

Boys will be Boys: A Field Experiment on Relative Performance Feedback Information

Ghazala Azmat Manuel Bagues Antonio Cabrales Nagore Iriberry

Queen Mary University Aalto University University College London Basque Country University

February 2014

Motivation

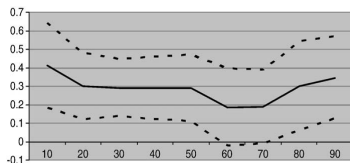
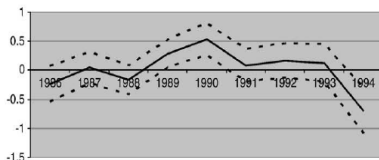
- The “true” motivation:
 - I was frustrated by my inability to persuade business leaders and policymakers to introduce scientific methods to evaluate their policies.
 - I told this guy about my frustration:



- His reaction was:
 - **Antonio, how many policies have you evaluated in that way at Carlos III? You have to lead by example!**

Motivation

- Inspired by the results of a natural experiment: Azmat, Iriberry, Jo. of Pub. Econ, 2010, 94, 435-452.
- The students of a high school in 1990/91 were informed besides their own grades as usual, about the average grade in their cohort and the average in their class.
- Class grades and university entrance examination improve, for the whole distribution.

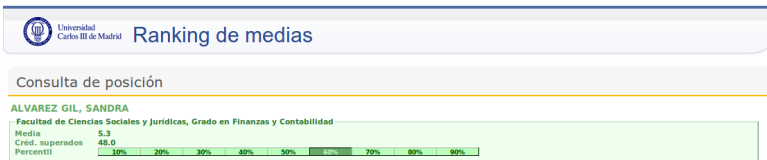


Our study

- The aim is to see if those results have external validity at the college level and in an environment with more control.
- We are also interested in the potential heterogeneity of the effect. 25 groups (1,000 students) belonging to the 2009 cohort are randomly assigned to treatment and control groups.
- Treatment started to be implemented at the beginning of the second year (October 2010).

Our study

- The treated students receive a message from a corporate account saying:
 - This e-mail is part of a pilot project of academic management. If you want to see your average grade, and your relative position in terms of average grade among the students that started the degree the same year you did, you can do it by clicking here:
<http://www.eco.uc3m.es/~acabrales/proyectomediarank.html>
 - And after the login in the Mediarank application they see:



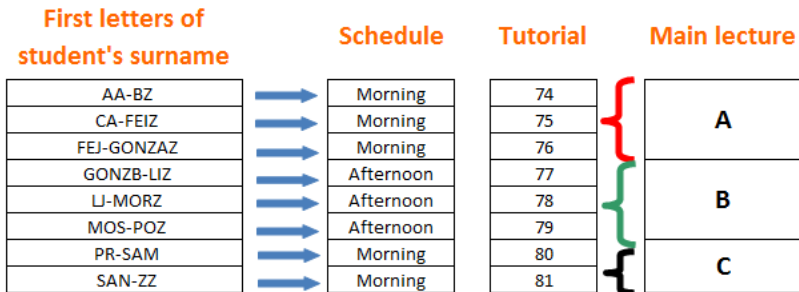
Universidad Carlos III de Madrid

- The university offers several degrees in three different campuses
- Some degrees offer a Spanish track and an English track
- Since 2008, Carlos III follows the Spanish version of the "Bolonia" system
 - Four year degrees
 - Two 14-weeks terms
 - Each week students attend a **lecture** and a **tutorial**
 - Subjects in the first and in the second year are **compulsory**. The third and the fourth year most courses are optional
 - The final **grade** is a weighted average of the grade obtained in a common final exam (at least 60%), midterm evaluations (about 20%) and group presentations/assignments (up to 20%)

Assignment of students

- Within each year, degree, campus and language, students are typically assigned to different tutorial and lecture groups based on their surname alphabetical order.

Figure : Example for 2010, first year, Business Administration, Getafe, Spanish



Data

- **Academic years:** 2008/2009 through 2012/2013
- **Five Degrees:** Business, Economics, Finance, Law, and joint degree in Business and Law.
- **Two Campuses:** Southern Madrid (Getafe) and Northern Madrid (Colmenarejo)
- **Two Languages:** Business, Economics and Finance also offer an English track.
- We consider only **1st and 2nd year courses** (for class size study).

Data

- Student level (N=4,960)
 - Gender, entry grade, entry mode, geographical origin
- Tutorial groups (N=2,550)
 - Identity of the professor, schedule, class size
 - Teaching evaluations are available for years 2010/2011 and 2011/2012
- Lecture groups (N=1,110)
 - Schedule, class size
 - Identity of the professor and teaching evaluations are **not available**
- Student x subject level (N=68,372)
 - Final grade
 - Disaggregated information on grades (final exam, midterms, term paper) is available for year 2010/2011

Individual characteristics

	Mean	SD	Min	Max
Female	0.54	0.5	0	1
Foreign	0.04	0.2	0	1
Surname order	0.5	0.29	0	1
Entry grade	7.19	1.02	5	9.94
Vocational training	0.04	0.2	0	1
High School	0.95	0.21	0	1
Madrid Center	0.32	0.47	0	1
Other Region	0.29	0.45	0	1
Madrid South	0.19	0.39	0	1
Madrid West	0.13	0.34	0	1

Notes: The table includes information on 4,960 students, except for information on citizenship and surname order which is only available since 2009 (3,238 observations). The variable *Surname order* takes values between 0 and 1, such that value 0 corresponds to the individual who ranks first in terms of surname alphabetical order.

Course level characteristics

Tutorial	Mean	SD	Min	Max
Class size	27.61	8	6	64
Weekday	3.15	1	1	5
Time	3.62	2	1	7
Main lecture				
Class size	63.41	24	7	152
Weekday	2.64	1	1	5
Time	3.59	2	1	7

Notes: The table includes information from 2,550 tutorial groups and 1,111 lecture groups. The sample includes all first and second year courses with at least four credits. Weekday equal to one corresponds to Monday. Time is codified as follows: 1=9:00, 2=10:45, 3=12:30, 4=14:30, 5=16:15, 6=18:00, 7=19:45.

Outcome variables

	N	Mean	SD	Min	Max
Academic Performance					
Took exam	68372	0.89	0.31	0	1
Passed exam	68372	0.67	0.47	0	1
Grade	68372	5.35	1.92	0	10
Teaching Evaluations					
Response rate	3301	0.39	0.22	0.03	1
Satisfaction	34787	3.76	0.82	1	5
Hours of study	34787	2.78	0.45	1	4.5
Grading is adequate	34787	3.41	0.66	1	5
Professor is helpful	6274	3.51	0.8	1.44	5
Ability acquired	6274	3.48	0.64	1.38	4.67

Notes: Teaching evaluations are only available for a subset of observations. Variables *Satisfaction*, *Grading is adequate*, *Professor is helpful* and *Ability acquired* are codified between 1 and 5, where 1 corresponds to ‘not at all’. *Hours of study* takes value ‘1’ if the individual studied less than an hour per week; ‘2’, between one and four hours; ‘3’, four to seven hours; ‘4’, seven to ten hours and ‘5’ more than ten hours.

Results

Table : Who checks the ranking?

	num_checks	checked
female	-0.573	0.036
	-0.461	-0.047
Average	1.205***	0.127***
	-0.443	-0.045
Credits Passed	-0.028*	-0.001
	-0.017	-0.002
Percentile	0.007	0
	-0.02	-0.002
Constant	-3.156*	-0.035
	(1. 670)	-0.169
Adj. R-sq	0.114	0.176
N	344	344

Results

Table : Is treatment really randomized (on observables)?

	female	grade access	average 2010	credits 2010
treatment	0.021 -0.035	-0.005 -0.061	0.022 -0.091	-0.24 (1. 152)
Adj. R Sq	0.018	0.313	0.122	0.063
N	938	930	938	938

	female	grade access	average 2010	credits 2010
treat*check	0.053 -0.038	0.072 -0.066	0.419*** -0.098	2.130* (1. 244)
Adj. R Sq	0.02	0.314	0.138	0.065
N	938	930	938	938

Results

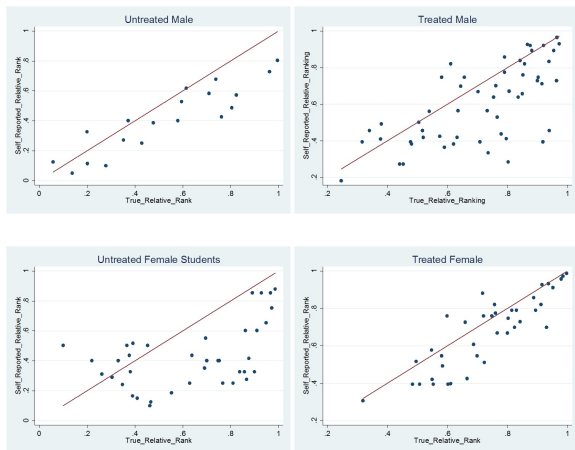
Table : Does information provided actually affect what students know?

VARIABLES	absolute_error	absolute_error
treatment	-0.0505** (0.0246)	-0.0526** (0.0256)
true_relative_rank		0.0466 (0.0582)
Constant	0.143*** (0.0187)	0.111*** (0.0334)
Observations	93	93
R-squared	0.085	0.091
Adjusted R-squared	0.0546	0.0497

Survey conducted in July and September 2013, after treatment was concluded, with Economics and Business students presenting (compulsory) degree thesis.

Results

But how was information related with reality for treated and control (by gender, which as you will see is important for results)?:



Results

Table : Outcomes end of 3rd year

	Ac. Year	Ac. Year	Grade Av.	Grade Av.	Percentile	Percentile	N. Credits	N. Credits
treatment	0.053* -0.03	0.052* -0.029	-0.005 -0.086	-0.022 -0.044	0.453 (1. 934)	0.069 (1. 129)	0.963 (3. 048)	0.911 (2. 010)
Average past		0.067*** -0.012		0.809*** -0.018		17.369*** -0.462		14.902*** -0.822
Credits past		0.003*** -0.001		0.001 -0.001		-0.028 -0.037		1.135*** -0.065
Constant	2010.726*** -0.018	2010.139*** -0.069	6.158*** -0.051	1.239*** -0.103	58.981*** (1. 136)	-44.108*** (2. 658)	159.224*** (1. 790)	3.317 (4. 732)
Adj. R Sq	0.039	0.111	0.076	0.76	-0.005	0.658	0.092	0.605
N	938	938	938	938	938	938	938	938

Similarly insignificant results when doing it for separate samples by gender, or by degree of studies, or by upper/lower half of GPA distribution at end of semester 2.

Results

Table : Females interacted with GPA

VARIABLES	gpa3	gpa4	gpa5	gpa6	gpa7	gpa8	gpa3	gpa4	gpa5	gpa6	gpa7	gpa8
gpa2	0.960*** (0.0237)	0.860*** (0.0350)	0.822*** (0.0393)	0.822*** (0.0389)	0.797*** (0.0428)	0.805*** (0.0424)	0.761*** (0.133)	0.738*** (0.125)	0.757*** (0.158)	0.749*** (0.173)	0.718*** (0.180)	0.742*** (0.185)
treatment	0.527 (0.480)	0.391 (0.471)	0.475 (0.561)	0.565 (0.603)	0.453 (0.632)	0.510 (0.643)	0.413 (0.391)	0.272 (0.440)	0.423 (0.526)	0.523 (0.553)	0.411 (0.583)	0.470 (0.596)
gpa2.treatment	-0.0901 (0.0746)	-0.0651 (0.0725)	-0.0746 (0.0862)	-0.0884 (0.0930)	-0.0690 (0.0972)	-0.0760 (0.0989)	-0.0735 (0.0608)	-0.0485 (0.0675)	-0.0673 (0.0805)	-0.0822 (0.0851)	-0.0628 (0.0894)	-0.0702 (0.0914)
percentile2							0.902* (0.528)	0.653 (0.471)	0.324 (0.599)	0.332 (0.670)	0.352 (0.690)	0.297 (0.714)
credits2							0.000222 (0.00194)	-0.00372* (0.00190)	-0.00111 (0.00249)	1.42e-05 (0.00277)	0.000323 (0.00300)	-0.000286 (0.00311)
Constant	0.226 (0.156)	0.886*** (0.232)	1.184*** (0.260)	1.196*** (0.257)	1.461*** (0.283)	1.463*** (0.282)	0.936** (0.463)	1.493*** (0.502)	1.468** (0.607)	1.459** (0.641)	1.732** (0.678)	1.705** (0.690)
Observations	528	528	528	528	528	528	528	528	528	528	528	528
R-squared	0.908	0.833	0.788	0.759	0.726	0.724	0.913	0.838	0.789	0.759	0.727	0.724
Adjusted R-squared	0.906	0.830	0.784	0.755	0.721	0.719	0.911	0.834	0.784	0.754	0.721	0.718

Results

Table : Male students interacted with GPA

gpa3	gpa4	gpa5	gpa6	gpa7	gpa8	gpa3	gpa4	gpa5	gpa6	gpa7	gpa8
0.858*** (0.0649)	0.777*** (0.0639)	0.749*** (0.0629)	0.766*** (0.0645)	0.757*** (0.0638)	0.772*** (0.0578)	0.581*** (0.210)	0.551*** (0.202)	0.585*** (0.206)	0.593*** (0.212)	0.628*** (0.216)	0.679*** (0.193)
-0.685* (0.364)	-1.063*** (0.395)	-1.307*** (0.432)	-1.210*** (0.438)	-1.306*** (0.447)	-1.305*** (0.434)	-0.630** (0.311)	-1.015*** (0.362)	-1.277*** (0.405)	-1.183*** (0.410)	-1.287*** (0.424)	-1.291*** (0.417)
0.101* (0.0575)	0.160** (0.0621)	0.203*** (0.0676)	0.189*** (0.0687)	0.203*** (0.0697)	0.206*** (0.0676)	0.0946* (0.0494)	0.155*** (0.0573)	0.200*** (0.0637)	0.185*** (0.0645)	0.201*** (0.0663)	0.204*** (0.0652)
						1.449* (0.788)	1.204 (0.763)	0.841 (0.784)	0.854 (0.808)	0.623 (0.822)	0.460 (0.741)
						0.000871 (0.00240)	6.96e-05 (0.00242)	0.00114 (0.00252)	0.00222 (0.00262)	0.00203 (0.00274)	0.00126 (0.00263)
0.839** (0.398)	1.368*** (0.393)	1.620*** (0.388)	1.539*** (0.398)	1.695*** (0.394)	1.643*** (0.358)	1.660** (0.727)	2.064*** (0.723)	2.082*** (0.731)	1.985*** (0.750)	2.012*** (0.761)	1.881*** (0.685)
449	449	449	449	449	449	449	449	449	449	449	449
0.873	0.801	0.764	0.757	0.736	0.733	0.885	0.810	0.768	0.762	0.739	0.734
0.871	0.797	0.759	0.752	0.731	0.727	0.882	0.805	0.762	0.756	0.732	0.728

The effect of class size on grades

Class size tutorial	0.0022* (0.0012)	0.0013 (0.0011)	0.0009 (0.0013)	0.0010 (0.0013)	0.0004 (0.0014)
Individual characteristics	No	Yes	Yes	Yes	Yes
Lecture fixed effects	No	No	Yes	Yes	Yes
Tutorial schedule	No	No	No	Yes	Yes
Tutorial professor fixed effects	No	No	No	No	Yes

Notes: OLS regressions. The table provides information from 68,372 individual x course observations. The left hand side variable is the standardized final grade obtained by the student in the course, rescaled to have a mean of zero and a standard deviation of one. All regressions include cohort fixed effects. Individual characteristics include entry grade, mode of entry, geographical origin and gender. Tutorial schedule includes weekday and time. Standard errors are clustered at the individual level.

The effect of class size on participation in teaching evaluations

Class size tutorial	-0.0015 (0.0030)	-0.0015 (0.0030)	-0.0006 (0.0029)	-0.0009 (0.0029)	0.0003 (0.0032)
Individual characteristics	No	Yes	Yes	Yes	Yes
Lecture fixed effects	No	No	Yes	Yes	Yes
Tutorial schedule	No	No	No	Yes	Yes
Tutorial professor fixed effects	No	No	No	No	Yes

Notes: OLS regressions. The table provides information from 3,301 individual x course observations. The left hand side variable is the proportion of students within the class that filled in the teaching evaluation. All regressions include cohort fixed effects. Individual characteristics include entry grade, mode of entry, geographical origin and gender. Tutorial schedule includes weekday and time. Standard errors are clustered at the tutorial group level.

The effect of class size on students' satisfaction

Class size tutorial	-0.0093*** (0.0035)	-0.0092*** (0.0035)	-0.0085*** (0.0032)	-0.0081** (0.0032)	-0.0094*** (0.0028)
Individual characteristics	No	Yes	Yes	Yes	Yes
Lecture fixed effects	No	No	Yes	Yes	Yes
Tutorial schedule	No	No	No	Yes	Yes
Tutorial professor fixed effects	No	No	No	No	Yes

Notes: OLS regressions. The table provides information from 34,787 individual x course observations. The left hand side variable is the standardized final grade obtained by the student in the course, rescaled to have a mean of zero and a standard deviation of one. All regressions include cohort fixed effects. Individual characteristics include entry grade, mode of entry, geographical origin and gender. Tutorial schedule includes weekday and time. Standard errors are clustered at the tutorial group level.

The effect of class size on grading standards

Class size tutorial	-0.0045 (0.0034)	-0.0045 (0.0034)	-0.0054 (0.0035)	-0.0053 (0.0035)	-0.0053 (0.0039)
Individual characteristics	No	Yes	Yes	Yes	Yes
Lecture fixed effects	No	No	Yes	Yes	Yes
Tutorial schedule	No	No	No	Yes	Yes
Tutorial professor fixed effects	No	No	No	No	Yes

Notes: OLS regressions. The table provides information from 34,787 individual x course observations. The left hand side variable is *grading standards*, as perceived by students, rescaled to have a mean of zero and a standard deviation of one. All regressions include cohort fixed effects. Individual characteristics include entry grade, mode of entry, geographical origin and gender. Tutorial schedule includes weekday and time. Standard errors are clustered at the tutorial group level.

The effect of class size on students' effort

Class size tutorial	0.0057 (0.0036)	0.0057 (0.0036)	0.0034 (0.0040)	0.0034 (0.0040)	0.0016 (0.0042)
Individual characteristics	No	Yes	Yes	Yes	Yes
Lecture fixed effects	No	No	Yes	Yes	Yes
Tutorial schedule	No	No	No	Yes	Yes
Tutorial professor fixed effects	No	No	No	No	Yes

Notes: The left hand side variable is the self-reported number of hours devoted to study weekly, excluding lectures, rescaled to have a mean of zero and a standard deviation of one. The table provides information from 34,787 individual x course observations. All regressions include cohort fixed effects. Individual characteristics include entry grade, mode of entry, geographical origin and gender. Tutorial schedule includes weekday and time. Standard errors are clustered at the tutorial group level.

The effect of class size on professors' help

Class size tutorial	-0.0284*** (0.0069)	-0.0283*** (0.0068)	-0.0167** (0.0069)	-0.0167** (0.0069)	-0.0244*** (0.0059)
Individual characteristics	No	Yes	Yes	Yes	Yes
Lecture fixed effects	No	No	Yes	Yes	Yes
Tutorial schedule	No	No	No	Yes	Yes
Tutorial professor fixed effects	No	No	No	No	Yes

Notes: OLS regressions. The table provides information from 6,274 individual x course observations. The left hand side variable is the *helpfulness of the professor*, as perceived by students, rescaled to have a mean of zero and a standard deviation of one. All regressions include cohort fixed effects. Individual characteristics include entry grade, mode of entry, geographical origin and gender. Tutorial schedule includes weekday and time. Standard errors are clustered at the tutorial group level.

The effect of class size on students' ability

Class size tutorial	-0.0165*** (0.0059)	-0.0163*** (0.0059)	-0.0151*** (0.0058)	-0.0150*** (0.0057)	-0.0202*** (0.0052)
Individual characteristics	No	Yes	Yes	Yes	Yes
Lecture fixed effects	No	No	Yes	Yes	Yes
Tutorial schedule	No	No	No	Yes	Yes
Tutorial professor fixed effects	No	No	No	No	Yes

Notes: OLS regressions. The table provides information from 6,274 individual x course observations. The left hand side variable is the *ability acquired*, as perceived by students, rescaled to have a mean of zero and a standard deviation of one. All regressions include cohort fixed effects. Individual characteristics include entry grade, mode of entry, geographical origin and gender. Tutorial schedule includes weekday and time. Standard errors are clustered at the tutorial group level.

Discussion and Future Research

- The experiment finished at the end of last academic year.
- We do not have a good theory for the results.
- There are also results about other variables.
- Suggestions about further research will be quite welcome!

Thank you for your attention!