	(0.157)	(0.158)	(0.162)	(0.152)	(0.153)
High school type dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region of residence dummies	Yes	Yes	Yes	Yes	Yes	Yes
Nr. Obs.	1124	1124	1124	1124	1124	1124
Pseudo R^2	0.146	--	--	0.146	--	--

Note:
1. Excluded instruments: averages of admission test, high school final grade, determined to do economics in the group of excluded peers who are not in one's peer group.
2. Normalised between 0 and 100. Average in the sample = 69.10
3. Normalised between 0 and 100 (pass = 60). Average in the sample = 86.3
4. If a student declares that household income falls in the highest income bracket no further information is collected therefore household income is coded to 1 for households in the last bracket and an ad-hoc dummy controls for this group.
5. Resident outside the province of Milan.
Robust standard errors in parentheses: p-values in square brackets.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Distribution of decision modes

		ABILITY INFLUENCE	
		YES ($g>1$)	NO ($g<1$)
PEERS' INFLUENCE	YES ($f>1$)	Coherent 25.12%	Peer driven 23.76%
	NO ($f<1$)	Ability driven 27.17%	Incoherent 23.95%

Table 11: Decision modes and academic outcomes

Dependent variable:	Av. Grade in non-common exams[1]		Graduation mark[2]		Time to graduation[3] (in years)
	[1]	[2]	[3]	[4]	[5]
Decision mode					
Peer driven	-0.170*	-0.154*	-0.618**	-0.575**	0.025
	(0.056)	(0.073)	(0.036)	(0.045)	(0.623)
Coherent	-0.023	-0.060	-0.049	-0.145	-0.080
	(0.794)	(0.479)	(0.867)	(0.614)	(0.122)
Incoherent	-0.205**	-0.146*	-0.565*	-0.409	0.112**
	(0.024)	(0.097)	(0.058)	(0.161)	(0.042)
Ability measures					
Av. grade all exams	--	--	--	--	-0.167***
					(0.000)
Av. grade common exams	0.644***	0.595***	2.877***	2.751***	--
	(0.000)	(0.000)	(0.000)	(0.000)	
Time to graduation	--	-0.479***	--	-1.252***	--
		(0.000)		(0.000)	
1=CLEP	-0.262***	-0.216**	-0.201	-0.083	0.087
	(0.003)	(0.012)	(0.529)	(0.798)	(0.119)
Admission test[4]	0.006	0.005	-0.021	-0.022	0.002
	(0.228)	(0.247)	(0.200)	(0.167)	(0.583)
High school final grade[5]	0.025***	0.024***	0.075***	0.072***	0.021
	(0.000)	(0.000)	(0.000)	(0.000)	(0.242)
High school type dummies	yes	yes	yes	yes	yes
Individual characteristics					
1=female	0.341***	0.267***	1.046***	0.853***	-0.113***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
Household income[6]	-0.011	-0.017	-0.052	-0.067	-0.013
	(0.730)	(0.543)	(0.634)	(0.494)	(0.385)
1=highest income bracket[6]	-0.030	-0.164	-0.414	-0.763	-0.284
	(0.935)	(0.611)	(0.746)	(0.508)	(0.120)
1=non resident[7]	0.086	0.089	0.271	0.280	0.023
	(0.194)	(0.155)	(0.219)	(0.188)	(0.691)
Region of residence dummies	yes	yes	yes	yes	yes
Nr. Observations	1027	1027	1027	1027	1027
R-squared	0.513	0.529	0.566	0.573	0.163

Note:
1. Range 0-30 (18 = pass). Average in the sample = 26.97
2. Range 0-111 (pass = 60). Average in the sample = 102.11
3. Official duration is 4 years. Average in the sample = 5.34
4. Normalised between 0 and 100. Average in the sample = 69.10
5. Normalised between 0 and 100 (pass = 60). Average in the sample = 86.3
6. If a student declares that household income falls in the highest income bracket no further information is collected therefore household income is coded to 1 for households in the last bracket and an ad-hoc dummy controls for this group.
7. Resident outside the province of Milan.
Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 12: Interval wage regressions

Dependent variable: **wage in the first job**[1]	[1]	[2]
graduation mark[2]	6.045***	3.718**
	(1.360)	(1.612)
1=full marks[3]		66.881***
		(25.013)
time to graduation[4]	-2.450*	-2.279
	(1.443)	(1.443)
1=female	-97.039***	-94.362***
	(17.360)	(17.368)
Household income[5]	-0.000	-0.000
	(0.000)	(0.000)
1=highest income bracket[5]	-2.276	-3.612
	(26.449)	(26.428)
1=post-graduate studies	-19.498	-18.686
	(19.099)	(19.078)
High school final grade[6]	-1.093	-1.214
	(0.893)	(0.894)
High school type dummies	yes	yes
Degree programme dummies	yes	yes
Contract type dummies	yes	yes
Nr. observations	3982	3982

Note:
1. Recorded in intervals
2. Range 0-111 (pass = 66).
3. 110 with or without honours
4. Recorded in quarters. Official duration is 4 years.
5. If a student declares that household income falls in the highest income bracket no further information is collected therefore household income is coded to -1 for households in the last bracket and an ad-hoc dummy controls for this group.
6. Normalised between 0 and 1 (pass = 0.6).
Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table A.1: Characteristics of the teaching classes

Course name	Semester	Class number[1]	Enrolled students	Number of completed student questionnaires	[5]/[4]	Attendance[2] (%)	Congestion[3] (1 to 5)
[1]	[2]	[3]	[4]	[5]	[6]	[8]	[7]
Management	I	4	169	109	0.64	86.28	3.61
Management	I	57	130	62	0.48	84.11	3.19
Management	I	20	131	67	0.51	87.23	3.27
Management	I	51	134	80	0.60	85.78	3.40
Management	I	30	133	68	0.51	87.24	3.51
Management	I	43	135	92	0.68	84.84	3.18
Management	I	7	134	77	0.57	85.70	3.21
Management	I	56	134	81	0.60	86.19	3.33
Management	I	84	136	85	0.63	84.08	3.39
Management	I	26	168	86	0.51	85.29	3.16
Average			**140.40**	**80.70**	**0.57**	**85.67**	**3.33**
Coeff. of variation			**0.11**	**0.17**	**0.12**	**0.01**	**0.05**
Mathematics I	I	50	160	147	0.92	83.56	4.18
Mathematics I	I	49	156	87	0.56	84.72	4.21
Mathematics I	I	56	125	35	0.28	83.30	3.00
Mathematics I	I	35	127	92	0.72	83.57	3.82
Mathematics I	I	35	128	109	0.85	82.81	4.15
Mathematics I	I	60	164	119	0.73	85.46	3.80
Mathematics I	I	55	128	69	0.54	86.51	3.35
Mathematics I	I	2	128	89	0.70	85.01	3.56
Mathematics I	I	55	131	28	0.21	81.39	3.11
Mathematics I	I	12	161	253	1.57	82.53	4.57
Average			**140.80**	**102.80**	**0.71**	**83.89**	**3.77**
Coeff. of variation			**0.12**	**0.62**	**0.53**	**0.02**	**0.14**
Private Law	I	32	510	104	0.20	83.89	3.23
Private Law	I	14	475	71	0.15	74.91	3.00
Private Law	I	39	233	67	0.29	83.27	3.12
Private Law	I	94	189	38	0.20	76.84	2.95
Average			**351.75**	**70.00**	**0.21**	**79.73**	**3.07**
Coeff. of variation			**0.47**	**0.39**	**0.27**	**0.06**	**0.04**
Accounting	II	74	258	215	0.83	82.26	4.08
Accounting	II	88	144	55	0.38	84.05	3.07
Accounting	II	73	164	83	0.51	84.11	3.24
Accounting	II	56	178	211	1.19	83.81	4.40
Accounting	II	64	136	54	0.40	85.22	3.02
Accounting	II	51	110	98	0.89	85.47	3.60
Accounting	II	94	110	57	0.52	85.06	3.02
Accounting	II	26	110	88	0.80	86.58	3.72
Accounting	II	41	109	74	0.68	86.01	3.31
Accounting	II	27	109	68	0.62	85.43	3.12
Average			**142.80**	**100.30**	**0.68**	**84.80**	**3.46**
Coeff. of variation			**0.33**	**0.61**	**0.37**	**0.01**	**0.14**
Economics I	II	9	280	111	0.40	86.84	3.87
Economics I	II	77	290	175	0.60	83.56	3.84
Economics I	II	41	316	317	1.00	84.78	4.82
Economics I	II	22	85	24	0.28	85.92	2.83
Economics I	II	82	184	138	0.75	84.31	3.40

Economics I	II	35	144	56	0.39	84.12	3.02
Average			**216.50**	**136.83**	**0.57**	**84.92**	**3.63**
Coeff. of variation			**0.43**	**0.76**	**0.47**	**0.01**	**0.20**
Public Law	II	39	528	44	0.08	82.65	2.89
Public Law	II	5	419	41	0.10	79.45	2.97
Public Law	II	51	243	15	0.06	83.17	2.67
Public Law	II	9	217	64	0.29	85.62	3.03
Average			**351.75**	**41.00**	**0.13**	**82.72**	**2.89**
Coeff. of variation			**0.42**	**0.49**	**0.80**	**0.03**	**0.06**
Economics II	III	58	160	110	0.69	84.14	3.33
Economics II	III	59	315	176	0.56	83.95	3.72
Economics II	III	90	381	142	0.37	84.92	2.91
Economics II	III	51	156	19	0.12	81.42	2.47
Economics II	III	70	163	106	0.65	86.80	2.73
Economics II	III	2	162	102	0.63	81.99	2.61
Average			**222.83**	**109.17**	**0.50**	**83.87**	**2.96**
Coeff. of variation			**0.45**	**0.48**	**0.43**	**0.02**	**0.16**
Management II	III	85	319	113	0.35	84.84	2.51
Management II	III	4	382	125	0.33	84.48	2.43
Management II	III	38	123	66	0.54	83.38	2.28
Management II	III	19	125	61	0.49	84.12	1.97
Management II	III	53	133	56	0.42	84.30	1.98
Management II	III	74	133	91	0.68	84.91	1.76
Management II	III	67	125	65	0.52	85.27	2.11
Management II	III	14	134	69	0.51	83.70	2.07
Average			**184.25**	**80.75**	**0.48**	**84.38**	**2.14**
Coeff. of variation			**0.56**	**0.32**	**0.24**	**0.01**	**0.12**
Statistics	III	77	370	157	0.42	83.31	3.54
Statistics	III	75	142	35	0.25	86.53	2.09
Statistics	III	4	404	203	0.50	86.09	2.86
Statistics	III	6	240	172	0.72	85.97	4.27
Statistics	III	74	248	157	0.63	86.21	4.02
Statistics	III	64	192	64	0.33	85.66	2.24
Statistics	III	57	336	180	0.54	85.18	2.65
Statistics	III	63	246	158	0.64	86.36	4.46
Average			**272.25**	**140.75**	**0.50**	**85.66**	**3.27**
Coeff. of variation			**0.33**	**0.42**	**0.32**	**0.01**	**0.29**

Note:

1. The true class identifier has been randomised to avoid identification of the individual teachers.

2. Self reported by the students.

3. Congestion is defined from students evaluations as the average answer given to the following question: *"For your learning, the number of students attending your class has been: insufficient (1), too low (2), ideal (3), too high (4), excessive (5)"* .

Table A.2: IV First-stage regressions

Dependent variable: fraction of peers choosing CLEP	All peers	All peers weighted	Restricted peers	Restricted peers weighted
	[1]	[2]	[3]	[4]
Instruments: Excluded peers'				
Admission test[2]	-0.348***	-0.357***	0.021***	0.023***
	(0.075)	(0.074)	(0.007)	(0.006)
High school final grade[3]	0.000***	0.000***	-0.333	-0.697
	(0.000)	(0.000)	(0.567)	(0.487)
Fraction of determined economics	-0.132***	-0.140***	-0.617***	-0.463***
	(0.018)	(0.018)	(0.136)	(0.106)
Individual characteristics				
Admission test[2]	-0.000*	-0.000*	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
High school final grade[3]	0.000	0.000	0.012	0.009
	(0.000)	(0.000)	(0.028)	(0.028)
1=determined economics	0.001	0.000	-0.012*	-0.013*
	(0.001)	(0.001)	(0.007)	(0.007)
1=female	-0.000	0.000	0.003	0.003
	(0.001)	(0.000)	(0.006)	(0.005)
Log household income[4]	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.002)	(0.002)
1=highest income bracket[4]	-0.003	-0.003	-0.014	-0.015
	(0.004)	(0.004)	(0.022)	(0.022)
1=non resident[5]	0.000	0.000	0.008	0.008
	(0.001)	(0.001)	(0.008)	(0.008)
High school type dummies	Yes	Yes	Yes	Yes
Region of residence dummies	Yes	Yes	Yes	Yes
Nr. Obs.	1141	1141	1141	1141

Note:

1. Excluded instruments: averages of admission test, high school final grade, determined to do economics in the group of excluded peers who are not in one's peer group.
2. Normalised between 0 and 100. Average in the sample = 69.10
3. Normalised between 0 and 100 (pass = 60). Average in the sample = 86.3
4. If a student declares that household income falls in the highest income bracket no further information is collected therefore household income is coded to 1 for households in the last bracket and an ad-hoc dummy controls for this group.
5. Resident outside the province of Milan.

Robust standard errors in parentheses: p-values in square brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A.3: Selected descriptive statistics of individual and peers' characteristics

Variable		Individual	All peers	Restricted peers	Excluded peers (all)	Excluded peers (restricted)
1=CLEP	*Mean*	0.127	0.127	0.128	0.123	0.123
	Std. dev	(0.333)	(0.012)	(0.088)	(0.019)	(0.019)
Determined economics	*Mean*	0.147	0.146	0.146	0.144	0.144
	Std. dev	(0.354)	(0.128)	(0.094)	(0.019)	(0.019)
Admission test[1]	*Mean*	69.1	69.1	69.1	69.1	69.0
	Std. dev	(7.4)	(0.3)	(1.9)	(0.5)	(0.5)
High school final grade[2]	*Mean*	86.3	86.4	86.4	86.3	86.3
	Std. dev	(11.2)	(0.3)	(2.7)	(0.5)	(0.5)

Table A.4: Robustness checks (Linear probability models)

	All peers			Restricted peers		
	OLS	**2SLS**	**2SLS weighted**	**OLS**	**2SLS**	**2SLS weighted**
	[1]	[2]	[3]	[4]	[5]	[6]
1. Baseline	1.000	5.785*	5.289*	0.150	1.260*	1.242*
	(0.768)	(3.171)	(2.899)	(0.105)	(0.698)	(0.643)
	[0.193]	[0.068]	[0.068]	[0.155]	[0.071]	[0.053]
2. Exogenous effects	--	--	--	0.143	1.317*	1.346*
				(0.107)	(0.746)	(0.699)
				[0.184]	[0.077]	[0.054]
3. Groups based on 5 exams	1.372**	4.379*	3.930	0.215	1.648**	1.643*
	(0.661)	(2.621)	(2.398)	(0.133)	(0.828)	(0.841)
	[0.038]	[0.095]	[0.101]	[0.107]	[0.047]	[0.051]
4. Exponential weights	0.516*	3.137**	--	0.113	1.247*	--
	(0.296)	(1.588)		(0.093)	(0.710)	
	[0.081]	[0.048]		[0.223]	[0.079]	
5. Placebo peer groups	0.822	0.466	0.149	-0.081	0.440	-0.287
	(0.864)	(4.646)	(4.722)	(0.058)	(0.674)	(0.794)
	[0.342]	[0.920]	[0.975]	[0.163]	[0.513]	[0.718]
6. Teacher quality controls	0.505	5.725*	5.231*	0.099	1.207*	1.109*
	(0.807)	(3.400)	(3.134)	(0.107)	(0.727)	(0.652)
	[0.531]	[0.092]	[0.095]	[0.357]	[0.097]	[0.089]
7. Course congestion	0.505	5.725*	5.231*	0.154	1.248*	1.234*
	(0.807)	(3.400)	(3.134)	(0.105)	(0.687)	(0.640)
	[0.531]	[0.092]	[0.095]	[0.145]	[0.069]	[0.053]

Robust standard errors in parentheses: p-values in square brackets.

Table A.5: Robustness checks (Probit models)

	All peers			Restricted peers		
	Probit	**IVProbit**	**IVProbit weighted**	**Probit**	**IVProbit**	**IVProbit weighted**
	[1]	[2]	[3]	[4]	[5]	[6]
1. Baseline	5.30	32.747**	30.873**	0.756	6.402***	6.059***
	(4.120)	(14.134)	(13.308)	(0.548)	(2.404)	(2.257)
	[0.198]	[0.021]	[0.020]	[0.167]	[0.008]	[0.007]
marginal effect	0.905	6.154	5.723	0.129	1.351	1.248
2. Exogenous effects	--	--	--	0.823	6.707***	6.594***
				(0.569)	(2.610)	(2.380)
				[0.148]	[0.010]	[0.006]
marginal effects				0.140	1.405	1.370
3. Groups based on 5 exams	7.224**	26.110**	24.053**	1.074	8.227***	8.413***
	(3.486)	(12.587)	(11.947)	(0.698)	(2.707)	(2.686)
	[0.038]	[0.038]	[0.044]	[0.124]	[0.002]	[0.002]
marginal effects	1.226	4.732	4.297	0.183	1.762	1.823
4. Exponential weights	2.937*	16.636***	--	0.610	6.206***	--
	(1.535)	(6.202)		(0.481)	(2.197)	
	[0.056]	[0.007]		[0.205]	[0.005]	
marginal effect	0.499	3.290		0.104	1.376	
5. Placebo peer groups	5.267	4.314	2.262	-0.459	2.755	-2.015
	(4.423)	(27.703)	(28.192)	(0.334)	(4.497)	(5.127)
	[0.234]	[0.876]	[0.936]	[0.170]	[0.540]	[0.694]
marginal effect	0.898	0.735	0.386	-0.078	0.576	-0.361
6. Teacher quality controls	2.565	33.423**	31.629**	0.497	6.297**	5.691**
	(4.453)	(14.904)	(14.023)	(0.562)	(2.564)	(2.397)
	[0.565]	[0.025]	[0.024]	[0.377]	[0.014]	[0.018]
marginal effect	0.428	6.232	5.819	0.083	1.308	1.132
7. Course congestion	4.988	31.201**	28.619**	0.778	6.330***	6.035***
	(4.113)	(13.620)	(12.833)	(0.549)	(2.391)	(2.253)
	[0.225]	[0.022]	[0.026]	[0.156]	[0.008]	[0.007]
marginal effect	0.851	5.834	5.255	0.133	1.328	1.240

Robust standard errors in parentheses: p-values in square brackets.