Discussion
Papers
in Economic
Behaviour

DPEB
05/18

Gender responses to competitive pressure in college: a regression discontinuity design

>Pilar Beneito
University of Valencia and ERICES

>Inés Rosell
University of Valencia

October, 2018

www.erices.es

Gender Responses to Competitive Pressure in College: a Regression Discontinuity Design

Pilar Beneito[a] and Inés Rosell

(University of Valencia and [a]ERICES)

Abstract

The proliferation of competitive college groups in Spain capturing highly qualified students has opened an interesting debate, motivating the study of how students react in such competitive environments. In this paper we provide empirical answers to this issue by comparing high achievement groups (in particular, *International Business* and *Law and Business*) with standard groups (*Business Administration*) at the University of Valencia, Spain. The co-existence of the two kind of groups sharing similar academic programs and the fact that they are separated by a particular value of the access-to-university score each year provide a suitable data source that allows us to identify the causal effect of peers by using a (fuzzy) regression discontinuity design. We implement this methodology to analyze peers' influences in terms of *learning externalities, competitive pressure*, or *requirement standards*, making special emphasis in gender disparities. Our results suggest that peer effects in college are negative and significant for students at the threshold, that is, for those who are ranked at the bottom of the high achievement groups. These findings are more remarkable for women and in *International Business*, where the level of competitive pressure is expected to be the highest among the three groups considered. We conclude that *competitive pressure* exerts a negative impact on threshold student`s grades, particularly women, a result that contributes to the recent literature documenting the lower preference of women for competitive contexts.

Keywords: high achievement groups, college education, competitive pressure, gender, fuzzy regression discontinuity.

JEL codes: C01, D9, I23, J16.

[*] We thankfully acknowledge the administration of the *University of Valencia* for providing the data. Pilar Beneito acknowledges financial support from the Spanish Ministerio de Economía y Competitividad (ECO2014-55745-R)

1. Introduction

In the wake of globalization, we live in a rapidly changing world where economic challenges emerge every day. Thus, it is extremely important to have motivated and educated human capital able to promote growth and to generate value added to our economies. We might as well aspire to foster an economic system that promotes an equally distribution of wealth, and we acknowledge that the only way to achieve this purpose is through education. In developed countries we have done well in decreasing inequality at many levels, but there is still work to be done to reduce gender inequality. In spite of the progress, the gender wage gap and the famous glass ceilings, the term used to reflect that women find it more difficult to get leading positions in companies, are still a reality. Hence, one of the main challenges of current economists and policy makers is to think about educational policies that help to foster the development of skills and knowledge of human capital as it will end up generating higher levels of productivity, growth, and development. Additionally, educational policies should as well promote an equalize world where someone`s career is not delimited by gender.

Since we believe that education is the driving force of economic development and definitely decreases inequality, we consider relevant to study the results that the current educational system is generating to better design future educational policies. For that reason in this paper we study the Spanish educational system. In particular, we analyze the atmosphere generated by peers in High Achievement college groups (HA hereafter) that exist in Spanish universities. To this end, and using administrative data from the University of Valencia (UV hereafter), we center our attention on the comparison of students in standards groups (SG hereafter) of the degree of Business, where entry is not conditional on particularly high entering marks, with students in the degrees of International Business and Law and Business where, on the contrary, admittance is conditional on passing a given threshold in the access-to-university exams (AU-score hereafter). Educational centers favor the creation of HA groups as they are considered to guarantee certain advantages not only at college level but also for future professional development.

To be more precise, our research questions here are: do these HA groups actually favor students in terms of academic achievement as compared to standard groups? And, if so, does this occur equally for male and female students? In this paper we propose that, although peers with higher academic ability may be, as widely argued, a source of *learning externalities* in a classroom, we may find additional effects from peers. Particularly at higher educational levels, HA groups may also imply higher *requirement standards* demanded by teachers, if only because teachers end up adjusting the distribution of marks in a relative manner. In addition, once students are in a HA group, they are not just exposed to higher educational requirements and more prepared mates but also to a certain atmosphere created in class, usually taking the form of higher *competitive pressure*. Hence, we believe that there are at least three different channels that may impact student`s performance in HA college classes which are *learning externalities*, *requirement standards* and *competitive pressure*.

While higher academic standards would be expected to exert a ‘cost’ on all students in HA groups equally, we consider, on the contrary, that competitive pressure may affect students in more subjective manners. Motivated by the subject literature that has recently emphasized that women tend to suffer relative more in competitive environments, we focus on female students to understand if at this stage of their education they share certain traits that may delimit their performance in their proper professional careers. In other words, we aim to understand whether being a part of a HA group has a positive or a negative impact on college women`s grades as compared with men.

The co-existence of standard groups and HA groups sharing similar academic programs, and the fact that they are separated by a particular value of the AU-score each year provide a suitable data source that allows us to identify the causal effect of peers by using a regression discontinuity design (RD henceforth), as we explain below. As the cutoff mark is really specific, students who were eligible to enroll in HA groups for a tiny score difference and students who nearly overpass the cutoff mark but ended up in standard groups should have similar academic abilities and thus provide a comparable sample of individuals. The RD methodology is based on the comparison of these two groups of individuals, mainly differentiated by the kind of group where they have been finally allocated. We use administrative data that refer to individual grades obtained in several subjects by 4 waves of entering students in the abovementioned groups at the UV from 2011 to 2014.

The way peers impact student`s achievement is a latent topic in education economics because being with certain colleagues may have an important impact on one´s future achievements. People´s confidence in HA groups has increased over the last years as long as top students have found an opportunity to outperform, take advantage of learning externalities created from other students, and to acquire better jobs. But whether students who were able to enroll in a HA group but whose AU-score was near the cutoff mark have benefited from positive spillovers such as their colleagues is still under discussion. There is an open debate about this issue because some people argue that instead of being at the bottom of a highly competitive class they would rather be at the top of the less competitive class and outperform, commonly explained by the expression: "I'd rather be a big fish in a small pond than a little fish in a big pond". The thing is that some students in HA classes may end up underperforming due to the higher academic requirements or the competitive pressure.

More interestingly, may peer effects be negative and the channel of competitive pressure affects students more than learning spillovers women will probably be more harmed than men. Literature shows that women perform worse in competitive environments by cause of competitive pressure. One of the possible reasons could be the existence of the so-called *stereotype threat* effect, which might be affecting women more than men: when women perceive that those who surround them may perform better than they do, they end up underperforming not because of their lack of ability but of self-confidence. Be this the case, women laying just above the AU-score threshold in HA groups would tend to underperform more than men would in that range of academic ability.

This fact may be reflected in several environments but definitely will delimit the possibilities of a women to develop a successful professional career. The prevailing gender wage gaps and the so-called glass ceiling have very much to do with the recognized fact that women seem more reluctant than men to occupy positions of great responsibility in companies and institutions. One of the reasons why this might be so is the fact that women react negatively in highly competitive environments. Analyzing whether this issue is a reality already at college level is essential, as we should work to solve it from the first educational steps. Policy makers and company owners are already designing initiatives to empower girls by showing children that women have been successful in many disciplines and to explain that the best results are achieved when combining all kind of personal characteristics.

Thus, the main goal of our paper is twofold. First, the set of mentioned arguments reinforces the concern to study the impact of being part of a HA group. It is important to understand whether peers with certain characteristics are affecting students in a positive or a negative way, emphasizing the consequences of competitive pressure in both genders. Second, our research seeks to understand the responses of women relative to men in the more competitive environments created by such HA groups. This knowledge seems to us crucial for economists and policy makers to better design educational policies, a pillar for future growth and development.

In next Section 2 we summarize the lines of research in the literature more closely connected with our research and set up more specifically the contribution of our paper to these lines. Then, Section 3 describes the dataset together with some relevant concepts in our study. In Section 4 we explain the methodology. Section 5 reports the main results, whereas Section 6 shows some robustness checks. Finally, we present our main conclusions in Section 7.

2. Literature Review

A growing literature covers, on the one side, the analysis of how peers affect student's achievements and, on the other side, women responses towards competitive pressure in different environments. As well, the RD methodology has been broadly used in recent research. Hereunder we provide a revision of relevant literature related to our research.

Several studies on peer effects have been published by researchers on education economics. Peer effects refer to the idea that individuals are influenced by their classmates in a number of ways, more prominently in the form of knowledge spillovers from more able peers to their less able mates. Sacerdote (2001) states that peer effects exist in college by exploiting that roommates are conditionally randomly assigned in Dartmouth college. Ding *et al*. (2007) find evidence for positive and nonlinear peer effects in China`s educational system. Vardardottir (2013) examines the ability of peer effects among teenagers in Iceland and find positive effects interpreted as learning externalities. All these works center the attention on the estimation of positive learning externalities, perhaps the effects more clearly present at earlier stages of education.

Duflo *et al*. (2011) analyze peer effects and the impact of tracking in Kenya introducing the analysis of teachers' incentives and behavior. They find that, at first sight, the direct impact of high-achievement peers is positive. However, a deeper exploration presents that dividing people in different classes by level benefits indirectly lower-achievement students since teachers may adapt the difficulty of the lectures. These results back our hypothesis that the higher requirement standards can imply lower grades for students at the bottom tail of the ability distribution in HA groups as compared to those at the top of less competitive classes.[1]

In the context of the gender literature and the growing line of behavioral economics research, a number of papers have studied the behavior of women in certain situations. Gneezy *et al*. (2003), in an influential paper published in *The Quarterly Journal of Economics*, analyze gender differences in performance in competitive environments. Their findings disclose that in

[1] Although academic achievement is perhaps one of the most widely-analyzed contexts, the study of peer effects is also present in other setups. As an example, Cornelissen et al. (2017) analyze such effects in the workplace, and find that more motivated workers influence positively their peers in terms of productivity, and hence, wages.

such kind of environments women perform worse than men, and the underperformance is more remarkable when their rivals are men. The authors emphasize the influence of the stereotype threat, persistent in the social psychology literature, and conclude that women perceiving themselves less able than men underperform more than under self-confidence. Such stereotype threat effect refers to a situation where a person is concerned with a negative stereotype about their social group, and was firstly established by Steele and Aronson, (1995). This concern may cause the person to perform worse when competitors are thought to be better than her/himself. Steele (1997) stated that the anxiety that the stereotype threat originates increases the probability of chocking under pressure. Spencer *et al*. (1999) found that, unlike men, women's performance in math is negatively affected by the risk of being judged according to the negative stereotype that women have weaker math ability. They hypothesized that the apprehension it causes may disrupt women's math performance.

Recent studies in this field revise previous literature and provide new findings pointing out in this same direction. For example, Niederle and Vesterlund (2007) also conclude that women are more likely to avoid competitive environments than men, and Gill and Prowse (2014) show that women and men react differently to positive or negative feedback about the ability of their peers. Iriberri and Rey-Beil (2017), develop a deep exploration on the issue, again emphasizing the role of the stereotype threat. They find that women underperform men when the task in which they compete is perceived as favoring men and they are explicitly informed of the presence of a strong rival. The same authors, Iriberri and Rey-Beil (2018), have recently published the results of an experimental research in which they analyze two-stage math contests, and find again that women underperformed worse in the stage two of the contest as a result of competitive pressure.

Hence, there exists recently contrasted evidence that women are particular sensitive to competitive pressure in their academic and professional contexts, particularly when they consider themselves disadvantaged as compared with their peers. As far as we know, there is no evidence, however, from real data on academic results at the university level. We consider that our paper is an interesting brand-new contribution to this line of research, since the degree of competition among students very likely achieves its higher intensity at such higher educational levels.

From the methodological point of view, the regression discontinuity design becomes an attractive tool to study the effect of peers on one's performance. Identification of the causal effect of peers entails certain difficulties since several sources of endogeneity may appear in their estimation.[2] The relatively new RD approach, whose details we provide in Section 4, is based on the comparison of two groups of individuals separated by a tiny difference in the value of a particular quantitative variable determining 'treatment'. In our setting, these are students just below and students just above the AU-score cutoff, the treated students being those in a HA group, and the control ones those who remain in a standard group. The attractiveness of the RD methodology responds to the weak assumptions it requires to recover causal effects (Cattaneo, 2016).

We find some examples of the RD implementation in past literature. Sa *et al*. (2014) analyze the impact of gifted and talented programs on students by employing the RD design. They exploit that gifted students are exposed to higher achievement peers although they find no significant improvement in achievement for students benefiting from these programs. More precisely, our research setup bears a close resemblance to the paper published by Vardardottir (2013) that examines the ability of peer effects among teenagers, by employing a fuzzy RD approach where the source of identifying information is the assignment into high-ability classes in an Iceland high school.

Alike for the study of peer effects, it must be said that our procedure has been generally used to estimate other effects of interest and also in other areas of social science. Again in education economics, the methodology has been recently implemented by many analysts who have taken advantage of the rules in worldwide education systems to study how different setups affect educational achievement. Some examples are the works of Angrist and Lavy (1999), Hoxby (2000), Urquiola (2006), and Urquiola and Verhoogen (2009), who study the effect of cohort size on test scores, or Canton and Blom (2004) and Chay *et al*. (2005) who study the effect of eligibility score in college outcomes and school aid respectively. Finally, Dobkin and Ferreira (2010) study if age based school entry laws affect educational attainment and labor market outcomes. [3]

---

[2] Selection bias and other unobserved influences are two of these sources of endogeneity. In the setting of peer effects, the so-called *reflection problem* is a particular case of selection problem. It refers to the endogenous self-selection of individuals into their groups of influence, that is, individuals approach those whose unobservable characteristics are systematically related to theirs. More notably, the peer effect is inherently endogenous: if peers' outcomes affect individual outcomes, individual outcomes will affect peers' outcomes, leading to simultaneity bias.

[3] In topics other than ours, such as political science, RD designs are particularly common. For example, in studies based on elections the discontinuous assignment of victory in close races often provides a credible research design to make causal inferences about mass or elite behavior. Some examples in this line are the works of Brollo *et al*. (2013), Klašnja (2015), and Klašnja and Titiunik (2017). Also criminal justice academics have used the RD design to study several topics related to their field (e.g., Berk and de Leeuw (1999), and Chen and Shapiro (2004) who identify the causal effect of prison conditions on recidivism

In spite of the fact that the methodology implemented in this research has been used in several studies that analyze peer effects, to our knowledge no other researcher has explored the impact of peers in college emphasizing gender differences. Our contribution lies in the fact that we use a fuzzy RD design to examine how peers impact college achievement of women in competitive HA groups, as compared with men and other women in standard and less competitive groups.
To summarize, our paper contributes to the existing subject literature in at least three fronts. First, while the bulk of the empirical literature on peer effects focuses on the earlier stages of the educational system, our paper provides brand-new evidence from administrative data at the university level. Second, the consideration of more mature students at the university level permits us to extend the channels through which peer effects are expected to affect one student's performance, namely *learning externalities*, *requirement standards* and *competitive pressure*. Finally, our focus on women performance as compared to men in HA groups is, to our knowledge, an unexplored issue in education economics that may help understand the behavior of women in competitive environments.

In addition, our paper constitutes a new contribution to the growing, though still recent, applied literature in social sciences that implements a regression discontinuity approach to identify causal effects of some educational and economic programs.

Our econometric analysis of grades obtained in the same subjects by students of similar ability in standard groups and in HA groups renders two main findings, which we advance here. First, students at the lower tail of the grades distribution in HA groups underperform as compared to students in the upper tail in standard groups. Second, these differences are quantitatively larger and statistically significant in the case of women, suggesting a competitive pressure effect underlying our results which is more evident in the case of female studens. rates by exploiting a discontinuity in the assignment of federal prisoners to security levels). One of the first RD applications was in international economics by van der Klaauw (2002).

---

rates by exploiting a discontinuity in the assignment of federal prisoners to security levels). One of the first RD applications was in international economics by van der Klaauw (2002).

3. Data

We use data on 4 years of entering students at the UV, Spain, concretely at the Facultat d'Economia (Faculty of Economics) to study how peer effects in terms of *learning externalities, competitive pressure*, and *requirement standards* affect students. Our data ranges from 2011, the year the degrees of *International Business* and *Business & Law* were set up, to 2014,.

In Spain, students at the last two years of high-school study general and specific subjects related to the major of their interest. When students finish high-school, they are able to choose their major based on the grade obtained to access the University. The AU-score is computed as the weighted average of the exam of access to the university and the averaged two last years of high school. The average of the last two years of high school is worth 40 percent. The exam of access to the university has two parts. The first part comprises general subjects, is compulsory to enroll in any university and is worth 60 percent. In the specific part, where students must complete exams related to the field of study they are looking forward to register, the access grade can increase up to 4 points. In total, students can get a maximum of 14 points at the AU-score which will determine their selection into High Achievement classes. [4]

Acceptance into degrees such as International Business and Business & Law is conditional on the student´s AU-score, which has to pass the cutoff mark demanded by each degree. This cutoff is delimited by the supply of places by the university and the demand of places by students. Consequently, the cutoff mark is higher in high achievement groups (HA) than in Standard groups of business (SG), and may vary yearly.

We select students in three different majors who started their college period from 2011 to 2014. We consider observations of common subjects in all the majors. The Faculty of Economics is the public college to study business and economic-related fields at the University of Valencia (Spain). The institution offers a wide range of well-recognized four or five years majors, taught by experienced academics. We focus on three different majors, which are Grado en ADE (Business Administration), Grado en International Business (International Business), and Grado en DADE (Business Administration and Law). The data set consists of 2324 students, 1156

[4] In the general part of the AU exam, students must complete language exams (Spanish, a foreign one and the regional one if applicable), a history or philosophy exam, and an exam oriented to the field they aim to study. The specific part comprises geography, applied mathematics, and economics in case of economics-related majors. In this last part if students fail the exam they get zero points, while in case they pass they can get up to two points per exam. In case of students attending to a foreign school in Spain they perform the foreign access to the university exam and the Ministry of Education establishes the basis to convert this grade into a comparison out of ten, and they perform as well the specific part that can sum up to 14.

females and 1168 males and, thus, it can be considered to be evenly distributed by gender. Students are usually 17-20 years old. Business Administration, which we denote here as Standard Groups, has around 370 students per year and it is not subject to specially high AU-score for enrollment. International Business and Business Administration and Law are considered High Achievement degrees. The number of places offered per year in these degrees accounts for around 130 and 90, respectively. Throughout this paper we will refer to Business Administration as *Business A.*, International Business as *Intern B.*, and Business and Law as *Law and B*. The difference between the careers lies on the international orientation of *Intern B.* together with the fact that students are obliged to study abroad for at least a year, and the extra law classes studied by *Law and B.* students. Students in *Intern. B* choose the destiny for the year depending on the grades achieved in the first year of the degree, which exerts an extra pressure for them and increasing competition among peers. Thus, we consider *Intern. B* to have an extra of competitive pressure.

This being said, although the main objectives and the material of the classes in many subjects coincide, HA groups are more demanding in terms of exam correction criteria. However and specially in the first two years, there are common subjects that present almost equal curriculum. These are the classes included in our study, as considered the appropriate ones to center the analysis. We classify them in three groups based on their similarities and their applications. The first group named '*Economics*' includes Introduction to Economics, Microeconomics, and Macroeconomics. '*Business*' includes Introduction to Business, Financial Accounting, Cost Accounting, and Strategic Management. The third group, '*Maths*', encompasses mathematics and financial mathematics.

The outcome variable of interest is the grade obtained in common subjects. We use different criteria to determine this outcome variable. i) First we use the pooled grades of all common subjects within each degree; ii) Second, we take the average grade across all the common subjects so that we only have one observation per individual. iii) Finally, to provide a more homogeneous sample, we divide the subjects by fields and create the three different groups specified above (*Economics, Business and Maths*) and run again the pooled regression and the regression of the average grade within each field. The latter case is our preferred outcome variable in the estimation below, as it takes the average grade per individual within homogeneous groups of subjects and, thus, it is probably the most representative measure. In addition to the mentioned, we have information on socioeconomic variables of the students parents, the age, the gender, and the place of residence of the students.

In tables 1 to 3, we provide descriptive statistics to better understand our sample and the characteristics of the individuals selected for our study. The cutoff mark appears in Table 1, where it can be observed that enrolling in *Business A.* depends as well on overpassing a cutoff grade but much lower than for enrolling in HA groups. For instance, in 2014 in *Business A.* the cutoff mark equals 7.809 whereas in GIB 10.87 and in Law and B. 11.25. Moreover, in HA has increased since 2011 because of an increase in the demand of these majors encouraging competition to enter. Table 1 also shows the mean of the AU-score and the AU-score with respect to the cutoff, respectively, for each of the groups. We may notice that the average of the AU-score of students in HA groups is much higher than for students in SG. For example*, Intern. B* students got in average 11.76 in 2014 and *Law and B.* students 12.16. These results are much higher than the 9.16 achieved by the average student in *Business A.* The disparities are equally noticeable in terms of the centered AU-score, that is, the AU-score deviated with respect to the official cutoff each year. All the numbers of *Intern B.* and *Law and B.* are positive whereas in *Business A.* the results turn out to be negative as most of the students do not overpass the cutoff mark.

Table 2 displays the total number of students broken down by classes and the female percentage ratio. There are more registered in *Business A.* than in *Intern. B* and *Law and B.* More concretely around 370 opposed to 130 and 90, respectively. As mentioned, the sample is evenly distributed taken into account gender as 49.7 percent of the individuals are women and 50.3 percent are men. As indicated above, our analysis will focus on examining gender differences in the context of peer effects.

Finally, Table 3 encompasses the differences between students in *Business A.* and *Intern. B* and *Law and B.* in terms of socioeconomic variables of the parents, and other personal characteristics of interest. Here we appreciate some disparities. For example, the parents of students in HA groups tend to be on average less unemployed and hold more frequently a university degree. In addition, there are more girls and more full time students in *Intern. B* and *Law and B.* and the average age is lower than in SG. Many of the observed differences turn out to be statistically significant, as indicated in Table 3. In the econometric section below, we will check for robustness of our main results to the inclusion of some of these socioeconomic controls in the RD regressions.

Table 1. AU-scores: Official cutoffs per year and Students' average AU-scores

	Official entrance cutoffs				Average AU-scores (in parenthesis AU-scores centered at the cutoff)			
	2011	2012	2013	2014	2011	2012	2013	2014
Intern. B	9.2	10.35	10.57	10.87	10.55 (1.35)	11.31 (0.96)	11.76 (1.18)	11.76 (0.89)
Law & B	10.3	10.8	11.01	11.25	11.75 (1.10)	11.87 (0.98)	12.15 (0.84)	12.16 (0.91)
Business	7.954	8.092	8.244	7.807	9.17 (-1.46)	9.29 (-1.60)	9.26 (-2.05)	9.16 (-2.10)

Cutoff grades vary by group and by year depending on the supply by the University and the demand by students. They are recorded on a 0-14 scale. AU-score is defined as the average of student`s results in the access to the university exam together with the average of the two last high-school years. AU-score centered equals the AU-score minus the high-ability assignment threshold.

Table 2. Number of students and female students share per degree

	Number of students					Female students share			
	2011	2012	2013	2014	TOTAL	2011	2012	2013	2014
Intern. B	126	132	126	128	512	52	54	54	63
Law & B	86	86	93	88	353	49	49	48	48
Business	367	369	352	371	1459	50	48	51	44
TOTAL	579	587	571	587	2324				

Table 3. Socioeconomic and educational variables

Percentage of students with:	Business	Intern. B	Law & B.
Father with University studies	17.7	33	45
Mother with University studies	16.6	35	42
Father Unemployed	13.9	10	9
Mother Unemployed	30	21	24
Father working as civil servant	43.1	46	46
Mother working as civil servant	33.2	40	36
Full time student	83.9	90	94
Female students share	46.1	54	56
Student living in hometown	85.5	77	84
Age	19.3	18.05	17.57

The reported differences of the values for Intern. B and Law and B. with respect to Business A. have been checked to be statistically significant at conventional levels.

4. Empirical model and identification strategy

In this paper we rely on regression discontinuity (RD henceforth) as an econometric design to identify the causal effects of belonging to a HA group (treatment group) on students' academic achievement (outcome variable) as measured by grades. RD identification is based on the idea that there exists a known cutoff in a quantitative variable that is perfectly observed (the assignment or *run* variable) which determines that a particular subpopulation of individuals falls into one regime or group while the rest fall in a different one. In the neighborhood of the cutoff individuals could be considered, in fact, identically eligible for the treatment but, almost arbitrarily, some of them fall in the treated group whereas some other do not. In this sense, it is much like an experimental design, though the allocation of individuals to the treatment is not randomly assigned by the researcher.[5] This framework enables us to understand and compare how the different groups of individuals perform.

More specifically, the RD method studies the existence of a jump or discontinuity in the conditional mean of the outcome variable at a threshold given by a certain date, a score, or a frontier. In our paper, the *run* or assignment variable is the AU-score, such that students just above and just below the threshold determining eligibility for a HA group can be considered students of identical academic ability. 'Randomly' some of these have entered a HA group while the rest remain in a SG. Which differences does the fact of belonging to a particular type of group create on the grades of comparable students? This is, broadly speaking, our research question.

The basic assumption to implement RD techniques is that the treatment must be a discontinuous function of the run variable x_i. The outcome, y_i, also depends on x_i and, in absence of treatment, the relationship between the outcome and the run variable is continuous. Thus, we can perform a regression to estimate the relationship between y_i and x_i and any jump or discontinuity will be attributed to the treatment.

There are two types of RD design, *sharp* and *fuzzy*. In sharp designs the treatment is fully determined by the assignment variable, so that treatment occurs whenever the running variable overpasses the cutoff. In the fuzzy RD case it is the probability or intensity of treatment that

[5] To whether policies and social order affect overall welfare, randomized experiments are the most desired technique to make causal inference due to the ability of the researcher to design the experiment and control for all variables. However, this is difficult in real life when dealing with social data. A more realistic setup is the possibility of developing a natural experiment. In natural experiments, the assignment of treatments to subjects has been made by nature, not by the researcher. Luckily, "rules that constraint the role of chance in human affairs often generate interesting and valuable experiments" (Angrist and Pischke, 2015, p. 147) so that the researcher may take advantage of the setup to properly estimate a model.

jumps at the cutoff.

More specifically, in sharp designs, just the knowledge of the running variable is necessary to know the treatment status, which can be formalized as a dummy that takes the value of 1 when the running variable overpasses the cutoff and zero otherwise. Moreover, the treatment status is a discontinuous function of the running variable as no matter how close the running variable is to the cutoff, the treatment remains unchanged until they equal.[6] In other words, no one below the threshold is in the treatment group, and no one above remains in the control one. Hence, the treatment dummy, D_i, in the sharp RD design can be defined as a deterministic function of the run variable, x_i, as follows:

$$D_i = 1 \; if \; x_i \geq c$$

$$D_i = 0 \; if \; x_i < c \tag{1}$$

where c is the cutoff or threshold determining treatment.

A simple regression that illustrates the idea of what we aim at capturing is:

$$y_i = \alpha + \beta x_i + \rho D_i + e_i \tag{2}$$

where ρ is the causal effect of interest, that is, the difference in grades caused by the fact of belonging to the treated group. Notice that the particular feature of this regression is that D_i not only is correlated with x_i but it is a deterministic function of it. As a generalization of specification in equation 2 we may allow the relationship between y_i and x_i to be nonlinear in some function $f(x_i)$, for example a pth-order polynomial:

$$y_i = f(x_i) + \rho D_i + e_i \tag{3}$$

In our setting, $D_i = 1$ if student i belongs to a HA group, and $D_i = 0$ if the student is in a SG. In a standard parametric OLS estimation of equation 3, accounting for the possible nonlinearity of the relationship is crucial to rule out that what looks like a jump at the cutoff might be simply an unaccounted nonlinearity in the relationship (see, e.g. Angrist and Pischke (2009) , p. 255 for

[6] Examples of sharp RD designs are those where the treatment depends on having or not a determined age. A classic example is the case considered by Carpenter and Dobkin (2009), where the authors link (legal) alcohol consumption to mortality. In their sharp RD design the cutoff is the minimum drinking age: the compliers overpass the cutoff clearly depending on they have or not the determined age.

more details on this). To reduce such possibility, we can rely on a purely nonparametric approach looking only at data in a neighborhood around the discontinuity, and just comparing the average of grades to the left and to the right of the cutoff.

$$E(y_i \,/\, x_o - \Delta < x_i < x_0) \cong E(y_{0i} \,/\, x_i = x_0)$$

$$E(y_i \,/\, x_o \leq x_i < x_0 + \Delta) \cong E(y_{1i} \,/\, x_i = x_0)$$

so that

$$\lim_{\Delta \to 0} E(y_i \,/\, x_0 \leq x_i < x_0 + \Delta) - E(y_i \,/\, x_0 - \Delta < x_i < x_0) = E(\, y_{1i} - y_{0i}\, x_i = x_0) \qquad (4)$$

Differently to the sharp case, in this paper we implement a fuzzy RD design due to the specific characteristics of our experiment. In particular, considering that some students may overpass the cutoff mark but choose to study Business Administration in a SG due to personal preferences, it is not the treatment status that jumps at the threshold but the probability of it.[7] In other words, we exploit here *discontinuities in the probability* of treatment conditional on x_i, which can be specified as:

$$P\,(\,D_i = 1 \,/\, x_i\,) = \begin{cases} g_1(x_i) & if\ x_i > c \\ g_0(x_i) & if\ x_i \leq c \end{cases} \qquad (5)$$

Where g_1 y g_0 differ at $x_i = c$.

Equation 5 shows that the probability of belonging to a HA group is a function of the assignment variable x_i = AU-score, and it should be clearly different at the right and at the left of the cutoff c. In other words, the probability jumps at the cutoff.

In the fuzzy RD, the basic regression:

$$y_i = \alpha + \beta x_i + \rho D_i + e_i \qquad (6)$$

differs from the sharp RD design because D_i is no longer a deterministic function of x_i. Instead, D_i can be estimated (predicted) as a function of a dummy variable T_i that indicates when the

[7] For example, in International Business it is compulsory to go abroad for at least 1 year, and some students prefer not to commit themselves to this. Regarding Law and Business, some students dislike law subjects and prefer not to be enrolled in the dual degree.

cutoff is overpassed:

$$T_i = 1 \quad (x_i \geq c) \tag{7}$$

To understand the fuzzy RD we may think on it as an instrumental variables kind of set up in which the dummy variable T_i is used as instrument for the treatment indicator D_i. The first stage of this IV setting corresponds to an equation like the following one:

$$D_i = \delta + \gamma x_i + \pi T_i + \varepsilon_i \tag{8}$$

And the second stage is illustrated by:

$$y_i = \mu + k_i x_i + \rho\pi T_i + \zeta_i \tag{9}$$

where the second stage coefficient on T_i, that is $\rho\pi$, divided by the first stage coefficient, π, gives the treatment:

$$\frac{\frac{\partial y_i}{\partial T_i}}{\frac{\partial D_i}{\partial T_i}} = \frac{\rho\pi}{\pi} = \rho \tag{10}$$

In the non-parametric set up, we have a first stage defined by:

$$E(D_i \,/\, x_o \leq x_i < x_0 + \Delta) - \; E(D_i/x_o - \Delta \leq x_i < x_0) = \pi \tag{11}$$

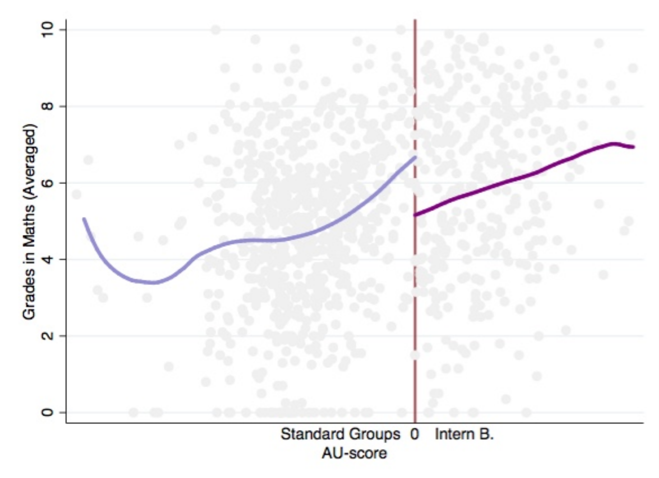
In the second stage, the reduced form conditional expectation of y_i near c is:

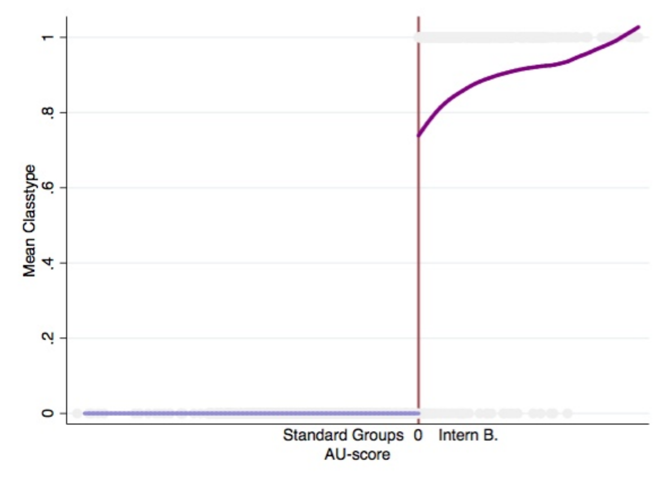
$$E(y_i \;/\; x_o \leq x_i < x_0 + \Delta) - \; E(y_i \;/\; x_o - \Delta \leq x_i < x_0) = \pi * \rho \tag{12}$$

Therefore, we can define a $Wald$ type estimation as:

$$Wald = \frac{E(y_i \;/\; x_o \leq x_i < x_0 + \Delta) - \; E(y_i \;/ x_o - \Delta \leq x_i < x_0)}{E(D_i \;/\; x_o \leq x_i < x_0 + \Delta) - \; E(D_i \;/ x_o - \Delta \leq x_i < x_0)} = \frac{\pi * \rho}{\pi} = \rho \tag{13}$$

This $Wald$ Estimator, which provides an estimate of the treatment effect of interest, has a numerator that is the jump in the outcome (grades) occurring at the cutoff, divided by a denominator that is the jump in the probability of belonging to a HA group that also occurs at the cutoff. The *Wald estimator* captures the causal effects on compliers, that is, individuals whose treatment status changes as we move the value of x_i from just the left to the right of the cutoff c. We present an example from our main results to observe graphically both components of the *Wald*, the jump in the outcome and the jump at the probability occurring at the cutoff. We see that they fit the treatment allocation rule of the fuzzy RD in all cases, since there is a clear jump at the threshold, at zero. The first graph shows that at the cutoff at zero there is a discontinuity, a jump on grades caused exclusively by the treatment. The *Wald Estimator* captures this jump at the numerator. The second graph illustrates the denominator of the *Wald Estimator*, the jump in the probability. We observe that for individuals who obtained AU-scores below the cutoff the probability of enrolling in an HA group is zero. As well, all the individuals assigned to a HA group overpass the cutoff mark. However, there are some students that would be eligible to enroll in HA group as they overpass the cutoff mark but they rather prefer to be assigned to a SG due to unknown reasons. This latter characteristic makes it necessary to implement a fuzzy RD setup instead of a sharp design.





In a fully nonparametric approach, running the regression only on the discontinuity sample (observations in a small neighborhood of the cutoff), we do not need to care about the functional specification of the regression equation. The main drawback in this case is that working only with observations around the cutoff may imply that we work with too few data. An alternative solution, which we adopt in this paper, is to estimate y_i by local linear regression (Hahn, *et al*. 2001) which amounts to weighted least squares (WLS) estimation with more weight given to points close to the cutoff. Hence, our preferred estimation method, whose results we report below, is a weighted local linear regression (LLR) of the type presented in equation 13, as that used in Hahn *et al*. (2001) and Porter (2003).

In this paper, the outcome variable y_i is the student's grade in different subjects. We first run a regression discontinuity with pooled and averaged grades in subjects that are common to both SG and HA groups. As well, we perform another analysis grouping pooled grades by fields classified into *Economics, Business, and Mathematics* in order to group the outcome variable into more homogeneous categories. Finally, we focus on average grades by field student and across all comparable subjects within each degree, what provides a sample with only one outcome per individual.

The running variable in our case is the AU-score. In our case, as shown in section 3, the cutoff varies over the years. We follow in this paper the most common practice in this case, consisting on centering the run variable in the cutoff and use the zero cutoff for all observations to estimate a pooled RD treatment effect.[8]

5. Results

In this section, we present our main results statistically and graphically. The crucial assumption to apply the regression discontinuity design is the existence of an observable assignment variable on which assignment is based and a discontinuity at some cutoff value of the assignment variable at the level of treatment. In our case, the assignment variable is AU-score, and the threshold is the cutoff mark, different by degree and per year. To be able to set the treatment effect as the increment with respect to zero, we subtract to each AU-score the cutoff mark so that each year the run variable mark is at the cutoff.

First of all, we present the results from the RD regressions on each of our outcome measures. In Table 4 and Table 5 we include all the common subjects together, pooled and averaged respectively. We extend these results in Tables 6 and 7 by showing pooled and averaged grades by fields (*Economics, Business, and Mathematics*). In each table, we report results for all the

[8] Cattaneo, et al. (2016) propose sophistications to the common practice case when multiple cutoffs exist that allow to exploit further the information contained in the different cutoffs.

sample as well as separately for women and men to explore gender differentials. The estimate of interest is the effect in grades at the discontinuity that results from being part of a HA group instead of being part of a SG. As already emphasized, the effect is just applicable to students around the threshold for being the individuals that can be considered observationally identical in terms of academic ability. The treatment effect is measured through a Wald-type estimator, obtained from dividing the jump in outcome over the jump in the probability. The jump in the outcome is the numerator of the Wald estimator and measures the effect of the AU-score overpassing or not the cutoff mark on grades. The jump in probability estimates the assignment to a HA group resulting from your AU-score. Only individuals with an AU-score higher than the cutoff mark will qualify to access to a HA group. Since the probability of being assigned to a HA group increases as your AU-score is higher the denominator of the Wald estimator is always positive and the sign of the treatment is going to depend on the jump on the outcome that constitutes the numerator. To implement the weighted local linear estimation of our RD design we use a triangular kernel and the optimal bandwidth proposed by Imbens and Kalyanaraman (2009) to minimize the sum of the squared bias and the variance of estimates (MSE).

In Table 4 we start presenting the results for grades pooled across common subjects. This entails letting each grade obtained by an individual in each subject to be a different observation. The results reveal that the treatment effect is negative and significant in all cases except for men in *Las and B*. Was this a result of the higher requirement standards demanded by teachers in HA groups, no expected differences would appear on the basis of gender. However, a further and interesting result is that the effect is higher and more significant in *Intern. B* than in *Law and B*. and much more accentuated for women than for men. For example, the differential in grades for women enrolled in *Intern. B* tends to be on average 1.529 points lower than for women in *Business A*., whereas men tend to get only 0.784 points less.

In Table 5, the outcome variable is the average grade per individual across all common subjects. We have a total of 2050 students in the case of comparing *Business A*. with *Intern. B* and a total of 1860 in the case of comparing *Business A*. with *Law and B*. Clearly, we count with fewer observations than in Table 4, as each student is assigned a unique grade. Nonetheless the results are broadly comparable to results in Table 4, showing that this does not affect the validity of our main conclusion above. Again, the treatment effect is negative and significant in all cases except for men in *Law and B*. As well, the effect is higher and more significant in *Intern. B* than in *Law and B*. The treatment effect in *Intern. B* is significant at the 10 percent level whereas in *Law and B* at the 1 percent level. Being part of a HA group has more impact on women who tend to get on average 1.468 points less than those in SG of Business A., whereas men tend to have only 1.114 points less. Focusing on the comparison between SG and Law & B. the effect is only significant for women.

Along with this, we understand that people studying the same major may share features that make them better at one specific field. In turn, subjects into different fields may share characteristics (more or less mathematically based, more or less theoretically or empirically oriented, and so on) that might explain differences on performance by gender.[9] Therefore, in Table 6 we contrast our results by performing an additional exploration organizing the nine subjects in fields, constituting more homogeneous groups. As we already anticipated in Section 3, we create three different groups of subjects, namely *Economis, Business* and *Maths*, that encompass similar subjects and then perform our RD estimation on all the outcome variables.

Alike in Table 4, in Table 6 we use pooled grades letting each grade obtained by an individual in each subject to be a different observation. In Economics-related subjects, students near the discontinuity belonging to a HA group get lower grades than students belonging to a Standard Group. In this case, the estimations are significant only when we focus on International Business. The treatment effect varies as well by gender whereas the size of this difference is smaller than in previous cases. The effect equals -1.131 for women and -1.071 for men studying Intern. B.[10]

When analyzing the results obtained for Business-related subjects, we observe the most noteworthy difference by gender. Overall, the treatment effect is negative. However, it is statistically significant for women at the 1 percent level in both majors whereas it is not significant at all for men in Intern. B and only significant at the 10 percent level for men in Law and B.

Focusing on the third group, *Mathematics*, we also appreciate relevant differences. The estimation is negative and statistically significant for students at the discontinuity sample studying *Intern. B*. Making contrasts, women in *Intern. B*. tend to get on average 1.831 points

[9] Beneito et al (2018) have shown gender differences in preferences among different subjects of economics.

[10] To preserve anonymity of teachers, we do not report the results by subjects separately. In exploratory work, however, we run the RD estimation for each of the 3 subjects composing the Economics group. In two out of the 3 cases, the estimates for women where clearly negative while for men no significant treatment effects were found. In one of the 3 subjects, the differences were neither significant for men nor for women.

less than those in SG of Business A. whereas men tend to get 1.326 points less. In *Law and B.* groups, the effect is only significant for women.

Finally, in Table 7 we focus our attention on the outcome measure that in our view provides the most representative results. Again here we run the regressions counting on only one observation per individual, the mean of the grades obtained in each field. This accurate exploration confirms that whereas overall students get higher grades in HA groups than in SG, students near the discontinuity assigned to HA groups obtain lower grades than those assigned to SG.

In other words, in general there is a significant negative jump on grades at the threshold. The results are significant for women in *Intern. B* groups and in two out of three cases in *Law and B.* groups, confirming that women belonging to a HA group in college tend to obtain lower grades than men with respect to students in SG. We only find significant results for the average grades obtained by men in *Intern. B* in *Economics* classes and by men in *Business* classes in *Law and B*. In the first group, *Economics*, the treatment effect is significant for women and for men, but the analysis does not reveal any significant difference between SG and *Law and B*. In *Business*, the effect is significant and stronger for women in both cases, *Intern. B* and *Law and B*. In the last group, the effect is significant for women in both majors.

To summarize, our results indicate that the influence of peers in HA groups is negative implying that peer affects in college do not impact as much through learning externalities as through requirement standards and competitive pressure. Moreover we know that competition in International Business is stronger as students choose the international destination depending on their grades at the first year of college what might explain why we find stronger negative effects in this case. As well, the literature proves that women perform worse in competitive environments.[11] As we find stronger evidence for a negative jump on grades at the threshold in *Intern. B* and for women, we can conclude that in HA groups in college the competitive pressure is big enough to overpass learning externalities for being in class with smart peers and students with AU-scores around the threshold end up performing worse than students with similar grades who had to enroll in a SG.

[11] In Section 2, we provide several examples of the literatura on this topic including Gneezy *et al.* (2003), Spencer *et al.* (1999), Niederle and Verterlund (2007), Gill and Prowse (2014), Iriberri and Rey-Beil (2017), and Iriberri and Rey-Beil (2018)

Table 4 . Fuzzy RD estimates. Grades pooled across common subjects.

	SG and Intern B. groups			SG and Law & B. groups		
	All	Women	Men	All	Women	Men
Treatment	-1.162***	-1.529***	-0.784**	-0.402*	-1.024***	-0.258
	(0.216)	(0.283)	(0.387)	(0.210)	(0.308)	(0.211)
Jump in outcome	-0.806***	-1.065***	-0.564**	-0.304*	-0.616***	-0.207
(Numerator)	(0.152)	(0.199)	(0.281)	(0.159)	(0.186)	(0.169)
Jump in prob	0.694***	0.697***	0.719***	0.756***	0.601***	0.803***
(Denominator)	(0.018)	(0.023)	(0.025)	(0.010)	(0.028)	(0.012)
Observations	15,207	7,418	7,789	13,573	6,544	7,029

Standard errors in parentheses. * p-value<0.1; ** p-value<0.05; *** p-value<0.01

Table 5. Fuzzy RD estimates. Grades averaged over common subjects.

	SG and Intern B. groups			SG and Law & B. groups		
	All	Women	Men	All	Women	Men
Treatment	-1.220***	-1.468***	-1.114**	-0.506*	-0.886*	-0.201
	(0.267)	(0.412)	(0.433)	(0.286)	(0.535)	(0.472)
Jump in outcome	-0.926***	-1.080***	-0.856**	-0.371*	-0.582	-0.137
(Numerator)	(0.263)	(0.400)	(0.340)	(0.212)	(0.354)	(0.323)
Jump in prob	0.759***	0.736***	0.769***	0.734***	0.657***	0.682***
(Denominator)	(0.038)	(0.060)	(0.049)	(0.036)	(0.069)	(0.062)
Observations	2,050	990	1,060	1,860	891	969

Standard errors in parentheses. * p-value<0.1; ** p-value<0.05; *** p-value<0.01

Table 6. Fuzzy RD estimates. Grades pooled by fields.

ECONOMICS

	SG and Intern B. groups			SG and Law & B. groups		
	All	Women	Men	All	Women	Men
Treatment	-1.051***	-1.131**	-1.071***	-0.013	-0.337	0.038
	(0.204)	(0.512)	(0.413)	(0.261)	(0.520)	(0.508)
Jump in outcome	-0.780***	-0.751**	-0.784**	-0.009	-0.206	0.025
(Numerator)	(0.197)	(0.345)	(0.306)	(0.178)	(0.319)	(0.331)
Jump in prob	0.743***	0.664***	0.732***	0.684***	0.611***	0.652***
(Denominator)	(0.022)	(0.047)	(0.034)	(0.026)	(0.047)	(0.042)
Observations	5,394	2,624	2,770	4,882	2,354	2,528

BUSINESS

	SG and Intern B. groups			SG and Law & B. groups		
	All	Women	Men	All	Women	Men
Treatment	-1.029***	-1.437***	-0.464	-1.044***	-1.317***	-0.803*
	(0.275)	(0.313)	(0.483)	(0.235)	(0.414)	(0.443)
Jump in outcome	-0.741***	-0.979***	-0.331	-0.678***	-0.804***	-0.548*
(Numerator)	(0.199)	(0.267)	(0.345)	(0.200)	(0.258)	(0.302)
Jump in prob	0.720***	0.681***	0.712***	0.650***	0.611***	0.683***
(Denominator)	(0.023)	(0.035)	(0.040)	(0.027)	(0.040)	(0.037)
Observations	6,374	3,095	3,279	5,741	2,760	2,981

MATHS

	SG and Intern B. groups			SG and Law & B. groups		
	All	Women	Men	All	Women	Men
Treatment	-1.448***	-1.831***	-1.326*	-0.369	-1.077**	0.568
	(0.409)	(0.657)	(0.691)	(0.416)	(0.496)	(0.414)
Jump in outcome	-1.082***	-1.281***	-0.838*	-0.249	-0.828**	0.393
(Numerator)	(0.311)	(0.458)	(0.447)	(0.282)	(0.380)	(0.370)
Jump in prob	0.747***	0.699***	0.632***	0.676***	0.769***	0.691***
(Denominator)	(0.032)	(0.054)	(0.055)	(0.034)	(0.026)	(0.046)
Observations	3,439	1,699	1,740	2,950	1,430	1,520

Standard errors in parentheses. * p-value<0.1; ** p-value<0.05; *** p-value<0.01

Table 7. Fuzzy RD estimates. Averaged grades by fields.

ECONOMICS

	SG and Intern B. groups			SG and Law & B. groups		
	All	Women	Men	All	Women	Men
Treatment	-0.982***	-1.057**	-1.040**	-0.226	-0.369	0.144
	(0.308)	(0.515)	(0.462)	(0.338)	(0.641)	(0.438)
Jump in outcome	-0.838***	-0.879**	-0.894**	-0.166	-0.241	0.117
(Numerator)	(0.265)	(0.431)	(0.400)	(0.247)	(0.420)	(0.355)
Jump in prob	0.853***	0.831***	0.859***	0.731***	0.654***	0.809***
(Denominator)	(0.018)	(0.029)	(0.025)	(0.036)	(0.070)	(0.033)
Observations	2,050	990	1,060	1,860	891	969

BUSINESS

	SG and Intern B. groups			SG and Law & B. groups		
	All	Women	Men	All	Women	Men
Treatment	-1.041***	-1.393***	-0.225	-0.845***	-1.280***	-0.973**
	(0.243)	(0.389)	(0.544)	(0.260)	(0.493)	(0.396)
Jump in outcome	-0.792***	-1.156***	-0.163	-0.683***	-0.834**	-0.659**
(Numerator)	(0.223)	(0.329)	(0.394)	(0.212)	(0.335)	(0.333)
Jump in prob	0.760***	0.830***	0.723***	0.808***	0.651***	0.678***
(Denominator)	(0.036)	(0.029)	(0.073)	(0.021)	(0.073)	(0.062)
Observations	1,990	957	1,033	1,801	858	943

MATHS

	SG and Intern B. groups			SG and Law & B. groups		
	All	Women	Men	All	Women	Men
Treatment	-1.527***	-1.978***	-0.911	-0.070	-1.195**	0.764
	(0.503)	(0.696)	(0.594)	(0.361)	(0.579)	(0.510)
Jump in outcome	-1.121***	-1.421***	-0.772	-0.052	-0.941**	0.618
(Numerator)	(0.380)	(0.496)	(0.510)	(0.268)	(0.455)	(0.406)
Jump in prob	0.734***	0.718***	0.847***	0.740***	0.788***	0.808***
(Denominator)	(0.045)	(0.067)	(0.027)	(0.034)	(0.034)	(0.032)
Observations	2,045	988	1,057	1,857	889	968

Standard errors in parentheses. * p-value<0.1; ** p-value<0.05; *** p-value<0.01

Together with the tables below, at the second part of Section 5 a more visual representation of data, in graphs 1 to 3. We plot the jump in the outcome variable (grades) both for women and men, and for the treated groups *Intern. B* and *Law and B* as compared with the Standard Group, *Business A*. Out of all the outcome variables we consider in our RD regressions, we choose to plot the average grades by fields, since they provide the most homogenous and representative outcome measure in our study.

So that Figures 1 to 3 plot the average grades by field as a function of the AU-score, and are convenient to explore visually whether there is jump in averaged grades around the threshold. Interestingly, we can see that in most cases there is a negative jump at the cutoff. In the two cases out of 12 where the jump in the trend is positive, the size of the jump does not look really high visually. These plots reveal two findings, consistent with the explained at the first part of this Section. A first one corresponds to the expected and intuitive result that on average students in HA groups perform better than those in SG. The mean of the points to the right of the cutoff is higher than the mean of the points to the left. However if we focus on the compliers of the local RD estimators, those around the threshold who had similar AU-scores and are considered comparable individuals, we arrive to different conclusions. Students who made it to HA groups have tend to perform worse in terms of grades achievement over the two first years than those who did not make it to HA groups. It could then be concluded that, in terms of grades, it pays for students on the threshold to enter a SG rather than a HA group. Moreover, the jump is always noticeable in case of *Intern. B* but not in all cases in *Law and B*. and much more prominent for women in all cases. This illustrates visually our findings supporting the idea that in the context of peer effects in college, competitive pressure and requirement standards tend to have more importance than learning spillovers in HA groups.

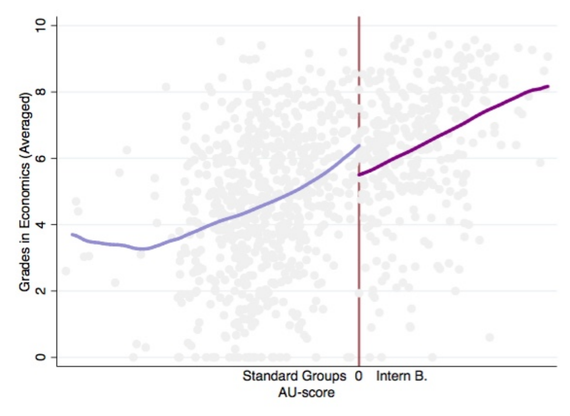
This inspection is valuable to make an overall screening of our results. However, it should be taken into account that graphs are only a way to visualize the relationship that exists between outcome and treatment, and to determine whether there is a jump in the outcome variable.

Next, we provide some robustness checks for these results.

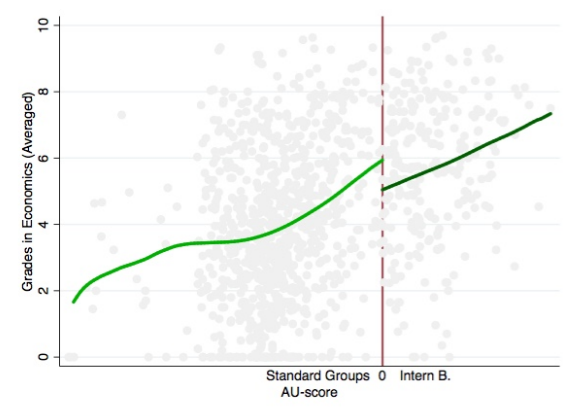
Figure 1. Economics-related subjects

International Business

WOMEN

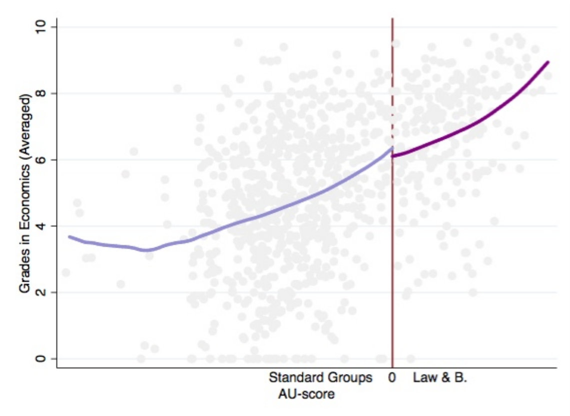


MEN



Law and Business

WOMEN



MEN

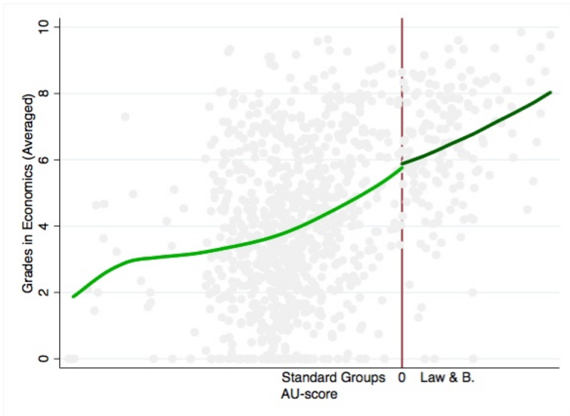
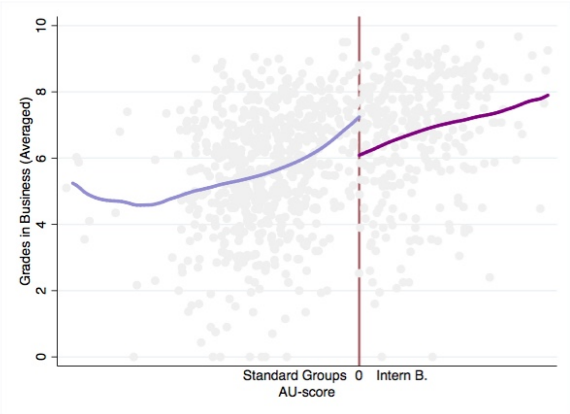


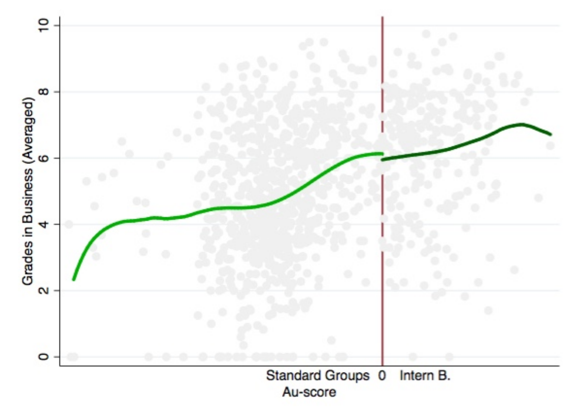
Figure 2. Business-related subjects

International Business

WOMEN

MEN

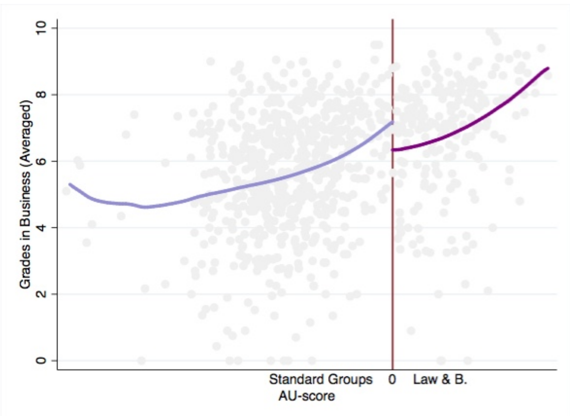




Law and Business

WOMEN

MEN



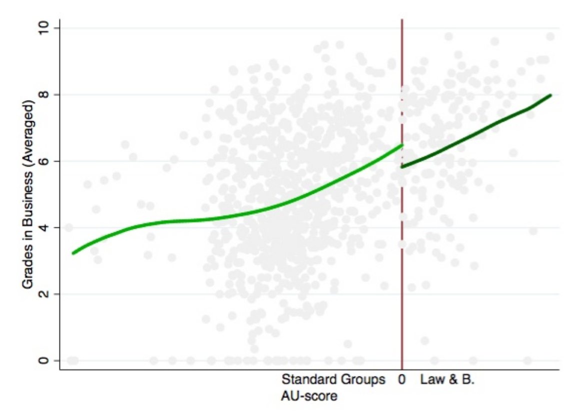
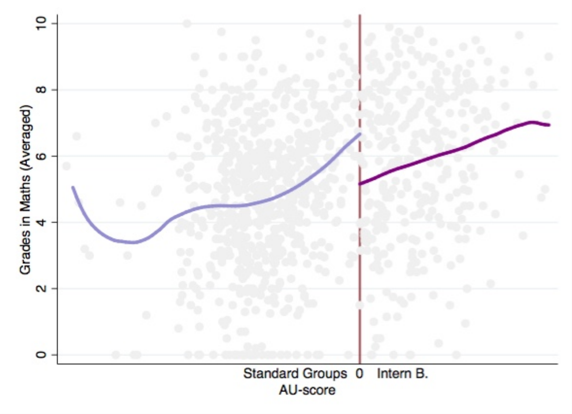


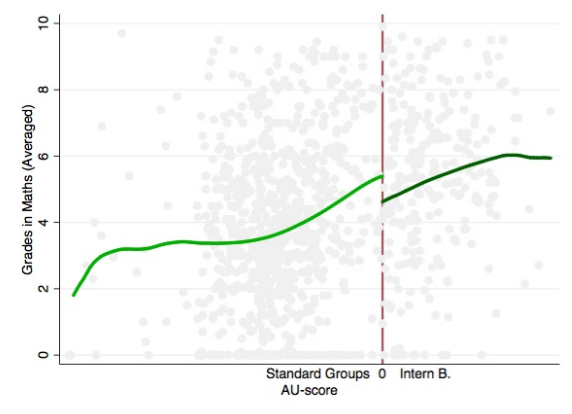
Figure 3. Maths-related subjects

International Business

WOMEN

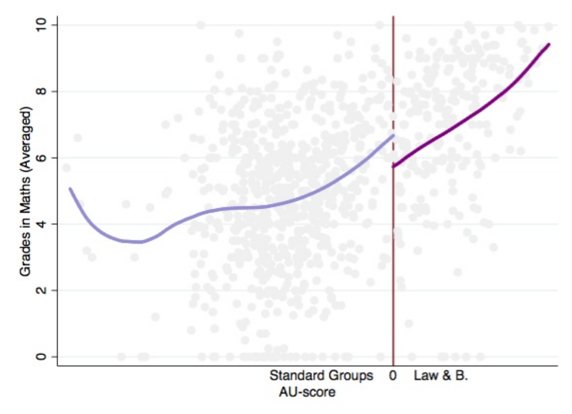


MEN

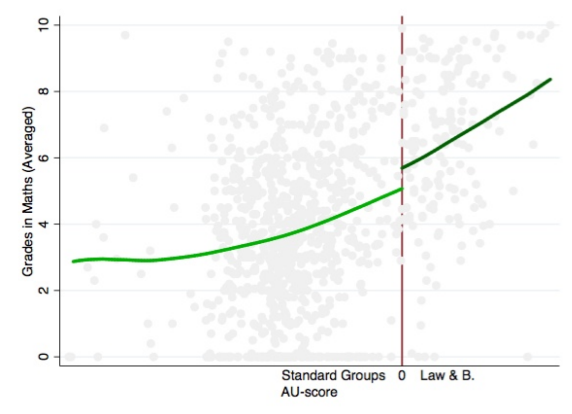


Law and Business

WOMEN



MEN



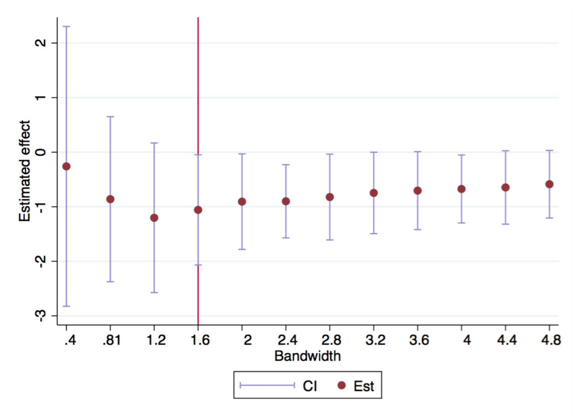
6. Robustness Checks.

In this section we provide some robustness checks to ascertain the sensitivity of our main RD results above to alternative specifications of the estimation equation and method.

Firstly, the use of the nonparametric LLR avoids the need of choosing a particular specification but forces the choice of bandwidth or smoothing parameter. In nonparametric kernel density estimation, the bandwidth is essentially a measure of how closely we want the density to match the empirical distribution.[12] The choice of bandwidth trade-offs bias and variance. We use the estimated optimal bandwidth suggested by Imbens and Kalyanaraman (2009), and check the robustness of our results to different values of it. [13]

Figures 4 and 5 below show that our results are not highly sensible to changes in the bandwidth parameter. As for graphs in Section 5, out of all the outcome variables we consider in our RD regressions, we implement this analysis by using the average grades by fields, since they provide the most homogenous and representative outcome measure in our study. We present the graphs for Economics and Mathematics-related subjects. The graphs illustrate the estimated effect obtained by applying the fuzzy RD for different bandwidths. As observed the estimated results are not too sensitive to the bandwidth used, and do not invalidate the conclusion drawn in Section 5.

Figure 4. Sensitivity of averaged grades in Economics-related subjects to changes in the bandwidth

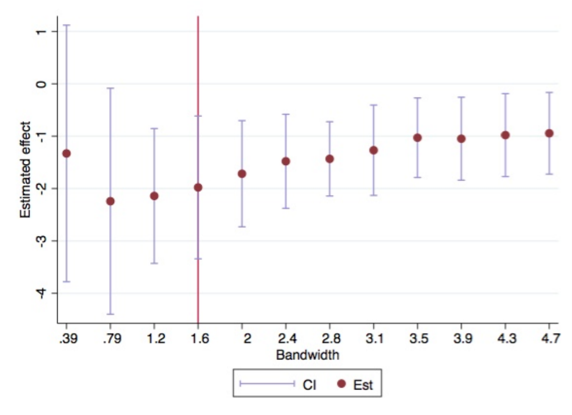


---

[12] More specifically, in our weighted local linear regressions we use a triangular kernel.

[13] See Imbens and Kalyanaraman (2012) for a recent review on bandwidth selection in the RD design.

Figure 5. Sensitivity of averaged grades in Maths-related subjects to changes in the bandwidth



Secondly, the key identification assumption of continuity in an RD design must hold: $f(x_i)$ in Equation 3 must be a continuous function. Intuitively, the idea is that the assignment to different groups is the only source of discontinuity of outcomes around the threshold. The violation of the continuity assumption may arise if, for example, students could have a direct control on grades. In our case, we discard the possibility that students have such a direct control on them. They may be able to study more if they want their performance to be slightly higher, but at the end the AU-score is the mean of different exams over several years so their control over them is limited. As well, the cutoff is unknown previously to the assignment on HA classes, as it changes each year based on the supply and demand of the different degrees and on the grades achieved by other people, clarifying the impossibility of designing a clear strategy to overpass the threshold.

To further convince that the assignment to different groups is the only source of discontinuity of outcomes around the threshold we need to check also that the sample is equal in both sides of the threshold, or at least, that different traits of treated and non-treated groups (other than the run variable) do not exert any difference in the outcome of interest. In other words, our results should be robust to the inclusion of control variables accounting for observed significant differences in the treated and non-treated samples.

Finally, we estimate also our fuzzy RD through polynomial parametric regression to evaluate how the estimated treatment varies when using the unweighted sample of data.

In Tables 8, 9, and 10 we run the regressions selecting averaged grades by fields as the variable of interest. We present the estimations for *Economics, Business, and Mathematics*-related subjects. Moreover, we distinguish among *Intern. B* and *Law and B*., and women and men. In the first columns, the tables reproduce the results presented in Section 5 of the RD local estimation for us to easily compare with the additional estimations we provide in the rest of columns.

On the one hand, RD estimations including as covariates the variables presented in Table 3 of Section 3 are shown in columns under the name of "*Controls*". These are dummy variables accounting for students with father/mother with university studies, father/ mother unemployed, father/mother working as a civil servant, full time students, student living in hometown, and age.

As we identify as the only source of discontinuity the assignment of students in different groups we expect our results to be indifferent to the inclusion of different personal traits of treated and non-treated groups (other than the run variable). Consistent with this expectation, these variables are not found to exert in general any statistically significant effect at any conventional level, so we conclude that including them do not add any additional information to our analysis. The only covariate found to be significant in some cases is the one indicating whether the student's father and mother have university studies. Thus, we decide to present the outcome obtained when considering *"father and mother University"* as a control. Although the Wald estimation of the effect of this regressor is statistically significant, the size of the treatment of interest does not vary notably with respect to the results in section 5. This confirms that including personal characteristics between groups other than the run variable does not affect the validity of our conclusions.

On the other hand, we present the outcome obtained from the estimations through polynomial parametric regression. We include combinations of first, second, and third-degree polynomials of the run variable indicating whether they are found to be statistically significant. Firstly, we perform the estimation considering only the first-degree polynomial. Then we include polynomials of all order first, second, and third together and finally just first and third-degree polynomials. All this information is summarized in the tables.[14] We include as well the regression results of the regressor of interest, the variable indicating whether someone is part of a HA group (treatment), that has been instrumented by whether the student´s AU-score overpasses the delimiting cutoff. Roughly speaking, first and third-degree polynomials appear

[14] At the final rows of the tables we indicate the polynomials included in the estimation equation and whether they are or not significant. A point indicates that this particular polynomial order has not been included at the determined estimation polynomial

to be statistically significant. Likewise, the estimation through this method produces quite comparable results to those obtained through fuzzy RD. When the treatment effect obtained through local RD is negative and statistically significant, the same holds at the majority of the cases obtained through polynomials. We find a couple of exceptions for men where, although the estimation through the RD local setup is significant, the model using polynomials does not produce significant estimations. Nevertheless in almost all cases the negative sign of the treatment is a recurring pattern and more noticeable in the case of women, that has been proven to be our strongest and more reliable group of analysis. We acknowledge that there are tiny differences in the size of the treatment, what we can partly attribute to the fact that different bandwidths can be chosen when applying the fuzzy RD methodology.

Table 8. ***Economics-related fields*. Robustness checks with socio-economic controls and parametric polynomial regression**

INTERN. B

	RD local	Polynomials			Controls	RD local	Polynomials			Controls	RD local	Polynomials			Controls
	All					**Women**					**Men**				
Treatment	-0.982***	-0.310	-0.817***	-0.814***	-1.056***	-1.057**	-0.387	-0.826**	-0.842**	-1.019***	-1.040**	-0.142	-0.671	-0.651	-1.023**
	(0.308)	(0.233)	(0.297)	(0.296)	(0.380)	(0.515)	(0.316)	(0.398)	(0.400)	(0.383)	(0.462)	(0.336)	(0.448)	(0.436)	(0.448)
Jump in grade	-0.838***				-0.790***	-0.879**				-0.771**	-0.894**				-0.781*
(Numerator)	(0.265)				(0.288)	(0.431)				(0.382)	(0.400)				(0.441)
Jump in prob	0.853***				0.748***	0.831***				0.757***	0.859***				0.763***
(Denominator)	(0.018)				(0.041)	(0.029)				(0.052)	(0.025)				(0.052)
Observations	2050	2050	2050	2050	2050	990	990	990	990	990	1060	1060	1060	1060	1060
Polynomials															
Order 1		Signif	Signif	Signif	Father U		Signif	Signif	Signif	Father U		Signif	Signif	Signif	Father U
Order 2		.	No S.	.	Moth U		.	No S.	.	Moth U		.	No S.	.	Moth U
Order 3		.	Signif.	Signif			.	Signif	Signif			.	Signif.	Signif	

LAW & B .

Treatment	-0.226	0.550**	-0.081	0.016	-0.137	-0.369	0.268	-0.290	-0.219	-0.211	0.144	0.919**	0.259	0.375	0.164
	(0.338)	(0.256)	(0.354)	(0.328)	(0.442)	(0.641)	(0.343)	(0.452)	(0.437)	(0.493)	(0.438)	(0.376)	(0.557)	(0.487)	(0.523)
Jump in grade	-0.166				-0.093	-0.241				-0.169	0.117				0.109
(Numerator)	(0.247)				(0.299)	(0.420)				(0.394)	(0.355)				(0.449)
Jump in prob	0.731***				0.676***	0.654***				0.797***	0.809***				0.665***
(Denominator)	(0.036)				(0.048)	(0.070)				(0.032)	(0.033)				(0.072)
Observations	1860	1860	1860	1860	1860	891	891	891	891	819	969	969	969	969	969
Polynomials															
Order 1		Signif	Signif	Sign	Father U		Signif	Signif	Sign	Father U		Signif	Signif	Signif	Father U
Order 2		.	No S.	.	Moth U		.	No S.	.	Moth U		.	No S.	.	Moth U
Order 3		.	No S.	Sign			.	No S.	Sign			.	No S.	Signif	

Table 9. *Business-related fields*. Robustness checks with socio-economic controls and parametric polynomial regression

INTERN. B

	RD local	Polynomials			Controls	RD local	Polynomials			Controls	RD local	Polynomials			Controls
	All					**Women**					**Men**				
Treatment	-1.041***	-0.357*	-0.752***	-0.752***	-1.008***	-1.393***	-0.385	-0.838**	-0.824**	-1.356***	-0.225	-0.195	-0.526	-0.567	-0.524
	(0.243)	(0.209)	(0.265)	(0.265)	(0.308)	(0.389)	(0.279)	(0.348)	(0.350)	(0.332)	(0.544)	(0.302)	(0.404)	(0.393)	(0.404)
Jump in grade	-0.792***				-0.750***	-1.156***				-1.026***	-0.163				-0.443
(Numerator)	(0.223)				(0.233)	(0.329)				(0.300)	(0.394)				(0.342)
Jump in prob	0.760***				0.744***	0.830***				0.757***	0.723***				0.846***
(Denominator)	(0.036)				(0.041)	(0.029)				(0.052)	(0.073)				(0.027)
Observations	1,990	1,990	1,990	1990	1990	957	957	957	957	957	1,033	1,033	1,033	1033	1033
Polynomials															
Order 1		Signif	Signinf.	Signinf.	Father U		Signif	Signif	Signif	Father U		Signif	Signif	Signif	Father U
Order 2		.	No S.	.	Moth U		.	No S.	.	Moth U		.	No S.	.	Moth U
Order 3		.	Signif.	Signinf.			.	Signif.	Signif			.	Signif.	No S.	

LAW & B .

Treatment	-0.845***	-0.261	-0.991***	-0.127	-1.091***	-1.280***	-0.372	-1.059***	-0.951**	-1.257***	-0.973**	-0.025	-0.809	-0.625	-0.803*
	(0.260)	(0.234)	(0.322)	(0.330)	(0.307)	(0.493)	(0.310)	(0.404)	(0.392)	(0.457)	(0.396)	(0.344)	(0.509)	(0.448)	(0.466)
Jump in grade	-0.683***				-0.726***	-0.834**				-0.843***	-0.659**				-0.534
(Numerator)	(0.212)				(0.248)	(0.335)				(0.319)	(0.333)				(0.379)
Jump in prob	0.808***				0.666***	0.651***				0.670***	0.678***				0.665***
(Denominator)	(0.021)				(0.049)	(0.073)				(0.068)	(0.062)				(0.071)
Observations	1801	1801	1801	1801	1801	858	858	858	858	858	943	943	943	943	943
Polynomials															
Order 1		Signif	Signinf.	Signinf.	Father U		Signif	Signinf.	Signinf.	Father U		Signif	Signinf.	Signinf.	Father U
Order 2		.	No S.	.	Moth U		.	No S.	.	Moth U		.	No S.	.	Moth U
Order 3		.	No S.	Signinf.			.	No S.	Signinf.			.	No S.	Signinf.	

Table 10. ***Mathematics-related fields*. Robustness checks with socio-economic controls and parametric polynomial regression**

INTERN. B

	RD local	Polynomials			Controls	RD local	Polynomials			Controls	RD local	Polynomials			Controls
	All					**Women**					**Men**				
Treatment	-1.527***	-0.361	-0.954***	-0.954***	-1.532***	-1.978***	-0.604*	-1.164***	-1.166***	-1.880***	-0.911	0.050	-0.547	-0.550	-0.960*
	(0.503)	(0.256)	(0.326)	(0.325)	(0.331)	(0.696)	(0.345)	(0.434)	(0.436)	(0.430)	(0.594)	(0.368)	(0.488)	(0.475)	(0.505)
Jump in grade	-1.121***				-1.162***	-1.421***				-1.423***	-0.772				-0.730
(Numerator)	(0.380)				(0.347)	(0.496)				(0.433)	(0.510)				(0.532)
Jump in prob	0.734***				0.759***	0.718***				0.757***	0.847***				0.760***
(Denominator)	(0.045)				(0.037)	(0.067)				(0.052)	(0.027)				(0.053)
Observations	2,045	2,045	2,045	2,045	2,045	988	988	988	988	988	1,057	1,057	1,057	1,057	1,057
Polynomials															
Order 1		Signif	Signif	Signif	Father U		Signif	Sign	Signif	Father U		Signif	Sign	Signif	Father U
Order 2		.	No S.	.	Moth U		.	No S.	.	Moth U		.	No S.	.	Moth U
Order 3		.	Signif.	Signif			.	Signif.	Signif			.	Signif.	Signif	

LAW & B .

	RD local	Polynomials			Controls	RD local	Polynomials			Controls	RD local	Polynomials			Controls
Treatment	-0.070	0.928***	-0.120	0.135	-0.391	-1.195**	-0.604*	-1.164***	-1.166***	-1.074**	0.764	0.050	-0.547	-0.550	0.647
	(0.361)	(0.281)	(0.386)	(0.359)	(0.432)	(0.579)	(0.345)	(0.434)	(0.436)	(0.526)	(0.510)	(0.368)	(0.488)	(0.475)	(0.595)
Jump in grade	-0.052				-0.309	-0.941**				-0.857**	0.618				0.434
(Numerator)	(0.268)				(0.342)	(0.455)				(0.418)	(0.406)				(0.534)
Jump in prob	0.740***				0.789***	0.788***				0.797***	0.808***				0.671***
(Denominator)	(0.034)				(0.024)	(0.034)				(0.032)	(0.032)				(0.071)
Observations	1,857	1,857	1,857	1,857	1,857	889	889	889	889	889	968	968	968	968	968
Polynomials															
Order 1		Signif	Signif	Signif	Father U		Signif	Signif	Signif	Father U		Signif	Signif	Signif	Father U
Order 2		.	Signif	.	Moth U		.	Signif	.	Moth U		.	Signif	.	Moth U
Order 3		.	No S.	Signif			.	No S.	Signif			.	No S.	Signif	

7. Conclusions

In this paper, we study peer effects at college level, by using a fuzzy RD approach and emphasizing the gender differences existing in the results obtained. We focus on different degrees at the UV (Spain) to compare between SG and HA groups. More precisely, we compare groups where peers must have better qualifications to enroll and tend to be more competitive, *Intern. B* and *Law and B.,* with standard *Business A*. groups. With the purpose of analyzing the impact of belonging to a HA group on student`s achievement, we choose as outcome variable student`s grades in common subjects of the first and second years of college.

In terms of methodology, we implement the relatively new RD approach. The method has been previously used by a few researchers in the study of peer effects, as being perfectly adequate for empirical studies on this topic provided a quantitative threshold determining treatment is available. However, this is the first contribution to the literature that uses administrative data at college level analyzing gender differences.

Was our purpose to analyze the impact of peers on HA students, we acknowledge that students who have always outperform and still have good results in college will for sure obtain advantages from being part of these prestigious groups. These benefits include creating interesting relationships with their colleagues, understanding the corporate environment by working in teams, and appreciating the importance of hard-work. However, we aimed to specifically assess whether the achievement of non-brilliant students (students at the threshold) enrolled in HA groups that may be comparable in terms of academic achievement with top students in SG improves as a result of having extremely qualified classmates. We find peers effect to have more impact in terms of *competitive pressure* as compared with *learning externalities* and *requirement standards*.

Whereas previous literature focusing on prior levels of education has documented that for comparable students the influence of HA classmates is positive, we obtain that peer effects at college level are negative. Hence, peers are not affecting students through *learning externalities* but through other sources. Moreover, we find disparities by gender and between *Intern. B* (a more competitive group) and *Law and B*. students that would not be expected if requirement standards were the main or unique source of peer effects. As the estimated effect for women is much higher than the estimation for men, and women tend to underperform in competitive environments, we can conclude that at college level peer effects affect in terms of competitive pressure. Henceforth, being surrounded by HA peers may be negative for students at the threshold.

We consider our study an important contribution to the literature analyzing the way people react in competitive environments. Numerous examples are found at company level but this research on the topic at college level seems to us necessary for the design of the proper education and gender equality policies. We acknowledge that we only have data on specific groups in Valencia, so in future research we may gather data from other universities in Spain to prove whether we can extend our conclusions. We also encourage other researchers to prove in other contexts the validity of our results. Finally, we find as well interesting to study whether women react the same way when being surrounded by men or by other women.

8. References

Angrist, J. D., and Lavy, V. (1999). "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement". *Quarterly Journal of Economics*, 114 (2), 533–575.

Angrist, J. D., and Pischke J., (2009). "Mostly Harmless Econometrics", Princeton University Press.

Beneito, P., Boscá, J. E., Ferri J. , and García, M. (2018). "*Women across Subfields in Economics: Relative Performance and Beliefs*". Working Papers 2018-06, FEDEA.

Berk, R. A, and Jan de Leeuw. (1999). "An Evaluation of California's Inmate Classification System Using a Generalized Regression Discontinuity Design". *Journal of the American Statistical Association*, 94 (448), 1045–1052.

Brollo, F., Nannicini, R. P., and Tabellini, G. (2013). "The Political Resource Curse". *American Economic Review*, 103, (5), 1759–96.

Bui, S. A., Craig, S. G., and Imberman, S. A. (2014). "Is gifted education a bright idea? Assessing the impact of gifted and talented programs on students". *American Economic Journal: Economic Policy*, 6(3), 30-62.

Canton, E., and Blom, A. (2004). "*Can Student Loans Improve Accessibility to Higher Education and Student Performance? An Impact Study of the Case of SOFES, Mexico*". World Bank Working Paper 3425.

Carpenter, C., and Dobkin, C. (2009). "The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age". *American Economic Journal: Applied Economics, American Economic Association*, 1(1), 164-82.

Cattaneo, M. D., Titiunik, R., Vazquez-Bare, G., & Keele, L. (2016). Interpreting regression discontinuity designs with multiple cutoffs. *The Journal of Politics*, *78*(4), 1229-1248.

Chay, K. Y., McEwan, P. J., and Urquiola, M. (2005). "The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools". *American Economic Review*, 95(4), 1237–1258.

Chen, M. K., and Shapiro, J. M. (2004). "*Does Prison Harden Inmates?: A Discontinuity-based Approach*". Cowles Foundation for Research in Economics.

Cornelissen, T., Dustmann, C., and Schönberg, U. (2017). "Peer effects in the workplace". *American Economic Review*. 107(2), 425-56.

Ding, W., and F Lehrer, S. (2007). "Do Peers Affect Student Achievement in China's Secondary Schools?". *The Review of Economics and Statistics*, MIT Press, 89(2), 300-312, 02.

Dobkin, C., and Ferreira, F. (2010). "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?". *Economics of Education Review*, 29 (1), 40–54.

Duflo, E., Dupas, P., and Kremer, M. (2011). "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya". *American Economic Review*, 101(5), 1739-74.

Gill, D., Prowse, V., (2014). "*Gender differences and dynamics in competition: the role of luck*". Quantitative Economics, 5, 351–376.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). "Performance in Competitive Environments: Gender Differences". *The Quarterly Journal of Economics*, 118 (3), 1049–1074.

Hahn, J, Todd, P., and van der Klaauw, W. (2001). "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design". *Econometrica*, 69(1), 201–09

Hoxby, C. M. (2000). "The Effects of Class Size on Student Achievement: New Evidence from Population Variation". *Quarterly Journal of Economics*, 115 (4), 1239–1285.

Imbens, G. W., and Kalyanaraman, K. (2012). "Optimal bandwidth choice for the regression discontinuity estimator". *Review of Economic Studies*, 79:, 933–959.

Iriberri, N., and Rey-Biel, P. (2017). "Stereotypes are Only a Threat when Beliefs are Reinforced: On the Sensitivity of Gender Differences in Performance under Competition to Information Provision". *Journal of Economic Behavior and Organization*, 135, 99– 111.

Iriberri, N. and Rey-Biel, P., (2018). "Competitive pressure widens the gender gap in performance: Evidence from a two-stage competition in mathematics". *The Economic Journal*. https://doi.org/10.1111/ecoj.12617

Klašnja, M. (2015). "Corruption and the Incumbency Disadvantage: Theory and Evidence". *Journal of Politics*, 77 (4), 928–942.

Klašnja, M., & Titiunik, R. (2017). "The incumbency curse: Weak parties, term limits, and unfulfilled accountability". *American Political Science Review*, *111*(1), 129-148.

Lee, D. S. (2008). "Randomized Experiments from Non-random Selection in U.S. House Elections". *Journal of Econometrics*, 142(2), 675–97

Lemieux, T, and Milligan, K. (2008). "Incentive Effects of Social Assistance: A Regression Dis- continuity Approach". *Journal of Econometrics*, 142(2), 807–28.

Niederle, M., Vesterlund, L., (2007). "Do women shy away from competition? Do men compete too much?" . *The Quarterly Journal of Economics*, 122, 1067–1101.

Porter, J. (2003). Estimation in the regression discontinuity model. *Unpublished Manuscript, Department of Economics, University of Wisconsin at Madison*, 5-19.

Sacerdote, B. (2001). "Peer Effects with Random Assignment: Results for Dartmouth Roommates". *Quarterly Journal of Economics*, 116(2), 681–704.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). "Stereotype threat and women's math performance". *Journal of Experimental Social Psychology*, *35*(1), 4-28.

Steele, C. M., (1997) "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance". *American Psychologist,* LII, 613–629.

Steele, Claude M., and Aronson, J. (1995) "Stereotype Vulnerability and the Intellectual Test Performance of African Americans". *Journal of Personality and Social Psychology,* LXIX, 797–811.

Urquiola, M. (2006). "Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia". *Review of Economics and Statistics*, 88(1), 171–177.

Urquiola, M., and Verhoogen, E. (2009). “Class-size Caps, Sorting, and the Regression-Discontinuity Design”. *American Economic Review*, 99 (1), 179–215.

Van der Klaauw, W. (2002). “Estimating the effect of financial aid offers on college enrollment: A regression–discontinuity approach”. *International Economic Review*, 43(4), 1249-1287.

Vardardottir, A. (2013). "Peer effects and academic achievement: a regression discontinuity approach". *Economics of Education Review*, Elsevier, 36(C), 108-121.