Good peers or good teachers? Evidence from a French University[☆]

Thibault Brodaty [a,b,*], Marc Gurgand [c,d]

[a] ERUDITE, Faculté de Sciences-Economiques et de Gestion - ERUDITE, Université Paris-Est Créteil, 61 Avenue du Général de Gaulle, Route de Choisy, Mail des méches, Créteil 94010, France
[b] TEPP: 5 boulevard Descartes, Champs sur Marne 77454 Marne-la-Vallée Cedex 2, France
[c] Paris School of Economics (CNRS), 48 Boulevard Jourdan, Paris 75014, France
[d] CREST: 15 Boulevard Gabriel Péri, 92245 Malakoff Cedex 1, France

ARTICLE INFO

ABSTRACT

Using a quasi-random allocation of students to classes in a French university, we are able to estimate peer effects and teacher effects, with a specific attention to non-linear peer effects. We find that teacher effects are strong, as found at other levels of the education system, but that peer effects have very limited impact. This implies that restricting student access to some universities is of no benefit to remaining students in terms of academic performance. In contrast, attention to teacher performance should be strong at the higher education level.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Students and teachers are the main inputs into the educational production function and both have received a great deal of attention among economists in the recent period. The allocation of students has been hotly debated. Mixing classes may have distributional impact, which depends on the very shape of peer effects, because sorting students would create winners and losers. Mixing classes can also be desirable for efficiency reasons: it is generally efficient to generate heterogeneous groups if peer effects are stronger for low ability students. Therefore, the existence, importance, and details of the structure of peer effects are decisive to several major policy issues. The empirical literature on peer effects is abundant but usually finds limited cognitive impacts of various allocation schemes (Brodaty, 2010; Sacerdote, 2010, 2014). On the other hand, teachers receive increasing attention from both researchers and policy makers, who consider incentives and training policies. Teachers are typically found to account for a large share of the variance in students' cognitive outcomes (Rivkin, Hanushek, & Kain, 2005).

In this paper, we simultaneously estimate the contribution of teacher and student class structure to cognitive outcomes of undergraduate university students at an

elite French university. Using a quasi-experimental setup, in an environment where students work exclusively in small classes, we estimate the impact on initially high- and low-performing students of having initially high- or low-performing peers in the class. We compare the impact of being allocated good students with that of being allocated good teachers. We find that teacher quality is far more important than peer quality as a determinant of cognitive improvement.

We use data from almost 3000 inflow undergraduate economics students over the academic years 2002/2003–2006/2007[1] with around 15 different teachers per subject. We allow the composition of academic quality in the classroom to affect students of different ability differently. As is well understood since Manski (1993) seminal paper, identification of peer effects is difficult. We do not attempt to estimate *endogenous effects*, that is to say, coordination in current behavior. We only consider a reduced form of the peer effects model and estimate the impact of predetermined measures of academic ability. Even then, the challenge lies in separating the effect of peer characteristics from unobserved individual qualities if peers are matched on the basis of their potential performance. This happens for instance when classes are formed by skill level. We will argue that in this university, group formation is *as good as random*, and we test for this. Accordingly, teacher allocation to groups is also as good as random. This allows straightforward identification of the set of peer and teacher effects.

The higher education context raises specific questions. It is usually more selective than compulsory education, so we should wonder whether it is efficient for universities to be strongly stratified? Also, access to higher education is growing in most countries: should this have any visible impact on performance because of the change in peer environment? Our results suggest that these are not first-order concerns.

Although there is recent empirical literature on peer effects in the classroom, some of which considers their very shape, there is only limited evidence on higher education. Arcidiaccono, Foster, Goodpaster, and Kinsler (2012), Braga, Paccagnella, and Pellizzari (2014), De Giorgi, Pellizzari, and Redaelli (2010) and De Paola and Scoppa (2007) find some peer effects in such a context. Some of these papers only document average effects. Non-linear effects are considered in detail in the context of higher education by Booij, Leuven, and Oosterbeek (2014), who do find non-linear peer effects. On the other hand, much of the literature on higher education considers social interactions between roommates, not classmates, which may be less relevant to the organization of education (Carrell, Fullerton, and West, 2009; Foster, 2006; Kremer & Levy, 2008; Lyle, 2007; Sacerdote, 2001; Stinebrickner & Stinebrickner, 2006; Winston & Zimmerman, 2003; Zimmerman, 2003). Generally, these significant effects are modest with respect to those found on other non-academic outcomes in higher education (Sacerdote, 2014).

In contrast, quantitative research on teacher effects at university is more limited, to the best of our knowledge. Carrell and West (2010), Braga et al. (2014) and Hoffmann and Oeropoulos (2009) all find teacher effects on student outcomes, although of a lower order than ours.

Although our peer effects can be precisely estimated with this sample, we find a very small and insignificant impact of class composition on individual student performance. Carrell, Sacerdote, and West (2013) have recently argued that endogenous social interactions within groups can hamper the potential benefits of mixing students by ability, something that may be happening here. In contrast, teacher effects are strong: a one standard deviation increase in teacher quality results in a more than 20% standard deviation increase in students' scores. Naturally, the external validity of the peer effects estimated in this context is questionable. This university is a strongly selective one by French standards, and one in which all teaching is given in small classes, an exceptional situation in undergraduate studies. As a result, neither the technology nor the population is typical. However, this is an exceptional laboratory for learning more about the very structure of peer and teacher effects and assessing them simultaneously.

The paper is organized as follows. Section 2 presents the institutional context and data, Section 3 introduces the model and discusses the identification strategy, Section 4 presents the empirical results and Section 5 concludes.

2. Institutional context and data

We consider an undergraduate economics program of an elite French public university, with a typical yearly inflow of 700–800 students. Students are assigned to small classes of about 25–30, and all the teaching is given at the class level. There are no lectures given to the whole cohort and the classes are fixed for the whole academic year. This is a very favorable situation for observing peer effects and teacher effects in higher education.

We consider the first year of the program (i.e., the first year of undergraduate studies) and we observe year-end exam grades and teacher tutorial marks in the following five subjects: Math (first semester), Microeconomics (first and second semester), Statistics (second semester) and Computer Science (second semester).[2] Many teachers teach the same subject in several classes, and occasionally also teach several different subjects. Table 1 shows that in each subject the team is composed of around 15 teachers, between 25% and 50% of them teaching two classes, depending on the subject and the year.

The year-end grades in each of the five subjects are based on a general exam that is common to all students. It is thus comparable across classes. However, it is different every year, so that between-year comparison may not

---

[2] These five subjects are quite homogeneous in the sense that they rely strongly on formalized mathematical skills: this makes reasonable the required assumption that given measures of individual and peer quality have similar impacts on marks in these subjects.

Table 1
Number of teachers.

		2002	2003	2004	2005	2006
Math	Number of teachers	15	15	16	16	14
	1 group teachers	9	9	8	11	11
	2 groups teachers	6	6	8	4	3
	3 groups teachers	0	0	0	1	0
Micro1	Number of teachers	16	14	18	17	13
	1 group teachers	12	7	12	12	9
	2 groups teachers	3	7	6	5	4
	3 groups teachers	1	0	0	0	0
Micro2	Number of teachers	15	15	18	17	13
	1 group teachers	9	9	12	12	9
	2 groups teachers	6	6	6	5	4
	3 groups teachers	0	0	0	0	0
Stats	Number of teachers	17	15	18	17	14
	1 group teachers	13	9	13	12	11
	2 groups teachers	4	6	4	5	3
	3 groups teachers	0	0	1	0	0
Comp. Sc.	Number of teachers	15	17	17	19	13
	1 group teachers	10	13	10	16	9
	2 groups teachers	4	4	7	3	4
	3 groups teachers	1	0	0	0	0

Table 2
Students quality, by cohort.

		2002	2003	2004	2005	2006
High school	[10 12]	35.1	46.1	19.1	8.9	5.4
Examinations	[12 14]	47.5	39.4	47.0	49.6	31.0
Grade (out of 20)	[14 20]	17.4	14.5	33.9	41.5	63.6
High school	Economics	34.2	40.4	42.1	34.3	33.3
Major	Science	65.8	59.6	57.9	65.7	66.7
Repeaters		8.7	11.8	17.6	11.7	11.8

Note: figures are percentages of students' high school outcomes for each yearly inflow into the University.

Table 3
Sample selection.

		2002	2003	2004	2005	2006
Initial sample	Number of students	870	738	849	817	635
	Non selected (%)	14.3	50.7	17.3	1.8	2.1
	Missing ability (%)	4.8	8.0	3.4	1.2	2.5
	Missing at exam (%)	8.2	12.7	7.7	6.0	6.9
	In special track (%)	19.3	12.3	11.3	12.1	23.2
Final sample	Number of students	609	525	672	661	442
	Sample size (exam)	3017	2623	3355	3290	2210
	Non selected (%)	11.0	47.6	14.1	0.6	0.0
	Number groups	21	21	24	22	17
	Average group size	29.0	25	28	30.0	26
	Std of group size	2.5	2.7	1.6	2.0	2.2
	Average marks at exam	10.4	8.6	9.2	9.5	9.9
	Std of exam marks	3.9	4.3	4.0	4.2	4.3
	Sample size (tutorial)	3085	2753	3452	3329	2240
	Average marks (tutorial)	10.9	9.8	10.6	11.1	10.8
	Std of tutorial marks	3.5	4.1	3.7	3.4	3.2

period, whereas the share of very good students has increased from 17% to 64%. Year 2003 clearly marks a break in this trend, as the share of non-selected (thus on average weaker) students was larger. The share of students who are currently repeating the first year follows a lagged and somewhat attenuated movement. Weaker changes can be observed for the high school major (Economics or Science), reflecting the university's willingness to balance these populations. As we cannot compare marks across years, these changes cannot be used for identification; but they generate significant heterogeneity, which contributes to statistical power.

We have collected administrative data over all of the 2002–2006 inflow students in the first year. Some of the students applied for admission through the selective process: we then have full information about their high school performance. Other students were admitted outside of the selective process: some did not submit a file (we therefore have no high school information for them); others had submitted a file, were rejected on the basis of it, but then managed to get admitted all the same, something that occurred frequently in 2003. The first panel of Table 3 shows the initial sample. In 2003, half of the students were admitted despite failing the selection procedure, but only 2% were in this situation in 2005 and 2006.

Our baseline outcome variable is the mark obtained in the final exam in each of the second semester subjects, this exam being common to all groups for each subject and year. We also observe marks received during the tutorials, although these are based on heterogeneous assignments. Some of the enrolled students (6–8% and almost 13% in 2003), did not take the final exams and are therefore dropouts. Obviously, the outcome is not defined for them. It is unclear whether they also contribute to peer effects, as we do not know how long they attended class: given that we concentrate on second semester subjects, we choose to exclude them from our baseline estimations. As a robustness check, we also run all main equations when missing marks are replaced with a very low mark (5/20): results are robust to this (Appendix A, Tables 13–18).

Information on initial (high school) academic performance is missing for the students that did not submit a

be a source of identification. Tutorial marks are specific to each class and given by the class teachers themselves.

The initial academic performance of students at entry can be measured with the average grade obtained in the high school final exam (the *baccalaureat*). This exam conditions entry into higher education, and is delivered through a formal national examination. In addition, we can also use information on the high school major, either Science or Economics: as we will see, Science students tend to perform better in the subjects considered here.

We use the years 2002–2006. The policies and legal status of this university evolved over this period. In 2002, it was not supposed to select students, although it did engage in some selection. As a result, in 2003, it was compelled by the educational authority to enroll any student with a high school degree from its district, even though they would not have passed the former selection process. Finally, since 2004, the admission process has become openly selective. The effect of this evolution on the recruitment is quite remarkable. Table 2 shows the high school performance of the inflow over those five years. The first three lines of the table give the share of the inflow whose average grade in the high school graduation exam was 10–12, 12–14, 14–16 or 16–20 out of 20 (below 10, the diploma is not delivered, and it is therefore impossible to enter university). In our sample, the share of weak students has decreased from 35% to 5% over the

file: there are 1.2–8.0% of them in the data, depending on the year; we also exclude them from our baseline specifications. As a robustness check, we run all main equations with the number of students with such missing information in the group, and results are also robust to this (Appendix B, Tables 19–24). Finally, 10–20% of the students follow special tracks (with Law or German as an option): they are not subject to the quasi-random formation of classes, so we entirely exclude the corresponding groups from the analysis.

The second panel of Table 3 describes the sample when these three groups (students who did not sit the final exam, those for whom the initial performance is missing and those on special tracks) have been excluded. We end up with 17–24 classes each year, with typically 25–30 students per class, representing 2909 students overall. This is our baseline working sample. We will argue below that class formation is as good as random (except for the special tracks). Therefore, dropping those students that have missing observations does not generate systematic bias in the measure of relative class quality. In addition to the robustness checks in Appendices A and B, we will test specifically that the proportion of students with missing information is balanced across classes.

3. Econometric framework

3.1. The statistical model

In this section we describe how we model students' grades. Consider an individual i from cohort c, in group g. Let y_{igmkc} denote the final grade of student i, in group g, taught by teacher k in subject m. This grade is assumed to depend on the initial academic quality of i, the initial quality of the peers in her group and her teacher for subject m. It also depends on a cohort and subject-specific effect, because the difficulty of the exams varies with time and subjects.

We first consider a reduced form of the Manski (1993) linear-in-means model; we then estimate a more general model where peer group quality can have different impacts for different types of students. Peer group quality is defined by predetermined variables, in our case high school exam outcomes. Because the information on this initial quality is discrete, and the number of ability categories that can be formed out of this information is further limited by the size of the sample, we form two groups: high- and low-performing students.

We note $H_i = 1$ when student i is high-performing at baseline. Our variable of interest is the proportion of high-performing students (excluding the individual of interest i) in a class, $\bar{H}_g^{-i}$. The general structure of this model is thus:

$$y_{imgkc} = a + \alpha H_i + \beta \bar{H}_g^{-i} + \delta_{mk} + \mu_{mc} + \nu_g + \epsilon_{imgkc} \quad (1)$$

for the linear model, and:

$$y_{imgkc} = a + \alpha H_i + \beta \bar{H}_g^{-i} + \gamma H_i \times \bar{H}_g^{-i}$$
$$+ \delta_{mk} + \mu_{mc} + \nu_g + \epsilon_{imgkc} \quad (2)$$

for the interacted model.

Parameters δ_{mk} capture the effect of teacher k teaching subject m; μ_{mc} is the specific effect of subject m for cohort c; ν_g is a group-wise random effect that is taken into account by clustering standard errors at the group level; and ϵ_{imgkc} is an iid shock.

In the interacted model, β measures the effect of having good peers on the low-performing students, whereas $\beta + \gamma$ measures the effect of having good peers on the high-performing students. In the corresponding tables, we will report these two effects, as well as the difference between the two, which is parameter γ.

Note that the initial quality of the students and of their peers is assumed to have the same impact for all subjects, conditional on teacher and cohort $\times$ subject effects. Sample size would not allow reasonably precise estimation of peer effects without this assumption. As mentioned earlier, we restrict the analysis to a set of subjects (Math, Microeconomics, Statistics and Computer Science) that require similar skills.

Teacher effects do not depend on the cohort c, which means that teacher quality is assumed to be time-constant. Therefore we do not take into account potential increases in quality with experience, and teacher effects must be understood as an average quality over the period. Teacher effect is specific to each subject, but for each subject, we normalize one teacher effect to zero. In other words, we only compare teachers within a subject, and we do not compare a teacher's performance across various subjects. In practice this means that teachers who are present in different subjects are considered as different teachers. Note that this model is similar to a value-added model (where H_i stands for the lagged score) from which to estimate teacher effects. As will be argued later, teacher assignment is as good as random, even unconditional on H_i, so that identification of teacher effects is possible (Rothstein, 2010).

However, when reporting aspects of the distribution of teacher effects (its variance or the lowest compared with the highest), we must take into account the fact that estimation errors inflate the variance of those estimators with respect to the true variance of the teacher effects. Therefore, we apply the following corrections (Jacob & Lefgren, 2008): to estimate variance of the teacher effects, $V(\delta)$, we take the empirical variance of our estimates, $V(\hat{\delta})$, and subtract the mean of the variances of the estimators of each $\hat{\delta}_{mk}$. When we estimate specific teacher effects (i.e. the min and the max), we use the shrinkage formula:

$$E(\delta_{mk}|\hat{\delta}_{mk}) = (1 - \lambda_{mk})\bar{\delta} + \lambda_{mk}\hat{\delta}_{mk} \quad (3)$$

where $\bar{\delta}$ is the average value of teacher effect estimates and λ_{mk} is the ratio of $V(\delta)$ and $V(\hat{\delta})$ defined earlier.

Our measure of initial ability, H, is constrained by the available information: track in high school (Science or Economics) and the mark received at the high school final exam (which we only observe in brackets: 10–12, 12–14, 14–16 or 16–20 out of 20). Because we do not know ex ante how much substitution there is between tracks and marks with respect to the measure of initial ability, we use the first semester subjects, Mathematics and Microeconomics, and regress the exam outcomes of those subjects on the full set of interacted high school exam dummies,

Table 4
Choice of ability measure.

High scool grade	High scool major	Math		Micro 1	
		Param	Std	Param	Std
[10–12]	Economics	–	–	–	–
[12–14]	Economics	1.18***	(.277)	1.26***	(.372)
[14–16]	Economics	1.72***	(.284)	2.23***	(.388)
[10–12]	Science	2.40***	(.265)	1.92***	(.359)
> 16	Economics	3.23***	(.461)	3.90***	(.604)
[12–14]	Science	3.83***	(.289)	3.73***	(.375)
[14–16]	Science	4.79***	(.313)	5.26***	(.391)
> 16	Science	6.95***	(.571)	7.13***	(.580)

Note: this table presents the parameters of the OLS regression of the exam grades in Math and Microeconomics (first semester) on the product of high school major and grade. Year dummies are also included in this regression. Standard errors between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively.

Table 5
Group quality heterogeneity.

Measure		2002	2003	2004	2005	2006
No Eco High ability group	Mean	36.9	29.5	45.6	58.5	61.9
	Std	11.9	7.1	9.4	12.8	8.4
	Min	11.1	13.0	25.9	22.5	48.0
	Max	58.3	42.3	61.5	85.7	86.3
Some Eco High ability group	Mean	37.7	30.1	46.8	62.3	67.6
	Std	11.4	7.1	9.1	12.1	8.6
	Min	11.1	13.0	28.5	25.8	53.5
	Max	58.3	42.3	66.6	85.7	90.9

Reading: in this table we compute the proportions of high ability students in each group. We then report summary statistics of these distributions for both ability measures.

tracks × mark brackets. Table 4 reports the coefficients of these two regressions. The order is as expected, with higher marks dominating lower marks in a given subject, and Science, which is more selective, dominating Economics for any given mark. However, the hierarchy between Economics × [14–16] and Science × [10–12] on the one hand and between Economics × [16–20] and Science × [12–14] on the other hand, is not perfectly robust across Mathematics and Microeconomics, although the differences are minor. As a result, we will systematically use two definitions of the high ability group (and thus also the low ability group). The first definition, called the "No Eco high ability group" includes all high school Economics major students in the low ability group, and includes only Science major students with marks of 12 or above in the high ability group. The second definition, called the "Some Eco high ability group" adds the highest performing Economics major students (marks above 16) into the high ability group (results are robust to the definition used).

Table 5 shows the distribution of the proportion of high ability students thus defined in the sample of classes. On average, it varies between 37% in 2002 and more than 60% in 2006, reflecting the dramatic increase in selectivity of this university over the years. Overall, there is a large variation in that proportion across classes: the standard deviation is around 10 points and the gap between the best and the worst class is typically 40–50 percentage points within years. Furthermore, over the years, the range varies between 11% and 91%.

Table 6
Random assignment tests.

Test	Type of test	Measure	KS	p-value
Students into groups	Carrell-West	No Eco high ability group	0.1020	0.190
	Graphical	No Eco high ability group	0.0536	0.905
	Carrell-West	Some Eco high ability group	0.0866	0.364
	Graphical	Some Eco high ability group	0.0506	0.938
	Graphical	Missing	0.0649	0.730
Teachers into groups	Carrell-West	No Eco high ability group	0.0589	0.202
	Graphical	No Eco high ability group	0.0254	0.984
	Carrell-West	Some Eco high ability group	0.0691	0.086
	Graphical	Some Eco high ability group	0.0205	0.999
	Graphical	Missing	0.0248	0.988

Note: this table presents the results of the tests of the random assignment of students and teachers into groups. Kolmogorov statistics in the KS column, and its *p*-value in the *p*-value column.

Table 7
Linear peer effects estimates, with quality measure "No Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Peer effect	−.0009	.0011	−.0027	−.003
	se	(.0033)	(.0023)	(.0024)	(.002)
	# obs	2893	2893	2966	2966
	R2	.09	.125	.07	.11
	F	45.49***	.	30.12***	.
Stat	Peer effect	.0016	.0017	.0022	.0035
	se	(.0024)	(.0022)	(.003)	(.0025)
	# obs	2898	2898	2974	2974
	R2	.188	.217	.083	.156
	F	68.64***	.	26.96***	.
Comp Sc	Peer effect	−.0023	−.0016	−.0005	−.0007
	se	(.0025)	(.0029)	(.0031)	(.0025)
	# obs	2901	2901	2961	2961
	R2	.1	.143	.08	.154
	F	36.19***	.	36.32***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, *F* is not defined when teachers fixed effects are included in the model.

3.2. Identification

The identification of the model depends on the rule governing the allocation of students into classes. Imagine for example that students self-select into classes and that this selection is based on ability: high ability students might end up with other high ability ones and poor ability students with other poor ability ones. Thus there would be a positive correlation between own ability and peer ability. In our model, one dimension of ability, H, is observed and another, ϵ, is not. In the presence of sorting, there would

Table 8

Non-linear peer effects estimates, with quality measure "No Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Effect on low ability	.0019	.0036	.0003	0
	se	(.0033)	(.0025)	(.0027)	(.0024)
	Effect on high ability	−.0043	−.0018	−.0063**	−.0062**
	se	(.0037)	(.0027)	(.0027)	(.0024)
	Difference	−.0062*	−.0054*	−.0067**	−.0063***
	se	(.0027)	(.0024)	(.0025)	(.0024)
	# obs	2893	2893	2966	2966
	R2	.093	.126	.072	.112
	F	42.37***	.	28.08***	.
Stat	Effect on low ability	.0032	.0033	.0038	.0053*
	se	(.0023)	(.0024)	(.0034)	(.0029)
	Effect on high ability	−.0004	−.0001	.0004	.0016
	se	(.0032)	(.0028)	(.0032)	(.0026)
	Difference	−.0036	−.0034	−.0034	−.0037
	se	(.0027)	(.0027)	(.0028)	(.0026)
	# obs	2898	2898	2974	2974
	R2	.188	.217	.083	.157
	F	60.83***	.	23.5***	.
Comp Sc	Effect on low ability	−.0007	.0003	.0023	.0019
	se	(.0025)	(.0031)	(.003)	(.0024)
	Effect on high ability	−.0042	−.0041	−.004	−.0043
	se	(.0034)	(.0034)	(.0037)	(.003)
	Difference	−.0034	−.0043	−.0063**	−.0062***
	se	(.003)	(.0028)	(.0024)	(.0023)
	# obs	2901	2901	2961	2961
	R2	.101	.144	.082	.156
	F	33.55***	.	33.82***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality depending on student own quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, *F* is not defined when teachers fixed effects are included in the model.

Table 9

Linear peer effects estimates, with quality measure "Some Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Peer effect	−.0011	.0006	−.004	−.0034
	se	(.0034)	(.0023)	(.0025)	(.0021)
	# obs	2893	2893	2966	2966
	R2	.099	.133	.074	.114
	F	55.14***	.	31.29***	.
Stat	Peer effect	.0014	.0017	.0037	.0043
	se	(.0025)	(.0024)	(.003)	(.0026)
	# obs	2898	2898	2974	2974
	R2	.191	.22	.089	.161
	F	70.03***	.	30.35***	.
Comp Sc	Peer effect	−.0026	−.0019	.0008	.0002
	se	(.0025)	(.0031)	(.003)	(.0024)
	# obs	2901	2901	2961	2961
	R2	.103	.146	.085	.158
	F	39.84***	.	37.18***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, *F* is not defined when teachers fixed effects are included in the model.

be a positive correlation between ϵ and the average observed group quality, which would upwardly bias peer effect estimates. In contrast, if assignment to groups with respect to student quality is as good as random, such correlation will not be present and students of similar observed ability H can be compared across groups of different quality, so as to infer peer effects.

In the university of interest, class formation is governed as follows. The undergraduate economic program offers compulsory subjects (which we use as outcomes) plus a variety of optional subjects, such as languages, international finance, etc. Students indicate their ordered preferences when enrolling. The timetable allows the formation of about 25 classes, some of which will be compatible with some of the options. There is not one option per class, but typically 4 or 5. The variety of options in classes is further determined by timetable constraints. There is no reason for some classes to concentrate "elite" options, for instance.

Students' preferred choices are treated sequentially, according to their date of enrolment. In order to maximize (approximately) the chances that each student will obtain an option high in his preferences, the staff allocates students to classes according to the options, but never allocates more than 5 students to a class at a time. In other words, when a class has received 5 students, it will not receive any additional students until all the other classes have 5 students. Then the same rule applies up to 10

Table 10

Non-linear peer effects estimates, with quality measure "Some Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Effect on low ability	.0009	.0024	−.0015	−.0009
	se	(.0034)	(.0025)	(.0027)	(.0024)
	Effect on high ability	−.0034	−.0012	−.0069**	−.0059**
	se	(.0037)	(.0025)	(.0028)	(.0024)
	Difference	−.0043*	−.0036*	−.0054**	−.005**
	se	(.0022)	(.002)	(.0023)	(.0021)
	# obs	2893	2893	2966	2966
	R2	.1	.134	.076	.116
	F	48.2***	.	28.91***	.
Stat	Effect on low ability	.0024	.0027	.0047	.0058*
	se	(.0024)	(.0026)	(.0034)	(.003)
	Effect on high ability	.0002	.0007	.0025	.0029
	se	(.0031)	(.0028)	(.0031)	(.0026)
	Difference	−.0022	−.0021	−.0022	−.0028
	se	(.0024)	(.0024)	(.0024)	(.0023)
	# obs	2898	2898	2974	2974
	R2	.191	.221	.089	.161
	F	61.06***	.	26.23***	.
Comp Sc	Effect on low ability	−.0012	−.0002	.0031	.0025
	se	(.0025)	(.0032)	(.003)	(.0024)
	Effect on high ability	−.0043	−.004	−.0018	−.0027
	se	(.0032)	(.0034)	(.0035)	(.0029)
	Difference	−.0031	−.0038	−.0049**	−.0052**
	se	(.0027)	(.0025)	(.0021)	(.0022)
	# obs	2901	2901	2961	2961
	R2	.103	.147	.087	.16
	F	35.59***	.	33.58***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality depending on student own quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, *F* is not defined when teachers fixed effects are included in the model.

Table 11

Teacher effects estimates, with quality measure "No Eco high ability group".

		Linear model		Non-linear model	
		Exam	Tutorial	Exam	Tutorial
Micro	P-value	1.78e−124	7.098e−65	2.95e−111	4.153e−66
	Max−min	.423	.493	.431	.484
	Std	.141	.157	.142	.155
Stat	P-value	0	1.11e−137	2.32e−216	1.56e−141
	Max−min	.774	1.853	.762	1.847
Computer Science	Std	.208	.357	.206	.357
	P-value	2.18e−186	1.81e−147	4.77e−187	1.44e−156
	Max−min	1.164	1.082	1.173	1.093
	Std	.247	.275	.25	.278

Note: this table tests the equality of all teacher dummies within each course. "*P*-value" lines give the *p*-values of this tests. We then report some descriptive statistics (max–min and standard errors) of the distribution of teacher dummies. Those parameters of the distribution of teacher effects have been corrected for the variance inflation resulting from estimation error.

Table 12

Teacher effects estimates, with quality measure "Some Eco high ability group".

		Linear model		Non-linear model	
		Exam	Tutorial	Exam	Tutorial
Micro	P-value	6.99e−196	3.521e−26	3.32e−143	7.987e−69
	Max−min	.452	.504	.455	.488
	Std	.147	.16	.148	.157
Stat	P-value	0	4.88e−138	3.54e−214	6.10e−134
	Max−min	.76	1.82	.755	1.816
Computer Science	Std	.207	.351	.206	.351
	P-value	2.56e−193	4.09e−139	2.46e−145	6.84e−157
	Max−min	1.171	1.097	1.18	1.109
	Std	.247	.274	.25	.278

Note: this table tests the equality of all teacher dummies within each course. "*P*-value" lines give the *p*-values of this tests. We then report some descriptive statistics (max–min and standard errors) of the distribution of teacher dummies. Those parameters of the distribution of teacher effects have been corrected for the variance inflation resulting from estimation error.

students, etc. This is in contrast to an algorithm allocating the most demanded options to the first students, which would result in the last treated students having options very low in their preferences.

Although this is not a strict algorithm, our interviews with the staff that actually implemented it during the years considered here, and with the Dean, indicate that it was generally followed. No information on student background was used and, in principle, students were not allowed to make specific requests, apart from their option preferences. Overall, this sequential allocation into classes should not concentrate good or weak students in the same

groups. In particular, students who ask for the same, say "difficult", subject will be distributed across several classes and some of them will have their second or third best choices in other classes. The only exception is for German, which had its specialized groups and, in 2006, the Law option. As mentioned earlier, we have excluded the corresponding classes entirely.

However, because this algorithm is partly meant to allocate students according to their preferences for options, high performing students may end up being grouped together, if they tend to prefer certain options more often than the low ability students. As we have no information on the ordered preferences, but only on the outcome of the process, we cannot test this directly. Therefore, we follow Carrell and West (2010) and test empirically that the distribution of the proportion of high ability students in the classes is compatible with the distribution that would occur under perfect random assignment. For each class in each year, we randomly drew 1000 possible classes of the same size from that year's student population without replacement. For each class, we then compute the proportion of simulated classes with values lower than that of the observed class. If there is random assignment, for a given year, any given value of this proportion is equally likely to be observed. Therefore, the empirical distribution of those proportions in the sample of classes should be uniform. We test the uniformity of the distributions of those proportions using a Kolmogorov–Smirnov test. To increase power, we pool all years.[3]

As mentioned previously, random allocation of teachers into classes is also necessary to identify teacher effects. The assignment works as follows. Once class timetables have been determined by the staff, teachers rank their preferred classes based on the timetables. At this point, they have no information about group quality. Consequently, there is no reason for good teachers to have high rather than low quality classes on average. We run a similar test to check empirically that teachers are allocated randomly to classes with respect to class quality. The distribution of the proportions of high ability students in the classes and the number of groups per teacher being fixed to the observed ones, we randomly assign, with 1000 repetitions, the groups to the teachers. We calculate for each teacher the proportion of simulations for which the proportion of high ability students is lower than the observed one. Again, if the assignment of teachers is random, the empirical distribution of those proportions should be uniform.

The results of these "Carrell-West" tests are presented in Table 6, in the upper panel for students and in the lower panel for teachers. The Kolmogorv–Smirnov statistics are given in the "KS" column and the corresponding *p*-value is presented in the last column of the table. For both ability measures, the tests cannot reject at 5% the random assignment of students to classes and to teachers.

Fig. 1. Observed and simulated distributions of high ability student proportions, by group (top graph) and by teacher (bottom graph), "No Eco high ability group" measure of ability.

To give a more visual insight of the assignment of student and teachers, we show the above simulated and observed distributions of high ability student proportions, per group and per teacher. The histograms corresponding to both ability measures are shown in Figs. 1 and 2. Observed and simulated distributions of high ability student proportions are visually very similar, by group and by teacher, for both measures. To confirm this impression statistically, we ran Kolmogorov–Smirnov tests that the observed and simulated distributions are equal. The results of these "graphical tests", presented in Table 6, show that we cannot reject at 5% the equality of observed and simulated distributions.

Finally, we test that students with missing observations (dropouts and missing baseline ability) are also randomly assigned with respect to groups and teachers. If this is the case, the measurement error in the proportion of high ability students per group will not bias our estimates. Because missing values are not frequent, some groups have none. The "Carrell-West" tests are therefore not applicable. Consequently we only use the "graphical" tests. Fig. 3 shows that the observed and simulated distributions of the proportions of missing observations per group and per teacher

---

[3] As the process is different across years and may happen not to be quasi-random in some years, we should test randomness separately for each year, at the cost of lower power. When we do this, we never reject randomness, the lowest *p*-value (across 50 values) being 0.14 (results available upon request).
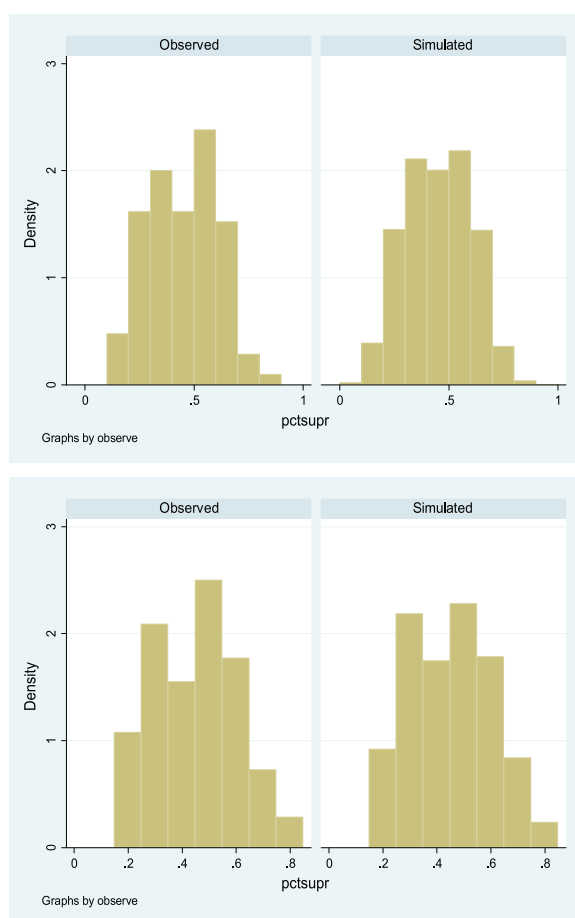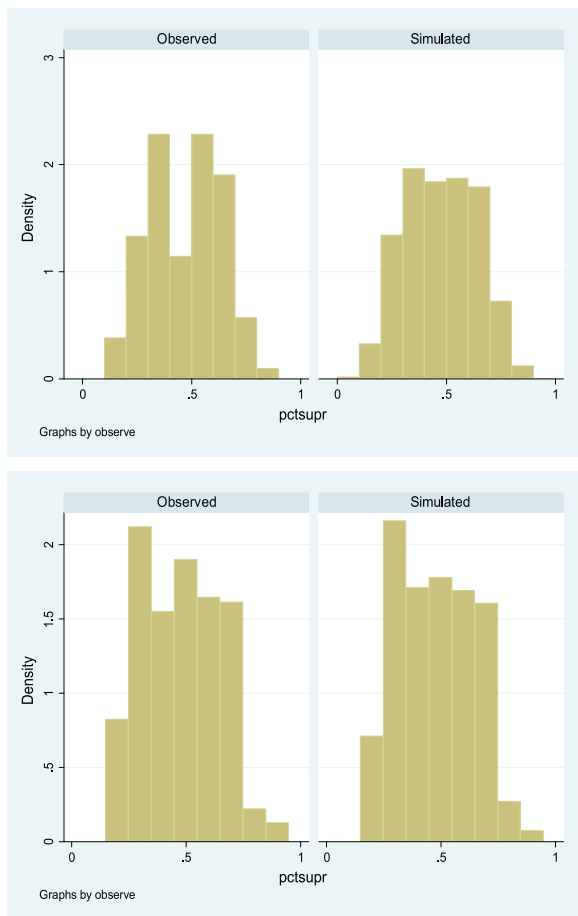
Fig. 2. Observed and simulated distributions of high ability student proportions, by group (top graph) and by teacher (bottom graph), "Some Eco high ability group" measure of ability.

Fig. 3. Observed and simulated distributions of missing value proportions, by group (top graph) and by teacher (bottom graph).

are visually very similar. The Kolmogorov–Smirnov tests presented in Table 6 confirm this impression, with *p*-values greater than 0.73.

In conclusion, we can assume that students, teachers and missing values are randomly assigned into groups. Consequently, the measurement error of peer characteristics, teacher effects and peer characteristics can be considered independent of ϵ. The parameters of our model can thus be identified using least squares estimation. Standard errors will be robust and clustered at the group level.

4. Empirical results

Table 7 presents the results of a model where peer effects are homogeneous in the sense that they are not interacted with student's own initial ability ($\gamma = 0$). In this table, high quality students are measured using the "No Eco high ability group" definition, meaning that students with high school majors in Economics are always counted in the low ability group. The outcomes are exam scores in the second semester subjects, Microeconomics, Statistics and
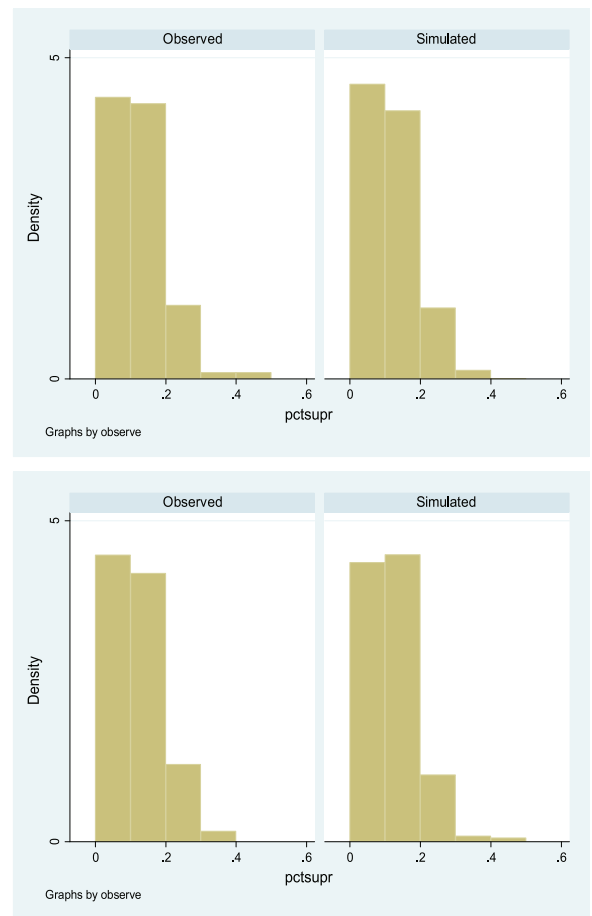
Computer Science.[4] All scores are normalized to variance one, so that coefficients can be interpreted as effect-size. The table reports the coefficient on the proportion $\bar{H}$, thus the effect of increasing the share of high ability students in the class by 1 percentage point. As the standard deviation of $\bar{H}$ is about 10% (Table 5), 10 times the coefficients measures approximately the effect of moving one standard deviation of the distribution of $\bar{H}$.

Table 7 first shows the peer effects separately for each subject. It also provides the estimation with and without teacher dummies. Because teachers' allocation to classes is quasi-random, the results hardly differ. For completeness, we also present results for the exam and for the tutorials, although the latter may not be homogeneous across classes, because they are based on teachers' own assignments and are marked by the teachers themselves. It is striking that the results are not very sensitive to the presence of teacher effects in the latter case either: this means that there is limited heterogeneity in marking behavior. In all cases, the impact of increasing the share of high-

---

[4] We do not use the first semester courses, Mathematics and Microeconomics, because they were used to define high ability vs. low ability students.

performing students is very small and insignificant at standard levels. The effect of an additional percentage point of that share is typically about 0.001 standard deviation in absolute value (and given our precision, we could hope to detect effects of about 0.005 standard deviation). Even if it was statistically significant, such a coefficient would mean that moving from the highest performing class ($\bar{H}$= 0.91) to the lowest performing class ($\bar{H}$ = 0.11) observed between 2002 and 2006 would only increase performance by 8% of a standard deviation.

Table 8 presents the estimation of the full model, where peer effects are interacted with students' own quality, still with the same definition of high-quality students. The first coefficient measures the impact of peer quality on initially low-ability students, and the second coefficient measures the impact of peer quality on initially high-ability students. The difference between the two parameters is also reported, and it corresponds to the interaction parameter γ. If anything, the impact of having good peers is slightly negative for high-performing students, and slightly positive for low-performing students. Although we usually do not have enough power to estimate precisely each of these effects, the difference between the two is sometimes significant, which means that good peers might tend to be more profitable to low- than to high-performing students. However, taken at face value, these effects remain very small.[5] With the alternative measure of baseline ability ("Some Eco high ability group", which includes in the high ability group the best high school students with an Economics major), the results shown in Tables 9 and 10 are qualitatively similar.

Although the peer effects literature on primary and secondary school students is rich, only a few recent papers document peer effects at the higher education level. We find much lower peer effects than these papers, but the literature usually considers residential peers, not classroom peers. Carrell et al. (2013) argue that endogenous social interaction within groups may imply that low-performing students may not be strongly influenced by the presence of better-performing peers. Although their paper considers squadrons and not classes (squadrons are mixed across classes), the same mechanism may apply here. This could be a reason for the low effects, especially when class heterogeneity is strong. Our result is, however, consistent with the idea that peer effects in higher education are more important for non-cognitive outcomes (Sacerdote, 2014).

Table 11 provides a summary of the teacher effects estimated in the model.[6] All effects are identified within subjects, so that we do not try to compare the efficiency of a teacher in Microeconomics and in Statistics, for example: the reason is that we do not have a sufficient number of teachers who teach several subjects at a time to identify such parameters properly. For each subject, the table first presents the *p*-value of a test of equality of all teacher effects. Teacher impact is extremely significant ($p < 0.0001$)

everywhere. The table then displays some elements of the distribution of these effects: the impact of having the best rather than the worst teacher and the standard deviation of the distribution of the effects. As explained earlier, these parameters of the distribution of teacher effects have been corrected for the variance inflation resulting from estimation errors.

The difference between the best and the worst teacher typically represents a gain of 42–117% of an exam score standard deviation. And a one standard deviation increase in teacher quality improves students' scores by between 14% and 25% of a standard deviation of scores, depending on the subject. Teacher effects are higher when looking at tutorial marks, which are left to each teacher, but not strongly so.[7] This is higher than (but of comparable order to) teacher effects found by Rivkin et al. (2005) in primary school (they find about 10% of a standard deviation). This is also significantly higher than recent evaluations in higher education by Carrell and West (2010), Braga et al. (2014), Hoffmann and Oeropoulos (2009), who all find effects representing around 5% of a standard deviation of scores.

5. Conclusion

In this paper, we estimate simultaneously teacher and peer effects in higher education in a French university. The details of the structure of peer effects is important for economic analysis and policy recommendation, and there is only limited evidence on this in the literature. It is therefore useful to provide additional results, in particular at the university level, to which it may not be possible to generalize results found at lower levels.

Identification of the peer effects relies on the fact that groups are set up by the administration following a rule that should not generate any systematic grouping of students by academic quality. This is tested, and confirmed, using the group distribution of observed predetermined characteristics.

On the one hand, we find no cognitive influence of classroom peers, which confirms that, at the higher education level, either peer effects mostly influence non-cognitive outcomes, or residential peer effects are more decisive. On the other hand, teacher effects are quite large, with the standard difference between two teachers representing 20–30% of an exam score standard deviation on average. From a policy point of view, this implies that expanding access to higher education should not be expected to have a negative impact for the category of students already attending. Specifically, the university policy of filtering access more and more strongly has no positive implications for the quality of teaching received by the students. In contrast, the importance of teacher effects raises the classical problem of identifying or improving teacher quality. However, at university, as at other levels of the education system, these findings confirm that significant attention should be focused on teacher performance, at the recruitment, training, evaluation and/or monitoring levels.

---

[5] These very small effects show that the weak instruments bias pointed out by Angrist (2014) does not occur in our data, which is consistent with the large heterogeneity in our group quality.

[6] The results presented in Table 12 for the alternative measure of baseline ability are qualitatively similar.

---

[7] The max/min extreme difference for tutorials in Statistics (1.85) results from one single teacher who has given particularly high marks in his own tutorial group.

Appendix A

As a robustness check, we replace in this appendix missing marks with a very low value (5/20).

Table 13
Linear peer effects estimates, with quality measure "No Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Peer effect	−.002	.0009	−.0036	−.0036*
	se	(.0034)	(.0025)	(.0025)	(.0021)
	# obs	3153	3153	3151	3151
	R2	.093	.122	.075	.112
	F	53.36***	.	33.85***	.
Stat	Peer effect	.0007	.0006	.002	.0015
	se	(.0024)	(.0022)	(.0032)	(.0025)
	# obs	3158	3158	3159	3159
	R2	.167	.194	.087	.149
	F	61.16***	.	31.75***	.
Comp Sc	Peer effect	−.0028	−.0016	−.0021	−.0032
	se	(.0026)	(.0029)	(.0031)	(.0025)
	# obs	3161	3161	3146	3146
	R2	.104	.142	.085	.147
	F	44.89***	.	43.61***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, *F* is not defined when teachers fixed effects are included in the model.

Table 14
Non-linear peer effects estimates, with quality measure "No Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Effect on low ability	.0005	.0033	−.0006	−.0006
	se	(.0033)	(.0027)	(.0027)	(.0024)
	Effect on high ability	−.0049	−.0016	−.0071**	−.0066**
	se	(.004)	(.0028)	(.0028)	(.0025)
	Difference	−.0054**	−.0049**	−.0064**	−.006**
	se	(.0026)	(.0024)	(.0025)	(.0025)
	# obs	3153	3153	3151	3151
	R2	.095	.124	.078	.114
	F	48.95***	.	30.71***	.
Stat	Effect on low ability	.0033	.0029	.0036	.0032
	se	(.0024)	(.0025)	(.0036)	(.0029)
	Effect on high ability	−.0021	−.0019	.0002	−.0003
	se	(.0032)	(.0027)	(.0032)	(.0026)
	Difference	−.0054**	−.0048*	−.0034	−.0035
	se	(.0027)	(.0028)	(.0027)	(.0025)
	# obs	3158	3158	3159	3159
	R2	.168	.195	.088	.15
	F	54.42***	.	27.38***	.
Comp Sc	Effect on low ability	−.0013	.0002	.0006	−.0007
	se	(.0027)	(.0031)	(.003)	(.0025)
	Effect on high ability	−.0045	−.0036	−.0052	−.006**
	se	(.0032)	(.0032)	(.0037)	(.0029)
	Difference	−.0032	−.0038	−.0059**	−.0053**
	se	(.0028)	(.0027)	(.0024)	(.0023)
	# obs	3161	3161	3146	3146
	R2	.105	.143	.087	.149
	F	41.54***	.	39.58***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality depending on student own quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, *F* is not defined when teachers fixed effects are included in the model.

Table 15
Linear peer effects estimates, with quality measure "Some Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Peer effect	−.0017	.0011	−.0052**	−.004*
	se	(.0035)	(.0024)	(.0026)	(.0022)
	# obs	3153	3153	3151	3151
	R2	.101	.13	.079	.116
	F	62.47***	.	33.91***	.
Stat	Peer effect	.0005	.0007	.0033	.0021
	se	(.0026)	(.0025)	(.0032)	(.0025)
	# obs	3158	3158	3159	3159
	R2	.17	.197	.092	.153
	F	60.91***	.	33.82***	.
Comp Sc	Peer effect	−.0034	−.002	−.0008	−.0022
	se	(.0027)	(.0032)	(.0031)	(.0025)
	# obs	3161	3161	3146	3146
	R2	.106	.144	.089	.151
	F	47.63***	.	42.35***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality (each co-efficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, *F* is not defined when teachers fixed effects are included in the model.

Table 16
Non-linear peer effects estimates, with quality measure "Some Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Effect on low ability	0	.0027	−.0026	−.0014
	se	(.0034)	(.0026)	(.0027)	(.0024)
	Effect on high ability	−.0034	−.0003	−.008***	−.0064**
	se	(.0039)	(.0026)	(.0029)	(.0025)
	Difference	−.0034	−.0031	−.0054**	−.005**
	se	(.0021)	(.002)	(.0022)	(.0022)
	# obs	3153	3153	3151	3151
	R2	.101	.13	.081	.117
	F	54.35	.	30.6	.
Stat	Effect on low ability	.0025	.0025	.0046	.0037
	se	(.0026)	(.0027)	(.0036)	(.0029)
	Effect on high ability	−.0015	−.0011	.002	.0005
	se	(.0031)	(.0028)	(.0032)	(.0025)
	Difference	−.004*	−.0036	−.0027	−.0031
	se	(.0024)	(.0025)	(.0024)	(.0023)
	# obs	3158	3158	3159	3159
	R2	.171	.198	.092	.153
	F	53.38	.	29.11	.
Comp Sc	Effect on low ability	−.002	−.0003	.0015	0
	se	(.0027)	(.0033)	(.003)	(.0025)
	Effect on high ability	−.0048	−.0037	−.0033	−.0046
	se	(.0032)	(.0034)	(.0036)	(.0028)
	Difference	−.0029	−.0034	−.0048**	−.0047**
	se	(.0025)	(.0024)	(.0022)	(.0021)
	# obs	3161	3161	3146	3146
	R2	.107	.145	.091	.152
	F	42.63	.	38.28	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality depending on student own quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, *F* is not defined when teachers fixed effects are included in the model.

Table 17
Teacher effects estimates, with quality measure "No Eco high ability group".

		Linear model		Non-linear model	
		Exam	Tutorial	Exam	Tutorial
Micro	*P*-value	2.221e–85	6.83e–109	1.603e–69	3.33e–106
	Max–min	.392	.436	.403	.432
	Std	.124	.146	.126	.144
Stat	*P*-value	1.10e–185	2.70e–124	2.63e–194	1.82e–104
	Max–min	.881	1.545	.87	1.541
Computer	Std	.214	.324	.211	.323
Science	*P*-value	1.88e–167	4.86e–173	1.75e–109	3.88e–170
	Max–min	1.034	.89	1.035	.893
	Std	.232	.232	.234	.234

Note: this table tests the equality of all teacher dummies within each course. "*P*-value" lines give the *p*-values of this tests. We then report some descriptive statistics (max–min and standard errors) of the distribution of teacher dummies. Those parameters of the distribution of teacher effects have been corrected for the variance inflation resulting from estimation error.

Table 18
Teacher effects estimates, with quality measure "Some Eco high ability group".

		Linear model		Non-linear model	
		Exam	Tutorial	Exam	Tutorial
Micro	*P*-value	6.074e–94	4.06e–122	7.78e–103	9.48e–124
	Max–min	.443	.441	.448	.434
	Std	.133	.148	.134	.146
Stat	*P*-value	2.59e–151	1.92e–135	2.318e–56	2.59e–128
	Max–min	.865	1.519	.865	1.516
Computer	Std	.213	.319	.211	.319
Science	*P*-value	7.21e–166	1.60e–172	2.49e–135	1.38e–169
	Max–min	1.041	.902	1.043	.911
	Std	.232	.232	.235	.234

Note: this table tests the equality of all teacher dummies within each course. "*P*-value" lines give the *p*-values of this tests. We then report some descriptive statistics (max–min and standard errors) of the distribution of teacher dummies. Those parameters of the distribution of teacher effects have been corrected for the variance inflation resulting from estimation error.

Appendix B

As a robustness check, we include in this appendix the number of missing values per group in the model.

Table 19
Linear peer effects estimates, with quality measure "No Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Peer effect	−.0007	.0018	−.0012	−.0019
	se	(.0033)	(.0022)	(.0025)	(.002)
	# obs	2893	2893	2966	2966
	R2	.09	.126	.073	.112
	F	39.75***	.	29.36***	.
Stat	Peer effect	.0022	.0022	.0021	.0026
	se	(.0024)	(.0023)	(.0031)	(.0026)
	# obs	2898	2898	2974	2974
	R2	.188	.217	.083	.157
	F	59.56***	.	23.12***	.
Comp Sc	Peer effect	−.0016	−.0015	.0001	−.0007
	se	(.0026)	(.0029)	(.0032)	(.0025)
	# obs	2901	2901	2961	2961
	R2	.101	.144	.081	.154
	F	30.79***	.	31.05***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, F is not defined when teachers fixed effects are included in the model.

Table 20
Non-linear peer effects estimates, with quality measure "No Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Effect on low ability	.0021	.0042	.0017	.001
	se	(.0034)	(.0025)	(.0027)	(.0023)
	Effect on high ability	−.0041	−.001	−.0048*	−.005**
	se	(.0037)	(.0026)	(.0028)	(.0025)
	Difference	−.0062**	−.0052**	−.0066***	−.006**
	se	(.0026)	(.0024)	(.0025)	(.0024)
	# obs	2893	2893	2966	2966
	R2	.093	.128	.076	.114
	F	37.24***	.	28.4***	.
Stat	Effect on low ability	.0038	.0038	.0036	.0043
	se	(.0023)	(.0025)	(.0035)	(.0029)
	Effect on high ability	.0002	.0004	.0002	.0007
	se	(.0032)	(.0029)	(.0033)	(.0027)
	Difference	−.0035	−.0034	−.0034	−.0036
	se	(.0027)	(.0027)	(.0028)	(.0026)
	# obs	2898	2898	2974	2974
	R2	.189	.217	.083	.158
	F	53.62***	.	20.58***	.
Comp Sc	Effect on low ability	−.0001	.0004	.0029	.002
	se	(.0026)	(.003)	(.0031)	(.0024)
	Effect on high ability	−.0035	−.004	−.0034	−.0042
	se	(.0034)	(.0034)	(.0038)	(.003)
	Difference	−.0034	−.0043	−.0063**	−.0062***
	se	(.003)	(.0028)	(.0024)	(.0023)
	# obs	2901	2901	2961	2961
	R2	.102	.145	.083	.156
	F	29.1***	.	29.51***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality depending on student own quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, F is not defined when teachers fixed effects are included in the model.

Table 21
Linear peer effects estimates, with quality measure "Some Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Peer effect	−.0009	.0013	−.0026	−.0023
	se	(.0034)	(.0022)	(.0026)	(.0021)
	# obs	2893	2893	2966	2966
	R2	.099	.134	.077	.117
	F	47.97***	.	29.62***	.
Stat	Peer effect	.002	.0022	.0036	.0033
	se	(.0025)	(.0025)	(.0031)	(.0026)
	# obs	2898	2898	2974	2974
	R2	.192	.221	.089	.161
	F	60.67***	.	26***	.
Comp Sc	Peer effect	−.002	−.0017	.0015	.0002
	se	(.0025)	(.0031)	(.0032)	(.0024)
	# obs	2901	2901	2961	2961
	R2	.104	.147	.086	.158
	F	33.95***	.	32.13***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, F is not defined when teachers fixed effects are included in the model.

Table 22
Non-linear peer effects estimates, with quality measure "Some Eco high ability group".

		Exam		Tutorial	
		Witout teacher	With teacher	Witout teacher	With teacher
Micro	Effect on low ability	.0011	.0029	−.0001	0
	se	(.0034)	(.0024)	(.0027)	(.0024)
	Effect on high ability	−.0032	−.0005	−.0055*	−.0047*
	se	(.0036)	(.0025)	(.0029)	(.0024)
	Difference	−.0043**	−.0034*	−.0054**	−.0047**
	se	(.0022)	(.002)	(.0023)	(.0022)
	# obs	2893	2893	2966	2966
	R2	.1	.135	.079	.118
	F	42.6***	.	28.34***	.
Stat	Effect on low ability	.003	.0033	.0046	.0048
	se	(.0024)	(.0026)	(.0034)	(.003)
	Effect on high ability	.0008	.0012	.0024	.002
	se	(.0031)	(.0029)	(.0032)	(.0026)
	Difference	−.0022	−.0021	−.0022	−.0028
	se	(.0024)	(.0024)	(.0024)	(.0023)
	# obs	2898	2898	2974	2974
	R2	.192	.221	.089	.162
	F	53.76***	.	22.94***	.
Comp Sc	Effect on low ability	−.0007	0	.0038	.0026
	se	(.0025)	(.0032)	(.0031)	(.0024)
	Effect on high ability	−.0037	−.0038	−.0011	−.0026
	se	(.0032)	(.0034)	(.0036)	(.0028)
	Difference	−.003	−.0038	−.0049**	−.0052**
	se	(.0027)	(.0025)	(.0022)	(.0022)
	# obs	2901	2901	2961	2961
	R2	.104	.148	.088	.16
	F	31.02***	.	29.65***	.

Note: this table presents the OLs regression of second trimester outcomes on peer quality depending on student own quality (each coefficient correspond to a different estimation); controls for own quality, year fixed effects and, when indicated, teacher fixed effects. Robust standard errors clustered at the group level between brackets. *, **, *** are significance levels at 10%, 5% and 1%, respectively. Because some teachers appear only one time in the sample, F is not defined when teachers fixed effects are included in the model.

Table 23
Teacher effects estimates, with quality measure "No Eco high ability group".

		Linear model		Non-linear model	
		Exam	Tutorial	Exam	Tutorial
Micro	*P*-value	5.65e−115	1.045e−63	4.98e−134	1.183e−77
	Max–min	.447	.532	.448	.518
	Std	.149	.156	.149	.152
Stat	*P*-value	3.26e−226	2.533e−95	5.61e−219	5.30e−155
	Max–min	.737	1.837	.723	1.831
Computer	Std	.202	.36	.2	.36
Science	*P*-value	7.33e−184	3.25e−154	5.60e−149	5.21e−157
	Max–min	1.151	1.059	1.156	1.071
	Std	.243	.273	.246	.277

Note: this table tests the equality of all teacher dummies within each course. "*P*-value" lines give the *p*-values of this tests. We then report some descriptive statistics (max–min and standard errors) of the distribution of teacher dummies. Those parameters of the distribution of teacher effects have been corrected for the variance inflation resulting from estimation error.

Table 24
Teacher effects estimates, with quality measure "Some Eco high ability group".

		Linear model		Non-linear model	
		Exam	Tutorial	Exam	Tutorial
Micro	*P*-value	3.41e−119	5.850e−49	1.11e−143	8.42e−107
	Max–min	.486	.547	.485	.527
	Std	.154	.159	.154	.156
Stat	*P*-value	5.87e−211	1.59e−151	3.66e−217	3.74e−155
	Max–min	.721	1.808	.716	1.805
Computer	Std	.201	.355	.2	.355
Science	*P*-value	6.22e−177	4.24e−148	2.65e−222	3.65e−161
	Max–min	1.158	1.075	1.164	1.087
	Std	.244	.273	.247	.277

Note: this table tests the equality of all teacher dummies within each course. "*P*-value" lines give the *p*-values of this tests. We then report some descriptive statistics (max–min and standard errors) of the distribution of teacher dummies. Those parameters of the distribution of teacher effects have been corrected for the variance inflation resulting from estimation error.

References

Angrist, J. (2014). The perils of peer effects. *Labour Economics, 30*, 98–108.

Arcidiaccono, P., Foster, G., Goodpaster, N., & Kinsler, J. (2012). Estimating spillovers using panel data, with an application to the classroom. *Quantitative Economics, 3*(3), 421–470.

Booij, A., Leuven, E., & Oosterbeek, H. (2014). *Ability peer effects in university: Evidence from a randomized experiment*. Mimeo.

Braga, M., Paccagnella, M., & Pellizzari, M. (2014). The academic and labor market returns of university professors. IZA DP no. 7902.

Brodaty, T. (2010). Les effets de pairs dans l'education: une revue de litterature. *Revue d'Economie Politique, 120*(5), 739–757.

Carrell, S., Fullerton, R., & West, J. (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics, 27*(3), 439–464.

Carrell, S., Sacerdote, B., & West, J. (2013). From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica, 81*(3), 855–882.

Carrell, S., & West, J. (2010). Does professor quality matter? evidence from random assignment of students to professors. *Journal of Political Economy, 118*(3), 409–432.

De Giorgi, G., Pellizzari, M., & Redaelli, S. (2010). Identification of social interactions through partially overlapping groups. *American Economic Journal: Applied Economics, 2*(3–4), 197–221.

De Paola, M., & Scoppa, V. (2007). Peer effects and the academic performance of Italian students. *Applied Economics, 42*, 2203–2215.

Foster, G. (2006). It's not your peers and it's not your friends: some progress toward understanding the educational peer effect mechanism. *Journal of Public Economics, 90*, 1455–1475.

Hoffmann, F., & Oeropoulos, P. (2009). Professor qualities and student achievement. *Review of Economics and Statistics, 91*(1), 83–92.

Jacob, B., & Lefgren, L. (2008). Can principals identify effective teachers? evidence on subjective performance evaluation in education. *Journal of Labor Economics, 26*(1), 106–136.

Kremer, M., & Levy, D. (2008). Peer effects and alcohol use among college students. *Journal of Economic Perspective, 22*(3), 189–206.

Lyle, D. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at west point. *Review of Economics and Statistics, 89*(2), 289–299.

Manski, C. (1993). Identification and endogenous social effects: The reflection problem. *Review of Economic Studies, 60*, 531–542.

Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *Quarterly Journal of Economics, 125*(1), 175–214.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics, 116*, 681–704.

Sacerdote, B. (2010). Peer effects in education: How might they work, how big are they and how much do we know thus far?. In *Handbook of economics of education*. Amsterdam: North Holland.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: Two steps forward? *Annual Review of Economics, 6*, 253–272.

Stinebrickner, R., & Stinebrickner, T. (2006). What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics, 90*, 1435–1454.

Winston, C., & Zimmerman, D. (2003). Peer effects in higher education. NBER Working Paper Series 9501.

Zimmerman, D. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economic and Statistics, 85*, 9–23.