NBER WORKING PAPER SERIES

BE AS CAREFUL OF THE COMPANY YOU KEEP AS OF THE BOOKS YOU READ:
PEER EFFECTS IN EDUCATION AND ON THE LABOR MARKET

Giacomo De Giorgi
Michele Pellizzari
Silvia Redaelli

Be as Careful of the Company You Keep as of the Books You Read: Peer Effects in Education and on the Labor Market
Giacomo DeGiorgi, Michele Pellizzari, and Silvia Redaelli
NBER Working Paper No. 14948
May 2009
JEL No. I21,J0

ABSTRACT

In this paper we investigate whether peers' behavior influences the choice of college major, thus contributing to the mismatch of skills in the labor market. Using a newly constructed dataset, we are able to identify the endogenous effect of peers on such decisions through a novel identification strategy that 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 also provide evidence on skills mismatch in terms of entry wages and occupation. We find that peers can divert students from majors in which they have a relative ability advantage, with adverse consequences on academic performance, entry wages and job satisfaction.

Giacomo DeGiorgi
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
and NBER
degiorgi@stanford.edu

Michele Pellizzari
Bocconi University
Department of Economics and IGIER
via Roentgen 1
20136 - Milan - Italy
michele.pellizzari@unibocconi.it

Silvia Redaelli
1818 H Street, NW
Washington, DC 20433 USA
sredaelli@worldbank.org

1   Introduction

The importance of peers in shaping individual behavior has been widely recognized in both the economic and the sociological literature. Numerous studies have produced empirical evidence documenting the existence of relevant peer effects in many areas, from schooling performance to criminal behavior and financial decisions (Katz and Case, 1991; Hoxby, 2000; Sacerdote, 2001; Duflo and Saez, 2003; Ammermueller and Pischke, 2006).

Despite the vast literature on the topic, 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; Soetevent, 2006).

The contribution of this paper is twofold. First, on the methodological side, we develop a new general strategy for the estimation of peer effects and particularly of their endogenous component, usually defined as the impact of the average peers' outcome on individual outcome. The many hurdles involved in the identification of such parameter have somewhat concealed its importance for policy interventions. The extent and power of endogenous interactions are, in fact, crucial to determine the magnitude of the effect of any manipulation of exogenous or group characteristics. Moreover, large-scale policy interventions that are likely to generate general equilibrium effects cannot be properly evaluated without knowledge of the endogenous interaction parameter. Finally, and perhaps most importantly, the literature already offers numerous examples of policy interventions that consist of a direct manipulation of the endogenous effect. This is the case, for instance, of the immunization program implemented in Kenya and analyzed in Miguel and Kremer (2004) or the PROGRESA program in Mexico (Todd and Wolpin, 2006).

The second contribution of the paper concerns the particular outcome considered: students' choices of college major. The analysis of social interactions in education has so far been focused almost exclusively on academic performance. The only exception, to our knowledge, is Sacerdote (2001), who does not find any significant influence of peers on major choice.[1]

Social interactions, and particularly the endogenous influence of peers, are likely to be crucial in determining students' choices above and beyond their documented effect on performance. For

---

[1]Notice, however, that in Sacerdote's paper, endogenous and exogenous effects could not be separated.

example, when deciding whether to go to college, which particular school to attend or in which field to specialize, peers may provide precious and otherwise costly information. Moreover, to many students, classmates also represent a natural reference group and conformist behaviors are extremely common during adolescence (see Akerlof and Kranton, 2002). Finally, there might simply be a utility gain from attending school or courses with one's friends.

Peer effects in schooling choices can be particularly important in explaining why such decisions are often somewhat inefficient, as documented in several branches of the economic literature. For example, recent evidence in the large literature on wage inequality (Goldin and Katz, 2007) suggests that most of the dynamics in wage inequality in the 1980s and 1990s has been driven by the slow response of (skilled) labor supply rather than by an acceleration in the demand for skills. Although our focus is on major choice and not specifically on college enrollment, these two decisions should not be fundamentally different, particularly regarding the role of peers.

Another dimension of inefficient schooling decisions, more closely connected to the choice of major, is skill mismatch. The fact that large fractions of workers are employed in jobs unrelated to their studies has been documented already some time ago (Farber, 1999), even if the issue has remained largely unexplored. Among the few recent contributions, Gottschalk and Hansen (2003) analyze the trends in college graduates that are occupied in "non-college" jobs, while Robst (2007) shows that being employed in a job outside one's major field negatively and significantly affects the returns to schooling. Bender and Heywood (2006) provide similar evidence for a sample of PhD graduates in the US.

Correctly understanding the dynamics of schooling choices is, thus, crucial for any policy intervention aimed at reducing these inefficiencies and peer effects are likely to play a key role in this framework.

In this paper we apply our identification strategy to the estimation of peer effects in the choice of major and we show that peers' influence can indeed lead students to make schooling decisions that contrast with their revealed abilities. Further, by observing students in their first job after graduation we are also able to estimate the effect of different major decision modes on wages and on the (self-reported) probability of job-mismatch. In particular, we find that those students who choose a major following their peers and in contrast to their revealed ability graduate with lower final marks, end up in lower paid jobs and are more likely to be mismatched.

Our findings do not directly support any specific policy intervention, with the possible exception of providing better information to students about the available schooling options. However, we contribute substantially to the understanding of the mechanism that underlies a large set of schooling choices, thus allowing a better design and evaluation of any policy aimed at improving the allocation of skills to jobs.

Our study is based on a newly constructed set of administrative data on undergraduate students from Bocconi University. In these data, groups of peers are nicely defined by institutional features. At the time covered by our data, Bocconi students initially enrolled in a common program and only at the end of their third semester did they choose whether to specialize in one of two majors: business or economics. During the first three semesters, all students took nine compulsory courses and attended lectures in randomly assigned classes for each course. This process of repeated random allocation naturally defines groups of peers that vary at the level of each single student and are thus immune from simultaneity (or reflection). Moreover, since the allocation into the classes is random, endogeneity in peer group formation is also excluded by construction.[2]

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 student i but did attend some courses with some of i's peers. The exogenous characteristics of such excluded peers are a natural set of instruments to overcome potential endogeneity generated by common (correlated) group effects.[3]

Individual-specific groups and the use of excluded peers as instruments define a new strategy for the correct identification of social interactions that, far from being peculiar to our specific setting, can be applied to numerous other contexts. In the real world one rarely belongs to only one fixed group of peers. On the contrary, we all have friends, mates and colleagues that may overlap, but very rarely perfectly coincide. Whenever the units of analysis are linked directly to some other units (the peers) but only indirectly (through peers) connected to others (the peers of peers), our identification could be fruitfully adopted. In the network literature (Calvó *et al.*, 2004; Jackson, 2006) this corresponds to the existence of links of (at least) diameter 2.

The spectrum of applications of our identification strategy is, therefore, potentially larger than

---

[2]Note that the exogenous formation of groups only simplifies our analysis, but it is not a necessary feature of the novel identification strategy that we propose. In principle, our approach can be applied to settings in which groups form endogenously, in which case one would have to take that process into account.

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

that of current empirical approaches for the estimation of peer effects. Perhaps, a limitation might come from the stricter data requirements. Our strategy, in fact, requires knowledge of all (or many) indirect links among peers of peers. However, data of this type are becoming more and more popular and easily available. In fact, it is likely that in many existing datasets our approach could be readily replicated. Some recent papers could already exploit our identification with no additional data requirement, like Bayer *et al.* (2004), a study of criminal behavior where peers' characteristics are weighted by the time spent in the same correctional facility; Calvó *et al.* (2006), who focus on the influence of the position of a player in a particular network on academic performance; Guiso and Schivardi (2007) who look at spillovers in employment and investments of firms located in the same district and Mas and Moretti (2008) in their analysis of social interactions among supermarket cashiers.

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), fixed effects (Hanushek *et al.*, 2003) or higher order moments (Graham, 2006). Laschever (2005) is, to our knowledge, the first application of a multiple group framework. In a paper developed independently and at the same time as ours, Bramoullé *et al.* (2006) discuss identification of peer effects in a network framework that is very close to ours.

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 final academic and labor market outcomes. Finally, Section 7 concludes.

2 Data and institutional details[4]

The analysis in this paper is based on administrative data from Bocconi University, and on a series of surveys conducted by the university on its recent graduates to collect information concerning their labor market outcomes. Our analysis uses information on one entire cohort of students who entered University in the academic year 1998/99 and enrolled in the CLEA/CLEP program, the most popular degree offered by Bocconi at that time.

Students in this program would first follow a common track of nine exams during the initial 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]

Excluding a few missing values for 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.[5] A few of them (less than 10%) have not graduated, either because they dropped out, moved to another university or are still enrolled and trying to graduate. The distribution across majors is strongly skewed towards business with only about 13% of students choosing CLEP.

[TABLE 2]

Table 3 reports some descriptive statistics on the ability and performance of these two groups of students. Considering all the common exams in the first three semesters, the 145 students choosing CLEP score on average almost 2 grade-points above the CLEA students (exams are graded on a

---

[4]In this section we describe only the most important features of the data and the institutional background. Further details can be found in Appendix A.

[5]We also exclude students transferring from other universities and students from abroad who were given reserved places.

scale of 0 to 30, with a passing grade equal to 18). This difference is even higher when the exams are disaggregated by field. As expected, CLEP students perform relatively better in economics and quantitative subjects, while the difference is considerably smaller for the average grade in business courses.

[TABLE 3]

Further, we exploit a number of surveys conducted by the university on its alumni, who are contacted and interviewed about 1.5 years after graduation. Unfortunately, the response rates are not particularly high, especially for the earlier years and for males, who often did their compulsory military service right after graduation.[6] For these reasons, we can recover labor market information only for 448 of the 1,027 students in our working sample who eventually graduated from Bocconi. This selection is obviously non-random; however most of it is driven by observable characteristics, namely the survey wave, the respondent's gender and the place of residence.[7] Notice, additionally, that in all our analyzes we always condition on a very large set of observable characteristics, many of which are often unobservable in other datasets (e.g. ability as measured by the entry test score).

The available labor market data include questions on monthly wages in the first job, the type of occupation and contract, and a number of questions on satisfaction with the job and the university. The last rows of Table 2 report some descriptive statistics for the variables that will be explored later in section 6.1. Wages are observed for all graduates who have had a job between the day of their graduation and the day of the interview (a large majority of 97%) and are recorded in 7 intervals of 250 or 500 euros. The lowest interval is for wages below 750 per month and the highest above 5,000. The reported mean and standard deviation refer to an imputed measure of wages computed at the mid-point of the interval indicated by the respondent. All monetary values are in euros evaluated at 2005 prices.

---

[6]The military service was 10 months long and university students could postpone it until graduation. Over the years, the reasons for complete exemption have been expanded (for example, around 2000 a set of new rules allowed permanent exemption from the service to students who enrolled in a PhD programme). The compulsory military service was abolished in 2001 for all citizens born after 1985.

[7]These variables alone explain about 30% of the probability of survey participation. Moreover, the indicator for survey participation is never significant in any regression with measures of performance or ability as dependent variables. This result is unaffected by the introduction of additional covariates.

Our labor market data also allow us to construct a more direct measure of job mismatch using a question regarding work conditions and difficulties encountered by the student in her first job. The specific question reads as follows: *"In your first job, have you experienced any of the following problems/difficulties?"* The list of possible answers is: tasks were too easy, tasks were too demanding, problems with team work, relationship problems with colleagues, difficulties with finding one's position in the organization, job is not secure, low pay, job does not fit personal attitudes. More than one item can be indicated. The measure we adopt is a simple dummy variable that equals 1 if the respondent indicates at least one problem and zero otherwise. However, as we will discuss at length in Section 6.1, our main results are extremely robust to the particular definition of the mismatch variable.

2.1 Lecturing classes

Within each of the nine compulsory courses students are randomly assigned to teaching classes.[8] The number of classes for each course depends on the number of available lecturers. Moreover, the capacity of the available classrooms at Bocconi varies considerably and the number of students in each class had to be determined accordingly. 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 the courses of the following term either at the administration desk or through computers located in the university buildings. Moreover, 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). For these reasons, the total number of students enrolled in each course (the sum over all the classes) may vary slightly across subjects.

When enrolling for a course, the algorithm would randomly assign the student to a class and communicate the allocated class number. By no means could the students interfere with the al-

---

[8]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 remainder of the paper we use the two terms interchangeably.

gorithm. 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.

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.[9] Additionally, students were forced to attend their assigned classes by various methods. First, lecturers were supposed to circulate attendance sheets at the beginning of each class for students to sign their presence. Obviously, with a large number of students in each class (the average class size was 202 students), it was relatively easy for those who wanted to attend a different class to have someone else 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), it would be in her interest to attend her assigned 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.[10]

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.

---

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

[10]As it is now customary in most universities, at the end of each course students are asked to evaluate both the teaching and the logistics of the lectures by filling a questionnaire. The congestion variable reported in Table 4 is computed from such questionnaires which are available in our dataset at the level of each single class.

This one off measure is most likely a lower bound of true attendance given that completing the questionnaire was not compulsory and no strong incentives were provided to increase reporting. Further, typically the form was distributed at the end of a two hours class and would require some effort from the students.

Attendance is also self-reported by the students in the questionnaire, where they 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 the low attendance in these courses. For these reasons we never use information on the law subjects for the definition of the peer groups.

The congestion variable provides an indication of how effectively the allocation into teaching classes was enforced in each course and we will use it later on in Section 5 for some robustness checks.

3   Peer group definition

Our definition of peers is based on students attending courses in 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.

Class attendance heavily influences how, where and with whom students spend most of their time. Each of the 9 compulsory courses is taught in 3 weekly sessions of 2 hours. Hence, students taking only one course together would be sitting in the same classroom for 6 hours per week for a term period of 12 or 13 weeks. Students taking all 9 courses together would spend as much as 700 hours in the same room over the 3 semesters.

While considering classes is standard in the literature on peer effects in high school, in our case effective attendance and the large size of the lecturing classes may 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, more importantly, also by weighting peers by the number of common courses attended together.

In our preferred specification, the weights are non linear as only students who attended at least 4 of the 7 common courses in the same classes are considered peers.[11] This restricted definition is particularly interesting because it leads to group sizes that are comparable to other papers in the literature, particularly those that look at high school classes. For completeness, however, we also present all our results using a simpler, but also looser, definition where any two students who have taken at least one course together are considered peers.

More 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 4 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, $\omega_{ij} \in (0,1]$, 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, $\omega_{ij} = 4/7$ if j attends 4 courses with i, and so on.[12] When computing means of variables at the group level the weights are normalized to sum to one within each group.

[TABLE 5]

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

---

[11] 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).

[12] The weights adopted in the core of the paper are linear in the number of courses attended together both for the restricted and the looser definition of peer group. We have experimented with many other specifications and the results are robust to the weighting scheme; see Section 5.

[TABLE 6]

As a first check of the validity of our definition of peers, Table 6 shows that, after the initial 3 semesters, students who have attended lectures in the same random classes also show remarkably similar academic patterns. In the upper panel of Table 6, for example, we show that students who were randomly assigned to the same classes are significantly more likely to graduate in the same session.[13] 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 most comprehensive definition (column 1) and it increases steadily as the definition becomes more stringent (columns 2 to 4). In our stricter definition the probability of graduating in the same session is almost twice (22.4%) that for non-peers in our sample. The differences are always strongly statistically significant.

In the middle panel of Table 6 we contrast the number of peers and non-peers who choose the same sub-major (i.e. field). Within each of the main majors - economics and business - students could further 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. Among students who have attended at least one of the 7 common courses in the same random class on average slightly more than 9.6% of peers choose the same sub-major. This compares to a marginally lower incidence of students making similar choices among the non-peers. As we restrict our definition of peers to students who have attended more and more courses in the same classes, the difference between peers and non-peers increases and becomes statistically significant. Only with the strictest definition (column 4) does this difference become smaller and insignificant again.

Finally, in the lower panel of Table 6 we look at the probability of choosing the same thesis supervisor (advisor). Once again, students who have been assigned to the same classes in the initial three semesters are substantially more likely to choose the same advisor roughly two years later, and such probability increases with the number of courses taken together.[14]

---

[13]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 committee which decides their final mark (based on both the dissertation and their GPA). Students could 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. See Garibaldi *et al.* (2007).

[14]Note that the pattern in the probability of choosing the same advisor might also be influenced by the fact that

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 that emerge for peers that have attended more and more courses supports the idea of our weighting scheme, which should indeed emphasize the most intense interactions.

Despite all our efforts, our peer groups could still be measured with error, an issue that we discuss more at length in Section 5. Nevertheless, a more general point can already be made here. Without some knowledge of the mechanism that generates social interactions, it is extremely hard to establish *a priori* who is going to influence whom, hence measurement error in the definition of the groups affects virtually all studies of peer effects. In fact, the level and degree of interactions are entirely specific not only to the context but also to the specific social mechanism. For example, if peer effects arise due to imitation, it is unclear that the definition of the groups should be limited to close friends, since one may in fact follow the behavior of the average person, thus including close friends as well as simple acquaintances.

In the literature the definition of peer groups varies substantially, from possibly the most comprehensive, i.e. same race in the State of residence in the US (Charles *et al.,* 2007*),* to the very restrictive roommate in a college dorm (Sacerdote, 2001). Our definition is to the restrictive end of the spectrum, and it appears as the most natural given the institutional setting. In section 5, we perform additional robustness checks by experimenting with alternative definitions and we also discuss the results of a simulation exercise, that is presented in fuller details in Appendix B.

4   Peer effects in major choices

4.1   Identification strategy

The identification of social interaction 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 their peers endogenously and, second, common unobserved shocks may hit the group as a

---

students who have attended the same classes have also met the same professors. However, students typically pick their advisors among the teachers of later elective courses, hence the evidence in Table 6 can hardly be explained solely by the fact that students have met the same professors in the initial compulsory courses.

whole (teacher effects are the usual suspect in studies of education). As a consequence, when we observe a significant correlation between individual and group outcomes it is hard to say whether this result is due to true peer effects or simply to endogenous group formation and/or 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' actions. Although particularly cumbersome, 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, and $E(\mathbf{x}|G_i)$ 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, and $\boldsymbol{\gamma}$ 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 4 courses in the same classes), however, B also studies with C (e.g., they attend some of the remaining 3 courses in the same class, different from A's class). A's peer group 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, so that every student has a distinct group of peers. The weighting scheme described in the previous section adds even more variation to such 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 be further clarified by going back to 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 into the following three equations:

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

Now, consider the corresponding reduced form equations:

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

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.[15]

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 y_C is redundant and, in fact, only observations with distinct groups of peers contribute to identification.[16]

Although this particular setting allows us to solve reflection, one might still worry about the presence of correlated effects, i.e. common unobservable shocks at the group level which could undo 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, class disruptions), and ε_i an i.i.d. random component.[17] If we were to substitute (4) into (1) we would 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 $E(x|G_i)$).[18] However, unobservable group shocks could still be present and induce endogeneity, i.e. $Cov\left(E\left(y|G_i\right), \theta^g\right) \neq 0$.[19] Even if our strategy effectively solves reflection, 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 own peer group. In fact, by construction, the $\mathbf{x}$'s of students who are excluded from i's peer group but included in the group of

---

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

[16] In fact, A and C here have the same peer group, $\{B\}$, although they are not peers to each other. In our data, however, there are no such cases and each single student has a distinct group of peers.

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

[18] 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.

[19] Note that correlated effects cannot induce endogeneity of the exogenous effect - $Cov\left(E\left(x|G_i\right), \theta^g\right) = 0$ - since the x's are determined prior to the allocation to the groups.

one or more of i's peers are uncorrelated with the group fixed effect of i and correlated with the mean outcome of i's group through endogenous interactions. 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.

Bramoullé *et al.* (2006), written parallely (and independently) to our paper, present a more general approach for the identification of social interactions that includes our specific case.

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. Importantly, we maintain the same definition of excluded peers also when working with groups defined over students who take at least 4 courses together. This guarantees that the excluded peers of any student i never attended any course in the same class of i, regardless of how we define actual 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.

[FIGURE 2]

To better document the absence of self-selection in our setup, Figure 2 compares the distributions of some selected individual characteristics in the entire population and in one randomly selected group of peers and excluded peers. In the upper panels of the figure we show the Kernel plots of the distributions of two important measures of ability and academic outcomes, namely the entry test score and the high school grade.

Not surprisingly the distributions of these variables are extremely similar in the population and in the randomly taken groups of peers or excluded peers. We also performed two-sample Kolmogorov–Smirnov tests for the equality of such distributions and in none of the cases we could reject the null of equal distributions.

Two features of the distributions of test scores and high school results are worth noticing. First, the sharp increase of the density of test scores around the (normalized) values of 55 is due to the practice adopted by Bocconi of rejecting students with particularly low test scores, regardless of the availability of admission places. Second, the evident right skewedness of the distribution of high school grades confirms the well known fact that Bocconi attracts a pool of positively selected students.

In the lower panels of Figure 2 we look at two other characteristics: gender and determinedness to major in economics (see Appendix A for a description of how this variable is constructed). Again the proportions of both female and economics determined students are very similar. Two-sample tests of proportions also fails to reject the null of equality.

4.2   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.[20]

To estimate the effect of peers on one's decision to specialize in economics versus business, we run a linear probability model similar to equation (1), where $y_i = 1$ if a student chooses economics and 0 otherwise. $E(y|G_i)$ is the (weighted) share of peers choosing economics 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 major in economics at enrollment, as described in Appendix A).[21]

[TABLE 7]

---

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

[21]We obtain very similar results with a probit model. However, we prefer the linear specification simply because it shows more clearly the features of our identification strategy. Furthermore, the results are robust to the controls added in the table.

Table 7 reports the results of the estimation of linear probability models for our two definitions of peer groups: our preferred one, based on the restricted set of peers who have attended at least 4 courses in the same classes (columns 1 to 4), and one based on students who attended at least one of the 7 common courses in the same class (columns 5 and 6).

For each of these definitions, we estimate the model under two different specifications: simple OLS and weighted IV, using the exogenous characteristics of the excluded peers as instruments and weighting each excluded peer by the number of courses attended with any of the student's peers (see Section 4.1 for a detailed description of the instruments).[22]

Moreover, given the randomness and the relatively large size of the peer groups, we have very little variation in $E(\mathbf{x}|G_i)$, especially when we adopt the largest definition of peers. In that case (columns 5 and 6) we cannot identify $\boldsymbol{\gamma}$ and the constant separately in equation (1) and we omit the average predetermined characteristics of the peer group.[23]

With our preferred restricted definition, instead, despite the random allocation the groups are sufficiently small to guarantee enough sample variation to achieve identification of all the effects. Even in this case the estimated $\boldsymbol{\gamma}$'s are never significantly different from zero but the endogenous effect is hardly affected by the introduction of $E(\mathbf{x}|G_i)$ in the specification.

Thus, in Table 7 we report in columns 1 and 2 estimates based on the restricted definition of peers and including the full set of exogenous effects, in columns 3 and 4 we maintain the restricted definition of the groups but we exclude the exogenous effects from the specification, so that results can be easily compared with those reported in columns 5 and 6, where we define peers using the looser definition (and the full set of exogenous effects cannot be included).

To facilitate the interpretation of our results, which is somewhat complicated by the weighting scheme, Table 7 reports the marginal effects computed for the average student and the average peer. So, for example, the reported estimates for the weighted fraction of peers choosing CLEP measure the effect for an average student of having one additional average peer choosing economics. And, according to the descriptive statistics presented in Table 5, two average restricted peers take 4.16 courses together while two average loose peers take only 1.57 courses together.[24]

---

[22]Weighting excluded peers by the number of links with the main student simply improves the efficiency of the estimates. Results are robust to weighting all excluded peers equally.

[23]The results in columns 5 and 6 are, however, robust to controlling for a subset of $\mathbf{x}$'s at the group level, and the estimated $\boldsymbol{\gamma}$ in that specific case is never significantly different from zero.

[24]These marginal effects are essentially computed by multiplying the actual coefficients by the weight of an average

While the OLS estimates are never significant in any of the specifications, the IV estimates are considerably larger and clearly indicate the presence of significant endogenous peer effects in the choice of major.[25] Our preferred specification is the one in column 2 where we use the restricted definition of peers, IV estimator and condition on the full set of group characteristics. In this specification, the estimates indicate that having one additional average peer opting for economics increases the probability of choosing CLEP by approximately 7.4 percentage points.

This is a very large effect given that on average only about 13% of students major in economics. Notice, however, that the average (weighted) size of a group of restricted peers is about 10 and among these only roughly 1 peer chooses CLEP. Therefore, one additional (average) peer opting for the economics major doubles the number of students choosing economics in one's group. Put it another way, a 10% increase in the share of restricted peers choosing CLEP raises the probability of an average student to choose CLEP by roughly 13%. An elasticity larger than 1 that confirms the importance of the multiplier effect.

Eliminating the exogenous effects from the set of control variables hardly affects the results, with an estimated effect of about 7 percentage points. This compares to a much smaller effect of about 0.8 percentage points that we obtain in column 6 adopting the looser definition of peers. This difference is consistent with the fact that, given the larger groups, $E(y|G_i)$ now varies a lot less. In fact, the standard deviation of the (weighted) share of peers choosing economics is now equal to 0.09, as compared to only 0.01 if the groups include all peers encountered in the 7 common courses that we consider.

Contrary to what is a common result, our IV estimates are larger than the OLS. In fact, it is generally thought that the OLS results over-estimate the actual size of the peer-effect because they cumulate the impact of endogenous effects and correlated effects or group shocks. However, this interpretation rests on the implicit assumption that the two effects influence the dependent variable in the same direction. Group shocks are assumed to affect 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

---

peer. According to the restricted definition, the average student has approximately 10.77 peers with whom she takes 4.16 courses together out of 7. Thus, the weight of the average peer is 4.16/(7x10.77)=0.0552. Using the loose definition, the average student has approximately 151 average peers with whom she sits together in 1.57 courses on average. Hence, the weight of the average peer is 1.57/(7x151)=0.0015.

[25]The IV coefficients are 5 to 6 times larger than the OLS ones.

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, the initial preference for economics. We selected only those students with either very many (top 90th percentile) or very few (bottom 10th 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. Consistent with this interpretation, in this selected sub-sample the IV estimates are smaller than the OLS (we omit the results for brevity). Notice additionally that the limited variation in the endogenous variable $E(y|G_i)$ exacerbates the downward bias of the OLS estimate.[26]

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 group shocks (not taken into account in Sacerdote's setting), the effect becomes sizeable and significant.

Overall, the estimates in Table 7 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 term.[27] The F-test of excluded instruments, reported at the bottom of Table 7, is always very large. Table C.1 in Appendix A shows the complete first-stage regressions for all our IV specifications.

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,

---

[26]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.

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

with t-statistics between 4 and 8. The only exception is the average admission test of the excluded peers that is insignificant when conditioning on exogenous effects.[28]

We also investigate whether the endogenous effect is heterogenous in some observable dimensions, namely in the share of economics-determined students, family income, gender composition and so on. This exercise may suggest some simple policy intervention if, as we will investigate later, the endogenous peer effects can potentially contribute to skill mismatch. There is an important dimension of heterogeneity that we can clearly identify and that is consistent with our prior: the share of determined economics students in one's group.

The endogenous effect is only significant and economically meaningful in peer groups where the relative majority of peers are not determined ex-ante to take economics and the student herself is not determined. For a student who is ex-ante determined to take economics, the endogenous effect is positive and significant only if a quite large share (top quartile in the sample distribution) of her peers are determined to take economics.

In the light of these results, if the university wanted to counteract a possible endogenous effect in the choice of major, it could implement a conditional random allocation making sure to have enough determined to take economics in a given group but not too many. However, we believe the above policy and the "optimal" design of groups to be beyond the scope of this work and we will leave it to further research. The point made here aims at opening up some room for potentially important policy suggestions on group formation arising from the apparent heterogeneity of the estimated effects along observable dimensions.

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 results presented so far.

First of all, we experiment further with how the groups of peers are constructed. The definition that we adopt throughout the paper is based on the assumption that students interact in the classroom and that the more courses they attend together, the more hours they will spend in the

---

[28]The J-tests for overidentification always confirm the validity of all the instruments.

same classrooms, the more intensively they will interact. The weighting scheme that assigns higher weight to students who attend more courses together is meant to capture such an intensity of the interactions.

[TABLE 8]

Our main results are based on either a very simple linear weighting scheme or on a more extreme one that assigns weight zero to any peer that has been encountered in less than 4 courses and increases linearly as the number of common courses increases. In Table 8 we present results obtained employing two alternative weighting functions.

First (in the second row of the table), we exclude the courses in quantitative subjects (math and statistics) from our definition of peer groups, thus considering only classmates that attended courses in economics and business (5 courses in total) in the same classes.[29] Additionally, in the third row of Table 8, 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. Like in Table 7, we always report the marginal effects of having one additional average peer choosing economics.

The estimated endogenous peer effects are highly comparable with our baseline specifications in Table 7, whose estimates are reported in the first row of Table 8. Hence, the weighting function does not seem to be crucial for identification in our setting.

We also performed the following thought experiment: assume that the peer groups we have defined have nothing to do with interactions and the effects that we estimate are generated by mere sample variation (or measurement error or anything else). We then construct placebo peer groups by artificially and randomly assigning students to hypothetical classes. We expect to find no significant effect when the groups are formed using this artificial allocation. In fact, in none of the many specifications reported in the fourth row of Table 8 is there an indication of significant social interactions, and the magnitude of the point estimates is much closer to zero then in Table 7.

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 what is plausibly the most important group shock: teacher quality. From the student evaluation questionnaires we can define

---

[29]We chose the courses most closely related to the choice of major.

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*.[30] The results reported in Table 8 (row 5) are obtained from models similar to those in Table 7, 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 Table 7, 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) suggests that in some cases the actual allocation of students into the classes might not have been maintained. Several anecdotes reveal 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 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).[31]

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 reports 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 variation across classes in this measure of congestion, i.e. there should be some classes with very many students and others with very few ones. 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

---

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

[31] Note that, for anonymization purposes, the class identifier is a randomized version of the true class number.

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 row of Table 8 reports the estimated endogenous peer effects under this particular weighting scheme and shows that the results are very similar to our baseline, suggesting that, in fact, problems of congestion were limited to a few cases.

To conclude this section, we briefly discuss the implications of measurement error in the definition of the peer groups for our estimates. Ammermueller and Pischke (2006) provide a rather comprehensive analysis of identification of peer effects in the presence of measurement error generated by either misresponses or missing values. Given that our data are drawn from the university administrative archives, the first type of error is probably negligible in our setting. Unfortunately, however, the second type of error is substantially more complex in our data than in Ammermueller and Pischke (2006). On the one hand, our groups might not include all relevant peers either because of the few missing observations or because there might be influential friends outside one's randomly assigned classmates. On the other hand, the relatively large size of our classes suggests that some of the randomly assigned peers might hardly ever talk to each other.

Moreover, given the complex network structure employed for identification, it is a priori unclear in which direction should measurement error bias our estimates. Even if random, as in Ammermuller and Pischke (2006), this double-faced type of error most likely biases OLS downwards but has very unclear implications for our proposed IV estimator. This is due to the fact that the measurement errors in the definition of peers and of excluded peers are necessarily correlated, thus invalidating our instruments. Without further assumptions, it is impossible to determine the direction of the bias in the IV estimator.

In order to study the implications of measurement error in our setting, we construct a simple Montecarlo experiment. We generate data from a structural process where the probability of being friends to each other increases with the number of classes taken together but is never equal to one. At the same time, we allow students to be friends with some positive probability even if they have never taken any course together.[32] We, then, look at the performance of our IV estimator vis a vis the simple OLS approach under different assumptions on the probability of being friends without

---

[32]In our simulation, the probability of being friends does not depend on any observable or unobservable characteristics of the individual. In this sense, we analyze the effect of measurement error in a setting where group formation is exogenous.

taking classes together and the probability of not being friends, even when taking classes together. Additionally, we also look at how our estimator performs when the intensity of the correlated effects varies.

The results of this simulation exercise show that in all scenarios the IV performs surprisingly well, and certainly better than the OLS, in recovering the true endogenous effect. Indeed, our simulations show that we might be underestimating the true parameter by a very small factor. All the details of the simulation and the results are reported in Appendix B.

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.[33]

To this end, 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

---

[33]For the analysis in this and the next section we adopt the restricted definition of peers. All the results are, however, robust to changes in the definition.

as the ratio between i's average grade in the exams of the economics area and i's average grade in the exams of the business area. Similarly for students who chose business. We normalize this measure by the relative performance of the full sample of students. Formally, g_i is defined as follows:

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

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, i 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, management 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. $g_i > 1$ and $f_i < 1$. The second group - the *peer driven* - are students who chose the same major 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, 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 9 summarizes these definitions.

[TABLE 9]

As the table shows, students are rather evenly distributed across the four groups. The largest (29.11%) is represented by the *coherent*, i.e. students who choose both according to their ability and their peers. *Peer driven* students, i.e. those who follow peers in contrast with the indication of their academic performance, are only slightly less numerous (27.56%). *Ability driven* students, i.e. those who choose against the relative majority of their peers and follow the signal of their revealed

performance, represent 23.13% of the sample, leaving a sizeable 20.16% 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_i 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 10]

The results of this exercise are presented in Table 10. Columns 2 and 4 extend the specification for average grades in the non-common courses and graduation mark with an additional covariate: 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.[34]

Our main interest is the comparison between the *ability driven* and the *peer driven*. These are in fact the only groups that face a trade-off between their relative aptitude for one subject over the other and the decisions of their mates. The coherent and the incoherent are somewhat uninteresting in this respect, as, whatever choice they make, it is either in accordance or in contrast with both their aptitudes and their peers' choices.

---

[34]We assume a basic version of the 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.

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 -0.17 to -0.20 of a grade point on the average grade in non-common exams (exam grades are given on a scale from 0 to 30 with a passing grade equals to 18) and of -0.64 to -0.74 on the final grade (given on a scale from 0 to 110 with passing a grade equals to 66).

6.1   Labor market effects

In this section we correlate the major decision modes of section 6 with measures of skill mismatch in the labor market. Depending on the mechanism that generates the endogenous social interaction effects estimated in section 4.2, one might expect different implications on the labor market. For example, if peers' interactions are mainly driven by the exchange of valuable information, the decision to follow one's peers should lead to better job matches. On the contrary, if peer-driven students (as defined in the previous section) simply enjoy spending time with their peers, this may in fact come at the cost of a worse job match in the labor market.

Our data allow us to explore these issues thanks to the possibility of linking the administrative information with the surveys of graduates. In particular, we concentrate on two outcomes. The first and most obvious is wages, which are observed for all graduates who have had a job between their graduation day and the day of their interview.

[TABLE 11]

The first two columns of Table 11 report results from simple regressions of (log) wages on the dummies for decision modes, as defined in the previous section. In column 1 we condition on a gender dummy, high school final grade, admission test results, household income at enrolment, degree type, contract type, a dummy for students who attended foreign high schools, a dummy equal to one if the actual degree obtained by the student corresponds to the one indicated as the most preferred upon admission, year dummies and survey wave dummies. In the second column we augment this set of controls with graduation mark and time to graduation (measured in quarters) to verify whether the different decision modes only affect labor market outcomes through their impact

on academic results (as documented in Table 10). Since wages are recorded in intervals, the results in Table 11 are produced with interval regressions.[35]

Results show that peer driven students earn approximately 13% less than their ability driven counterparts. This effect indicates that the implications of choosing a major against one's academic aptitudes go beyond the mere final graduation mark. This, in fact, supports our conjecture on the role of peers in driving towards inefficient outcomes. For comparison purposes, in column 3 we simply regress wages on academic results (graduation mark and duration of studies) to estimate their overall effect on earnings. In this regression we still condition on the large set of controls described above. Consistent with the small effect of decision modes on academic outcomes shown in Table 10, the coefficients on these variables do not change substantially and indicate a relatively small effect of graduation marks on wages: a one point increase in graduation mark leads to a 1.8% increase in the first wage.[36] The robustness of the results to the inclusion of additional regressors also corroborates the simple mean independence assumption that underlies these estimates and that rests essentially on the richness of the control set.

[TABLE 12]

The second measure of mismatch is a dummy variable for job satisfaction constructed from the alumni questionnaire.[37] Table 12 reports results from a set of probit regressions for this indicator of mismatch along the same lines of the wage analysis of Table 11. We find a strong and significant effect of being peer-driven on the likelihood of being mismatched in the first job, and this result is robust to the inclusion of academic outcomes as controls.[38] The marginal effects reported in the table measure the hypothetical experiment of an average ability driven student who changes status

---

[35]The same results have been produced with alternative econometric specifications (i.e. linear OLS using the mid-points of the intervals, quantile regression, ordered probit) and the magnitude and significance of the estimated effects vary only slightly.

[36]Theoretically, graduation marks range from 66 to 110, but the minimum observed value is 82 with a standard deviation of 7.5 over an average of 101.

[37]Namely, from the following question: *"In your first job, have you experienced any of the following problems/difficulties?"* The list of possible answers is: tasks were too easy, tasks were too demanding, problems with team work, relationship problems with colleagues, difficulties with finding one's position in the organization, job is not secure, low pay, job does not fit personal attitudes. More than one item can be indicated. The measure we adopt is a simple dummy variable that equals 1 if the respondent indicates at least one problem and zero otherwise.

[38]The results also change very little with the inclusion of wage categories.

and becomes peer driven. Such change would lead to an increase of 13 percentage points in the probability of reporting any difficulty in the first job (over an average of about 76%).

Interestingly, the graduation mark does not appear to correlate with mismatch, whereas time to graduation does. In our sample those who take longer to complete their degree (conditional on the final mark) end up in substantially better matches, most likely because they spend time searching during their last months in college. This result is confirmed in column 3, where we eliminate the decision mode dummies from the set of regressors. As in Table 11, the similarity of the estimates in the three columns supports the identification assumptions of these estimates and at the same time suggests that decision modes and academic results affect labor market outcomes through distinct channels.

7 Conclusions

In this paper we present a novel strategy for the identification of social interactions and we investigate whether peers' behavior affects the choice of college major, and the academic and labor market performances of young individuals. The available data and the 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 evaluate the labor market effects of choosing a major based more on peers' behavior than on one's revealed ability, and we find a negative impact on both wages and perceived mismatch of skills.

There can be many possible mechanisms generating 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, 2008), then 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 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 labor market penalties associated with the peer driven students is in contrast with this interpretation. It should also be noted that any combination of these explanations may actually be at the origin of the effects that we estimate.

If social interaction effects are important, as we show, they might actually lead to *low level social equilibria* and contribute to the sluggish supply of skills in the labor market (Goldin and Katz, 2007). From a policy perspective, our results highlight the relevance of social interactions that should, therefore, be taken into account in the design of policies. Moreover, our identification strategy emphasizes the importance of including information on one's social network when collecting data.

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] Akerlof, G. and Kranton, R., (2002), "Identity and Schooling: Some Lessons for the Economics of Education," Journal of Economic Literature, 40(4): 1167-1201.

[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] Bender, K. and J. Heywood, (2007),"Educational Mismatch among Ph.D.s: Determinants and Consequences," NBER Working Paper: 12693.

[7] Betts, J. R. (1996). "What do students know about wages? Evidence from a survey of undergraduates", Journal of Human Resources, 31(1): 27-56.

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

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

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

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

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

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

[14] Charles, K., Hurst, E. and Roussanov, N., (2007), "Conspicuous Consumption and Race", mimeo, University of Chicago

[15] Cipollone, P. and Rosolia, A., (2007), "Social Interactions in High School: Lessons from an Earthquake", American Economic Review, 97(3): 948-965.

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

[17] Duflo, E. and Saez, E. (2003), "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," Quarterly Journal of Economics, 118(3): 815-842.

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

[19] Farber H.S. (1999), "Mobility and stability: the dynamics of job change in labor markets", in Ashenfelter O., Card D. (edited by), *Handbook of Labor Economics*, Elsevier Science.

[20] Garibaldi, P., Giavazzi, F., Ichino, A. and Rettore, E., (2007), "College cost and time to complete a degree: Evidence from tuition discontinuities", NBER Working Paper: 12863.

[21] Goldin, C. and Katz, L. (2007), "Long-Run Changes in the U.S. Wage Structure: Narrowing, Widening, Polarizing", Brookings Papers on Economic Activity, forthcoming.

[22] Gottschalk, P. and Hansen, M. (2003), "Is the Proportion of College Workers in Noncollege Jobs Increasing?," Journal of Labor Economics, University of Chicago Press, vol. 21(2): 409-448.

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

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

[25] Guryan, J., Kroft, K. and Notowidigdo, M., (2007), "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," NBER Working Papers 13422, National Bureau of Economic Research.

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

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

[28] 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.

[29] Jackson, M., (2006), "The Economics of Social Networks," Chapter 1 in Volume I of Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society, edited by Richard Blundell, Whitney Newey, and Torsten Persson, Cambridge University Press.

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

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

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

[33] Mas, A. and Moretti, E., (2008), "Peers at Work", American Economic Review, forthcoming.

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

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

[36] Robst, J. (2007), "Education and job match: the relatedness of college major and work", Economics of Education Review, vol. 26 (4): 397-407.

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

[38] 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.

[39] 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):

[40] 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
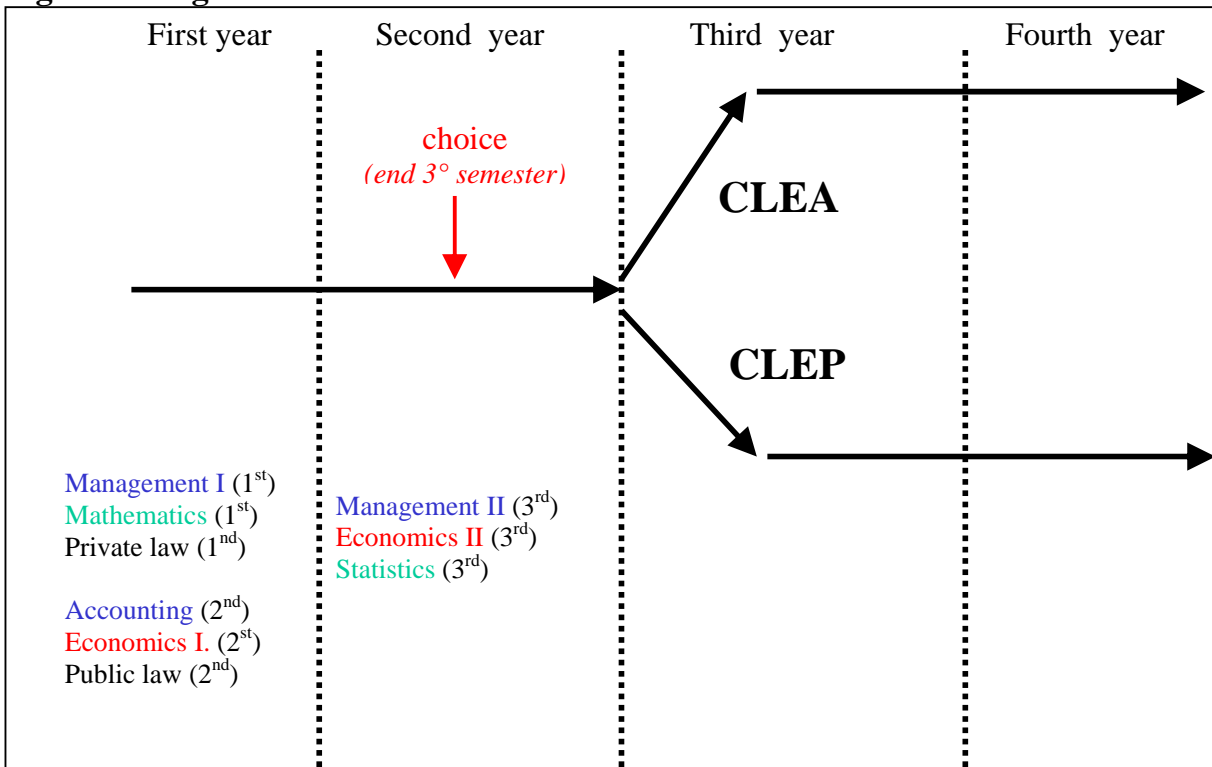
Figure 2: Distributions of selected characteristics
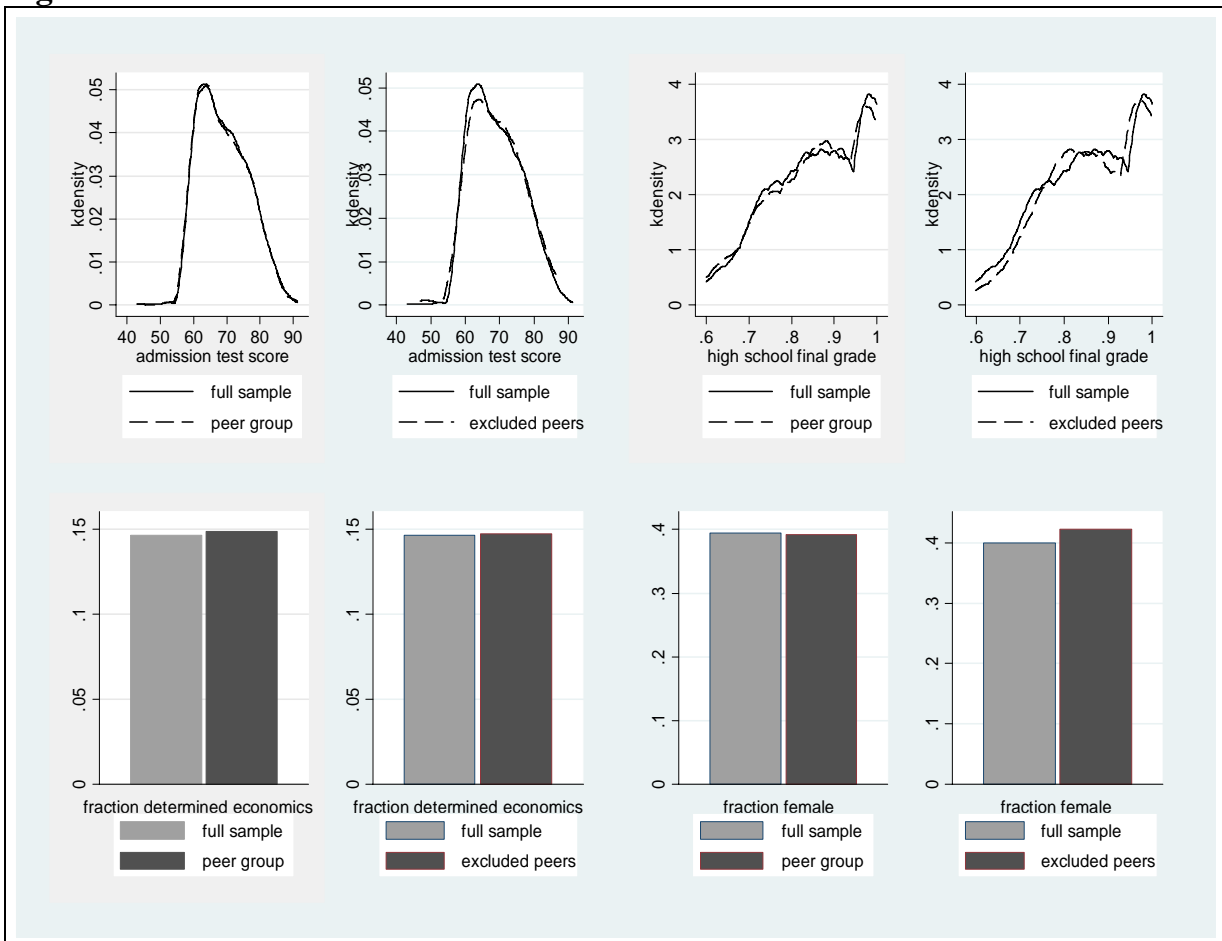
Table 1: Common exams CLEA/CLEP

	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 quantitative 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
Labour market outcomes					
wage in the first job[8]	942.7	(499.2)	750	5,000	432
1=mismatched[9]	0.748	-	0	1	448

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. Normalized between 0 and 100.

7. Normalized between 0 and 100 (pass = 60).

8. Wages are reported in intervals. The reported statistics refer to an estimated wages computed at the mid point of the interval indicated by the responded. All monetary values are evaluated in 2005 Euros.

9. Mismatched=1 if the respondent reports having had difficulties with her first job (tasks were too demanding, not enough demanding, pay was too low, job did not fit one's attitudes, et.).

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

Subject	Semester	Number of classes	Characteristics	Average	(s.d.)	Min	Max
Management I	I	10	Enrolled students	140.40	(14.92)	130	169
			Student questionnaires	80.70	(13.70)	62	109
			Average attendance[1] (%)	85.67	(1.12)	84.08	87.24
			Congestion[2] (1 to 5)	3.33	(0.15)	3.16	3.61
Mathematics	I	10	Enrolled students	140.80	(16.91)	125	164
			Student questionnaires	102.80	(63.86)	28	253
			Average attendance[1] (%)	83.89	(1.53)	81.39	86.51
			Congestion[2] (1 to 5)	3.77	(0.52)	3.00	4.57
Private Law	I	4	Enrolled students	351.75	(164.14)	189	510
			Student questionnaires	70.00	(27.02)	38	104
			Average attendance[1] (%)	79.73	(4.52)	74.91	83.89
			Congestion[2] (1 to 5)	3.07	(0.13)	2.95	3.23
Accounting	II	10	Enrolled students	142.80	(47.75)	109	258
			Student questionnaires	100.30	(61.17)	54	215
			Average attendance[1] (%)	84.80	(1.25)	82.26	86.58
			Congestion[2] (1 to 5)	3.46	(0.48)	3.02	4.40
Economics I	II	6	Enrolled students	216.50	(92.67)	85	316
			Student questionnaires	136.83	(103.78)	24	317
			Average attendance[1] (%)	84.92	(1.23)	83.56	86.84
			Congestion[2] (1 to 5)	3.63	(0.72)	2.83	4.82
Public Law	II	4	Enrolled students	351.75	(147.84)	217	528
			Student questionnaires	41.00	(20.12)	15	64
			Average attendance[1] (%)	82.72	(2.54)	79.45	85.62
			Congestion[2] (1 to 5)	2.89	(0.16)	2.67	3.03
Economics II	III	6	Enrolled students	222.83	(99.20)	156	381
			Student questionnaires	109.17	(52.42)	19	176
			Average attendance[1] (%)	83.87	(1.97)	81.42	86.80
			Congestion[2] (1 to 5)	2.96	(0.47)	2.47	3.72
Management II	III	8	Enrolled students	184.25	(104.07)	123	382
			Student questionnaires	80.75	(25.94)	56	125
			Average attendance[1] (%)	84.38	(0.63)	83.38	85.27
			Congestion[2] (1 to 5)	2.14	(0.25)	1.76	2.51
Statistics	III	8	Enrolled students	272.25	(90.00)	142	404
			Student questionnaires	140.75	(58.91)	35	203
			Average attendance[1] (%)	85.66	(1.04)	83.31	86.53
			Congestion[2] (1 to 5)	3.27	(0.93)	2.09	4.46

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)".

Table 5: Size of peer groups, various definitions.

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

1. Students who have been assigned to the same lecturing class in at least 4 of the 7 common courses considered.
2. Students who have been assigned to the same lecturing class at least once over 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

<u>Definition of peers:</u>				
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 graduate in the same session

	at least 1 [1]	at least 2 [2]	at least 3 [3]	at lest 4 [4]
Peers	13.438	13.890	16.346	22.418
Non-peers		12.523		
Diff.	*0.915****	*1.367****	*3.823****	*9.895****

Panel B: Percentage of students who choose the same sub-major

	at least 1 [1]	at least 2 [2]	at least 3 [3]	at lest 4 [4]
Peers	9.645	9.685	9.908	9.633
Non-peers		9.603		
Diff.	*0.042*	*0.082**	*0.306****	*0.030*

Panel C: Percentage of students with the same thesis supervisor

	at least 1 [1]	at least 2 [2]	at least 3 [3]	at lest 4 [4]
Peers	1.138	1.150	1.194	1.255
Non-peers		0.957		
Diff.	*0.181****	*0.193****	*0.237****	*0.298****

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

Table 7: Peer effects in the choice of major. Linear probability model

Dependent variable: probability of choosing CLEP	Restricted peers (with exogenous effects)		Restricted peers (no exogenous effects)		All peers (no exogenous effects)	
	OLS [1]	2SLS[1] weighted [2]	OLS [3]	2SLS[1] weighted [4]	OLS [5]	2SLS[1] weighted [6]
(weighted) Fraction peers choosing CLEP	0.008 (0.006) [0.184]	0.074* (0.039) [0.054]	0.001 (0.006) [0.155]	0.069* (0.035) [0.053]	0.001 (0.001) [0.193]	0.008* (0.004) [0.068]
Individual characteristics						
Admission test[2]	0.005*** (0.002)	0.005*** (0.002)	0.005*** (0.002)	0.005*** (0.002)	0.005** (0.002)	0.005*** (0.002)
High school final grade[3]	0.426*** (0.103)	0.410*** (0.106)	0.424*** (0.102)	0.416*** (0.106)	0.004*** (0.001)	0.004*** (0.001)
1=determined economics	0.099*** (0.031)	0.110*** (0.033)	0.098*** (0.031)	0.112*** (0.033)	0.095** (0.031)	0.089** (0.031)
1=female	-0.012 (0.020)	-0.017 (0.021)	-0.012 (0.020)	-0.016 (0.021)	-0.011 (0.020)	-0.012 (0.020)
Log household income[4]	0.005 (0.005)	0.006 (0.006)	0.005 (0.006)	0.006 (0.006)	0.005 (0.005)	0.006 (0.005)
1=highest income bracket[4]	0.043 (0.064)	0.056 (0.067)	0.043 (0.065)	0.057 (0.069)	0.044 (0.064)	0.053 (0.064)
1=non resident[5]	-0.002 (0.025)	-0.014 (0.027)	-0.000 (0.025)	-0.012 (0.027)	0.001 (0.025)	-0.002 (0.026)
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	--	0.12	--	0.12	--
Shea Partial R[2]	--	0.288	--	0.030	--	0.071
1st stage F-test	--	10.84	--	12.17	--	30.30

The estimates refer to the effect of having one additional peer with average characteristics opting for economics (CLEP).
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 8: Peer effects in the choice of major. Robustness checks

| | Restricted peers (with exogenous effects) | | Restricted peers (without exogenous effects) | | All peers | |
	OLS [1]	2SLS weighted [2]	OLS [3]	2SLS weighted [4]	OLS [5]	2SLS weighted [6]
1. Baseline	0.008 (0.006) [0.184]	0.074* (0.039) [0.054]	0.001 (0.006) [0.155]	0.069* (0.035) [0.053]	0.001 (0.001) [0.193]	0.008* (0.004) [0.068]
2. Groups based on 5 exams	0.011 (0.008) [0.137]	0.108* (0.060) [0.071]	0.012 (0.007) [0.107]	0.091* (0.046) [0.051]	0.002** (0.001) [0.038]	0.006 (0.004) [0.101]
3. Exponential weights	0.006 (0.005) [0.279]	0.074* (0.041) [0.069]	0.006 (0.005) [0.223]	0.069* (0.039) [0.079]	0.001* (0.000) [0.081]	0.005** (0.002) [0.048]
4. Placebo peer groups	-0.006 (0.005) [0.210]	-0.022 (0.062) [0.720]	-0.004 (0.003) [0.163]	-0.016 (0.044) [0.718]	0.001 (0.001) [0.342]	0.000 (0.007) [0.975]
5. Teacher quality controls	0.004 (0.006) [0.494]	0.068* (0.041) [0.096]	0.005 (0.006) [0.357]	0.061* (0.036) [0.089]	0.001 (0.001) [0.531]	0.008* (0.005) [0.095]
6. Course congestion	0.008 (0.006) [0.171]	0.073* (0.039) [0.058]	0.008 (0.006) [0.145]	0.068* (0.035) [0.053]	0.001 (0.001) [0.531]	0.008* (0.005) [0.095]

The estimates refer to the effect of having one additional peer with average characteristics opting for economics (CLEP).
Robust standard errors in parentheses: p-values in square brackets.

Table 9: Distribution of decision modes (based on restricted peers)

| | | ABILITY INFLUENCE | |
		YES ($g>1$)	NO ($g<1$)
PEERS' INFLUENCE	YES ($f>1$)	Coherent 29.11%	Peer driven 27.56%
	NO ($f<1$)	Ability driven 23.17%	Incoherent 20.16%

Table 10: 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	-0.205**	-0.168*	-0.737**	-0.641**	0.060
	(0.091)	(0.088)	(0.308)	(0.304)	(0.053)
Coherent	-0.058	-0.050	-0.330	-0.310	0.006
	(0.090)	(0.088)	(0.301)	(0.297)	(0.052)
Incoherent	-0.217**	-0.123	-0.776**	-0.534	0.182***
	(0.102)	(0.099)	(0.334)	(0.333)	(0.059)
Ability measures					
Av. grade all exams	--	--	--	--	-0.168***
					(0.014)
Av. grade common exams	0.643***	0.595***	2.875***	2.749***	--
	(0.020)	(0.021)	(0.070)	(0.072)	
Time to graduation	--	-0.479***	--	-1.245***	--
		(0.054)		(0.180)	
1=CLEP	-0.268***	-0.217**	-0.235	-0.103	0.097*
	(0.087)	(0.087)	(0.321)	(0.324)	(0.057)
Admission test[4]	0.006	0.005	-0.021	-0.023	0.001
	(0.005)	(0.005)	(0.016)	(0.016)	(0.003)
High school final grade[5]	2.534***	2.450***	7.497***	7.278***	0.359
	(0.452)	(0.428)	(1.552)	(1.508)	(0.253)
High school type dummies	yes	yes	yes	yes	yes
Individual characteristics					
1=female	0.340***	0.266***	1.042***	0.851***	-0.113***
	(0.065)	(0.063)	(0.220)	(0.218)	(0.039)
Household income[6]	-0.009	-0.015	-0.046	-0.061	-0.013
	(0.031)	(0.027)	(0.108)	(0.097)	(0.016)
1=highest income bracket[6]	-0.018	-0.152	-0.345	-0.693	-0.282
	(0.363)	(0.319)	(1.264)	(1.133)	(0.185)
1=non resident[7]	0.122	0.127	0.384	0.395	0.021
	(0.096)	(0.091)	(0.319)	(0.308)	(0.057)
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

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 11: Wages and decision modes

Dependent variable: log (monthly) wage in the first job[1]

	[1]	[2]	[3]
Decision modes			
Peer driven	-0.130	-0.136*	-
	(0.081)	(0.081)	
	[0.109]	[0.090]	
Coherent	-0.153**	-0.141*	-
	(0.076)	(0.074)	
	[0.045]	[0.058]	
Incoherent	-0.082	-0.096	-
	(0.087)	(0.086)	
	[0.343]	[0.265]	
Academic results			
graduation mark[2]	-	0.018***	0.018***
		(0.006)	(0.006)
		[0.002]	[0.002]
time to graduation[4]	-	0.094	0.089
		(0.057)	(0.057)
		[0.102]	[0.120]
Observations	427	427	427

1. Recorded in intervals. Estimation performed with interval regression.
2. Range 0-111 (pass = 66)
4. Measured in quarters (Official duration is 8 quarters)
All regressions include the following set of controls: a gender dummy, high school final grade, admission test results, family income, degree type, contract type, a dummy for foreign students, a dummy equal to one if the actual degree attended by the student corresponds to the one indicated upon admission, year dummies and survey wave dummies.
Robust standard errors in round parentheses and corresponding p-values in squared brackets.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 12: Mismatch and decision modes

Dependent variable: Probability of mismatch[1]

	[1]	[2]	[3]
Decision modes			
Peer driven	0.443**	0.438**	-
	(0.197)	(0.199)	
	[0.024]	[0.027]	
marginal effect[2]	*0.135*	*0.133*	
Coherent	0.241	0.242	-
	(0.180)	(0.180)	
	[0.179]	[0.178]	
marginal effect[2]	*0.074*	*0.073*	
Incoherent	0.362*	0.376**	-
	(0.190)	(0.191)	
	[0.056]	[0.049]	
marginal effect[2]	*0.106*	*0.108*	
Academic results			
graduation mark[3]	-	-0.002	-0.000
		(0.013)	(0.013)
		[0.896]	[0.979]
marginal effect[2]		-0.001	-0.000
time to graduation[4]	-	-0.410***	-0.407***
		(0.150)	(0.150)
		[0.006]	[0.007]
marginal effect[2]		*-0.135*	*-0.121*
Observations	436	436	436

1. Dependent variable is 1 if the respondent reports having had difficulties with her first job (tasks were too demanding, not enough demanding, pay was too low, job did not fit one's attitudes, et.). See text for details. Probit regression.
2. The marginal effects are compute at the sample averages of all the covariates excluding the decision modes dummies that are all set to zero.
3. Range 0-111 (pass = 66)
4. Measured in quarters (Official duration is 8 quarters)
All regressions include the following set of controls: a gender dummy, high school final grade, admission test results, family income, degree type, contract type, a dummy for foreign students, a dummy equal to one if the actual degree attended by the student corresponds to the one indicated upon admission, year dummies and survey wave dummies.
Robust standard errors in round parentheses and corresponding p-values in squared brackets.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix A: Additional institutional details

The first set of data extracted from the university archives has been used in Garibaldi *et al.* (2007) and kindly passed on to us. Our dataset is an extensively updated version of the same sample of students with information on admission tests, teaching classes, course evaluations, labor market outcomes, exchange programs, etc. The currently available dataset covers all students enrolled at Bocconi since 1989.

Until the academic year 1999/2000, Bocconi offered four other degree programs in addition to the CLEA/CLEP: 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).[39] These degree programs differ both in their curricula and in the number of students admitted in each academic year.[40]

In their application forms, prospective 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 allocated to her first choice if there were still places available in that program; 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, thereby excluding any strategic behavior in the reporting of preferences. This allows us to use this information to construct our indicator of ex-ante preferences. In particular, we consider students who indicated the DES degree - the more academically oriented version of CLEP - as a first or a second choice as "determined" to do economics since the beginning of their studies.[41]

Admitted candidates who decided not to register freed places for students further down in the

---

[39]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.

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

[41]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 who did not get a place in the DES.

ranking. However, only a few students (48 out of 753 in our cohort) 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 registration and accommodation.[42]

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), and CLAPI and DES (with, respectively, 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 the academic year 1999/2000 Bocconi introduced a major reform of its structure (the so-called "Bocconi 2000" plan). In particular, the CLEA/CLEP was abolished and students were forced to choose a specific degree upon entering the university with relatively limited chances to move across programs at later stages. Moreover, the information on the random allocation of students to classes has unfortunately been lost for the earlier cohorts of students and it is reliable only starting with the academic year 1998/1999. This forces us to use only the cohort of students enrolled in the CLEA/CLEP program in the academic year 1998/1999.

---

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

Appendix B: Monte-Carlo Simulations

To investigate the role of measurement error in the definition of the peer groups we design a simple Monte-Carlo experiment where we allow for the groups to be mis-measured in a random fashion. To avoid confusion, in what follows we define the *friends* of a generic individual i as those students who truly affect i's outcome and *peers* those students who take classes together with i. Hence, measurement error arises because not all friends are peers and not all peers are friends.

The specific setting is the following: we simulate a simple world where an outcome y_i is a function of the (weighted) mean outcome and the (weighted) mean exogenous characteristics of one's friends, as well as individual traits (x_i), the sequence of class-specific unobserved confounder (U_i^g) and an iid error (ϵ_i):

$$y_i = \alpha + \beta E(y_{-i}|F_i) + \gamma E(x_{-i}|F_i) + \delta x_i + U_i^g + \epsilon_i.$$

where F_i is the indicator of the group of friends of student i. For simplicity, in the simulation we consider only a one-dimensional x_i. Consistently with our empirical analysis in the main text, $E(y_{-i}|F_i)$ and $E(x_{-i}|F_i)$ are computed weighting each friend by the number of courses taken together. Friends who are not peers are assigned the weight of the average peer.

The correlated effect U_i^g is generated as to reproduce the sum of the macro-shocks (u_i^g) that student i cumulates in the 7 compulsory classes of our application:

$$U_i^g = \sum_{g=1}^{7} u_i^g$$

where the simplest example of an unobservable u_i^g shock would be teacher quality. This unobservable U_i^g represent the correlated effect in the Manski (1993) wording and it is the source of endogeneity of $E(y_{-i}|F_i)$ that our IV strategy is designed to address.

Next, we generate the vector of the x_i's, the class shocks u^g and the idiosyncratic errors ϵ_i for a sample of n students. Specifically, we assume $x \sim N(\bar{x}, \sigma_x)$, $u^g \sim N(0, \sigma_u)$ and $\epsilon \sim N(0, \sigma_\epsilon)$. Given the linear recursive structure of the model, we can solve for the full vector of the outcomes y, once the groups of friends are defined.

Consistently with our application, we fix the number of courses to 7 and the number of classes

for each course to 10, allowing for an uneven distribution of students in each class within each course.[43] Then, we allocate students to classes, and consequentially to peers, according to a totally random procedure and peers are defined, just like in the main text, as those students who take courses in the same classes.

Friends are defined by a probabilistic process that depends on the number of courses two students have taken together. Such process, and consequently the extent of measurement error, is regulated by three crucial parameters:

$$Pr(F_{ij} = 1|meet_{ij} = 0) = \pi_0$$
$$Pr(F_{ij} = 1|meet_{ij} = 1) = \pi_1$$
$$Pr(F_{ij} = 1|meet_{ij} = 7) = \pi_7.$$

where $F_{ij} = 1$ if i and j are friends and $meet_{ij}$ measures the number of courses i and j have taken together. Hence, π_0 is the probability of two students being friends, given that they never met in the classroom, π_1 is the corresponding probability for those students who met once, and finally π_7 gives the likelihood of friendship when the students met 7 times, i.e. the maximum in our simulated (and actual) data.

We assume, as it seems natural, that π increases (non-linearly) in the number of meetings: the more often two students meet the more likely they will be friends and interact. This structure allows a large degree of flexibility in the extent of mis-measurement we can generate in the simulations, while preserving a relatively simple structure. Each scenario will be defined by the triplet (π_0, π_1, π_7) and all the other values of $\pi_m = Pr(F_{ij} = 1|meet_{ij} = m)$ are computed under a simple linearity assumption: $\pi_m = a + bm$ for $m = 1, .., 7$.[44] For example, a scenario of no measurement error is one with $\pi_0 = 0$, $\pi_1 = \pi_7 = 1$, so that all peers would also be friends and vice versa. Increasing π_0 or reducing one or both π_1 and π_7 leads to more mis-measurement.

Throughout the simulation we set the number of observations to 1150 (as in the actual data) and the parameters β, γ, δ and σ_ϵ to the following reasonable values:

---

[43] Namely, we introduce some small random variation in class size within each course.

[44] This linearity assumption could also be modified, although in unreported results we find that it affects results only very marginally.

Parameter	Value
n	1150
β	0.9
γ	0.2
δ	0.5
σ_ϵ	0.2

Given that our main interest in the simulation is on measurement error and its relation to the endogenous macro-shock effect, we let π_0, π_1, and σ_u vary so as to construct a large set of scenarios. For convenience, we fix π_7 to a constant equal to 0.95. We then estimate the model from the simulated data by OLS and IV, where the IV's for student i, as in the paper, are the x's of the excluded peers. Given measurement error, these may not necessarily be also excluded friends. For each of the simulated scenarios we replicate the data 100 times, estimating the model at each replication. In Table B.1 we present results averaged over the 100 replications.

In the upper panel of Table B.1, we consider a scenario where the class shocks u_i^g have a relatively large variance and, thus, the OLS estimates are more heavily affected by endogeneity bias. Overall, the simulation shows that the IV strategy we introduce in the paper performs extremely well in all the scenarios, irrespective of the degree of mis-measurement. While the OLS always overestimates the true β by over 10%, the IV is never biased by more than 4%. Similar results are reported in the lower panel where we consider a scenario with $\sigma_u = 0.1$ (low endogeneity), although the OLS and the IV estimators are now a lot more similar.

A final caveat should be borne in mind when comparing our empirical results with the simulation. In fact, while in our analysis we consider a binary outcome and employ a linear probability model, the simulated results are generated by a simple linear-in-means model. In principle, it is possible to simulate a model that more closely mimics the one we use to produce our main results. However, solving such model recursively would be a lot more complicated (mainly because it may feature multiple equilibria) and the role of measurement error might be confounded by the specific functional form assumptions. Thus, we prefer to present simulation results based on a simpler model in order to focus attention on the role of mis-measurement per se.

Table B.1. Simulation results

Panel 1: High endogeneity (σ_u=0.2)

π_0	π_1	% of peers who are friends	% of friends who are not peers	% of peers who are not friends	Ratio OLS/true parameter	Ratio IV/true parameter
no measurement error		1.00	0.00	0.00	1.132 (0.010)	1.010 (0.010)
0	0.3	0.34	0.00	0.66	1.106 (0.014)	0.983 (0.014)
0	0.5	0.53	0.00	0.47	1.131 (0.012)	1.004 (0.013)
0	0.7	0.72	0.00	0.28	1.120 (0.014)	0.996 (0.015)
0.05	0.7	0.73	0.08	0.27	1.131 (0.014)	1.016 (0.014)
0.05	0.7	0.74	0.16	0.26	1.134 (0.012)	1.014 (0.012)
0.05	0.7	0.73	0.08	0.27	1.131 (0.013)	1.016 (0.013)
0.1	0.3	0.41	0.16	0.59	1.146 (0.011)	1.039 (0.011)
0.1	0.5	0.58	0.16	0.42	1.114 (0.011)	1.003 (0.012)
0.1	0.7	0.74	0.16	0.26	1.134 (0.011)	1.014 (0.013)

Panel 2: Low endogeneity (σ_u=0.1)

π_0	π_1	% of peers who are friends	% of friends who are not peers	% of peers who are not friends	Ratio OLS/true parameter	Ratio IV/true parameter
no measurement error		1.00	0.00	0.00	1.033 (0.006)	1.001 (0.006)
0	0.3	0.34	0.00	0.66	1.013 (0.007)	0.992 (0.008)
0	0.5	0.53	0.00	0.47	1.008 (0.007)	0.982 (0.007)
0	0.7	0.72	0.00	0.28	1.022 (0.007)	0.994 (0.008)
0.05	0.7	0.73	0.08	0.27	1.032 (0.006)	1.004 (0.007)
0.05	0.7	0.74	0.16	0.26	1.049 (0.006)	1.021 (0.007)
0.05	0.7	0.73	0.08	0.27	1.032 (0.006)	1.004 (0.006)
0.1	0.3	0.41	0.16	0.59	1.065 (0.007)	1.043 (0.007)
0.1	0.5	0.58	0.16	0.42	1.055 (0.006)	1.031 (0.006)
0.1	0.7	0.74	0.16	0.26	1.049 (0.006)	1.021 (0.006)

Appendix C: Additional results

Table C.1: IV First-stage regressions

Dependent variable: fraction of peers choosing CLEP	Restricted peers (with exogenous effects) [1]	Restricted peers (without exogenous effects) [2]	All peers [3]
Instruments: excluded peers'			
Admission test[2]	-0.423 (0.472)	0.023*** (0.006)	-0.357*** (0.074)
High school final grade[3]	0.021*** (0.006)	-0.697 (0.487)	0.000*** (0.000)
Fraction of determined economics	-0.418*** (0.104)	-0.463*** (0.106)	-0.140*** (0.018)
Individual characteristics			
Admission test[2]	-0.000 (0.000)	-0.000 (0.000)	-0.000* (0.000)
High school final grade[3]	0.017 (0.027)	0.009 (0.028)	0.000 (0.000)
1=determined economics	-0.009 (0.007)	-0.013* (0.007)	0.000 (0.001)
1=female	0.005 (0.006)	0.003 (0.005)	0.000 (0.000)
Log household income[4]	-0.001 (0.002)	-0.000 (0.002)	-0.000 (0.000)
1=highest income bracket[4]	-0.012 (0.019)	-0.015 (0.022)	-0.003 (0.004)
1=non resident[5]	0.007 (0.007)	0.008 (0.008)	0.000 (0.001)
High school type dummies	Yes	Yes	Yes
Region of residence dummies	Yes	Yes	Yes
Nr. Obs.	1141	1141	1141

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%