Multitask Agents and Incentives: the Case of Research and Teaching for University Professors

Marta De Philippis

Bank of Italy, CEP

November 29th, 2016

Bank of Italy, Human Capital Workshop

Motivation

- Most jobs involve different tasks
- Important to take it into account when providing incentives
- Theory predicts (Holmstrom and Milgrom, 1990):
  - behavioural distortions generated by incentives on one task only for multitask agents
  - stronger if no intrinsic motivation and if performance in other tasks difficult to measure
- No clean empirical evidence on the actual economic cost of these effects
  - Difficult to measure performance in many different tasks
  - High requirements in terms of data and exogenous variation

This Paper

- This paper looks at the case of university professors, who allocate time between many tasks (mostly, teaching and research)
- Strong incentives towards research: *"Publish or perish"*
- Case study: Bocconi University

Research questions

- What are the effects of research incentives on:
  1. (directly) on quantity and quality of publications?
  2. (indirectly) on teaching quality?
     $\rightarrow$ it depends on whether teaching and research efforts are substitute or complement in the professors' cost function
  3. on selection of faculty members in terms of teaching and research quality?
     $\rightarrow$ it depends on whether teaching and research skills are positively or negatively correlated
- Would the effect on teaching be smaller if the universities reward good researchers by reducing teaching load?

Bocconi is a suitable case study

- Bocconi is a private university, based in Milan
- Suitable institutional setting:
  1. Available measure of teaching quality in university setting:
     - standardized exams
     - randomization of students to teachers
  2. Exogenous variation:
     - sharp change in research incentives (monetary but especially in terms of promotions) announced in 2005
  3. Control group:
     - heterogeneity in professors' contracts
- I use a Diff-in-Diff strategy using as control groups:
  - external teachers (prof. in other uni, not academics)
  - Two alternative control groups (not today):
    1. for teaching: full prof (right before the reform)
    2. for research: professors in a comparable university (Bologna)

Findings: the Effects of Stronger Research Incentives

- More publications and working papers
- Worse teaching performance
  - Stronger effect for young faculty
  - Low ability students are hit the most
- Research and teaching efforts are substitute in the agents' cost function
- Selection of better researchers and therefore better teachers
  - Research and teaching skills are positively correlated
- Overall effect on teaching is ambiguous: positive selection effect countervails negative effect from effort substitution

Related papers and Contribution

Incentive literature:

- Few available empirical papers, usually on quantity-quality trade-off (Dumont et. al. 2008, Jacob 2005, Fryer at. al 2013)
- This paper:
  1. looks at two different activities (trade-off no obvious a priori)
  2. analyze selection effects as well (Lazear, 2000)
  3. disentangle effect of monetary incentives from career concerns

Education literature:

- On teachers' contracts (Figlio et. al. 2013, Bettinger et. al. 2010)
- On the effect of teaching incentives on research performance (Brikley and Zimmerman 2001, Payne at al. 2010)
- This paper:
  1. looks at research incentives (publish or perish)
  2. correlation between teaching and research skills

The Treatment and The Data

The new Bocconi policy on research

- In 2005 Bocconi boeard announced a new strategic plan. The objectives were:
  1. to improve research evaluation system for professors (independent committees, internationalization of criteria)
  2. to strengthen incentives towards research
     - changes in the hiring and promotion policies
     - introduction of monetary incentives
     - only for internal faculty members
  3. to create mechanisms to attract and keep best researchers.
- The focus switched explicitly towards research

Bocconi data

- Students-teachers matched panel dataset (2001-2011)
- Only undergraduate students (compulsory courses)
- Academic records of students:
  - demographic information
  - exam grades and dates, teaching classes, degree courses chosen
- Teachers:
  - courses/classes taught every year, numbers of hours taught
  - type of contract, age
  - yearly data on publications and working papers, collected from Web of Science and Google Scholar websites

Treatment and control group

- Treatment group=internal professors:
  - assistant and associate prof.: both career path and monetary incentives
  - full prof.: only monetary incentives
- Control group=external professors:
  - not subject to research incentives
  - so-called "non academic" (but expert in the subject) 57%
  - professors hired by other universities 43%
- in 2005 33% of considered teachers (in compulsory undergrad courses) are external

Descriptives: teachers

	Internal	External	Diff
Teachers' descriptives			
N teaching hours per class	38.91	33.91	5.47***
	(16.60)	(17.44)	(1.34)
Age	43.18	41.29	1.89**
	(9.45)	(7.80)	0.77)
% female	32.27	34.25	-0.20
	(0.47)	(0.47)	(0.045)
Teachers' Position			
% Assistant prof	50.04%		
% Associate prof	10.45%		
% Full prof	12.65%		
% Non academic		9.61%	
% Other univ prof		9.11%	
% Lecturers		7.76%	

Effect on Research Performance

Three measures of research outcomes

1. n of published papers from Web of Knowledge
2. (Not today) proxy of the index used by Bocconi to reward publications (n papers weighted by the importance of the journal/n coauthors)
   - based on Journal Quality List and on Impact Factor (examples: *Econometrica* A+=15 points, *Economic Journal* A=7 points, *Economic Letters* B=3 points etc)
3. (Not today) n of working papers from Google scholar

Diff in Diff for research

$$pub_{py} = \theta_y + \theta_p + \gamma_{res}(internal_p * post2006_y) + \gamma_4 Q_{py} + \eta_{py}$$

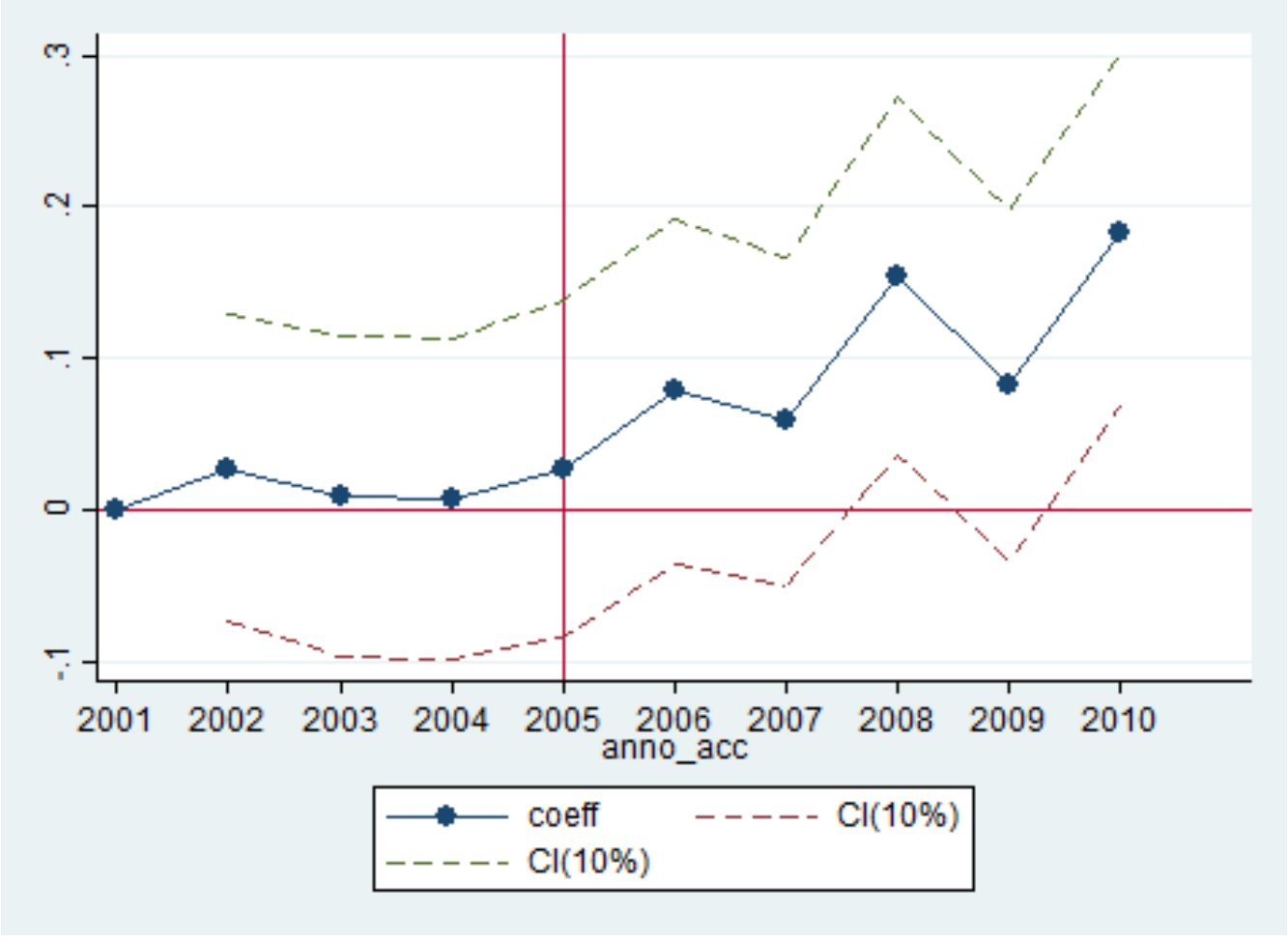
- where:
  - pub_{py} is the n of publications of professor p in year y
  - take care of switchers by using status in 2005. Robustness checks
  - θ_y are time fixed effects
  - θ_p are teacher fixed effects
  - Q_{py} are professor characteristics (age, age squared)
  - η_{py} is the error term

Summary statistics: Reserach

	Overall	Post 2006	Pre 2006
		N publications	
Internal 2005	0.539	0.680	0.302
sd	1.561		
External 2005	0.416	0.481	0.264
sd	1.318		

Graphical representation and parallel trends check

$$pub_{py} = \theta_y + \theta_p + \sum_{s=2002}^{2011} \gamma_s (internal_p * \theta_s) + \gamma_4 Q_{py} + \eta_{py}$$



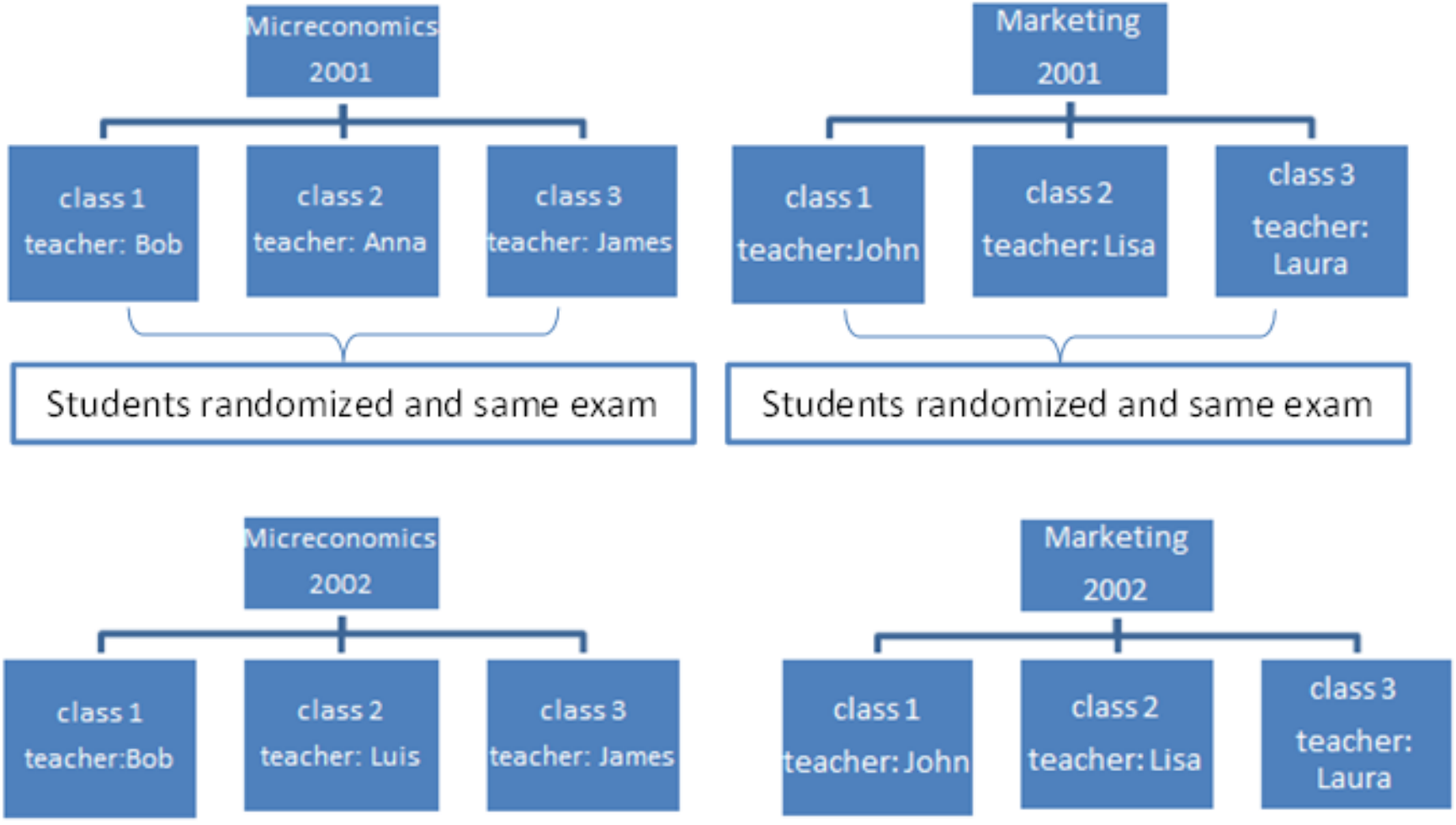
Research outcome improved

	[1]	[2]	[3]	[4]
	Dependent variable: Number Publications			
	n pub			
1=int*post06	0.206**		0.142**	
	(0.100)		(0.070)	
1=jun pr *post06		0.224*		0.157**
		(0.114)		(0.081)
1=full pr*post06		0.137		0.099
		(0.186)		(0.123)
N	5230	5230	5230	5230
Teacher fe	No	No	Yes	Yes

Overall n of published papers increased by 25% after the change in incentives

Effect on Teaching Performance

Teaching Performance Measure



Teaching Performance Measure: step 1

$$grade_{ipcy} = \beta_0 + \beta_1 HSgrade_{ipcy} + \beta_2 X_{ipcy} + \alpha_{pcy} + u_{ipcy}$$

- where:
  - $grade_{ipcy}$ is the grade taken by student i, with the teacher p in year t in the course c. Standardized by course-year
  - $HSgrade_{ipcy}$ is final high school grade
  - X_{ipcy} are individual characteristics (gender, age, whether Italian, whether from Milan, type of high school attended)
  - α_{pcy} is the parameter of interest (conditional mean, after controls)
  - u_{ipcy} is the error term

students' descriptives step1

Diff in Diff for teaching

$$\hat{\alpha}_{pcy} = \delta_p + \delta_{cy} + \gamma_{teach}(internal_p * post2006_y) + \gamma_2 Q_{py} + \epsilon_{pcy}$$

- where:
  - $internal_p$=1 if professor internal in 2005 (before announcement)
  - δ_p are teacher fixed effects
  - δ_{cy} are fixed effects for the interaction between time and courses (exam paper)
  - Q_{py} are time-varying professor characteristics (age)
  - ϵ_{pcy} is the error term
- similar results if regression directly at the student level and if just raw average ($\overline{grade_{pcy}}$), given randomization
- drop lecturer. Robustness check

Summary stat: teaching ($\hat{\alpha}_{pcy}$)

		$\hat{\alpha}_{pcy}$		
		Overall	Post 2006	Pre 2006
Internal 2005	*mean*	-0.020	0.146	-0.197
	sd	0.632		
External 2005	*mean*	0.074	0.239	-0.192
	sd	0.645		

Graphical representation and parallel trends check

$$\hat{\alpha}_{pcy} = \delta_p + \delta_{cy} + \sum_{s=2002}^{2011} \gamma_s(internal_p * \delta_s) + \gamma_2 Q_{py} + \gamma_3 Z_{pcy} + \epsilon_{pcy}$$

Results: teaching performance worsened

	Dependent variable: $\hat{\alpha}_{pcy}$			
	[1]	[2]	[3]	[4]
int*post06	-0.011		-0.037**	
	(0.012)		(0.018)	
jun pr*post06		-0.014		-0.042**
		(0.013)		(0.020)
full pr*post06		-0.001		-0.023
		(0.016)		(0.022)
N	3889	3889	3889	3889
Teachers fe	No	No	Yes	Yes
Year*course*degree pr fe	Yes	Yes	Yes	Yes

Control set: age, age squared. Se clustered by teacher.

Correlation research and teaching skills

Estimating research and teaching skills

- For research:

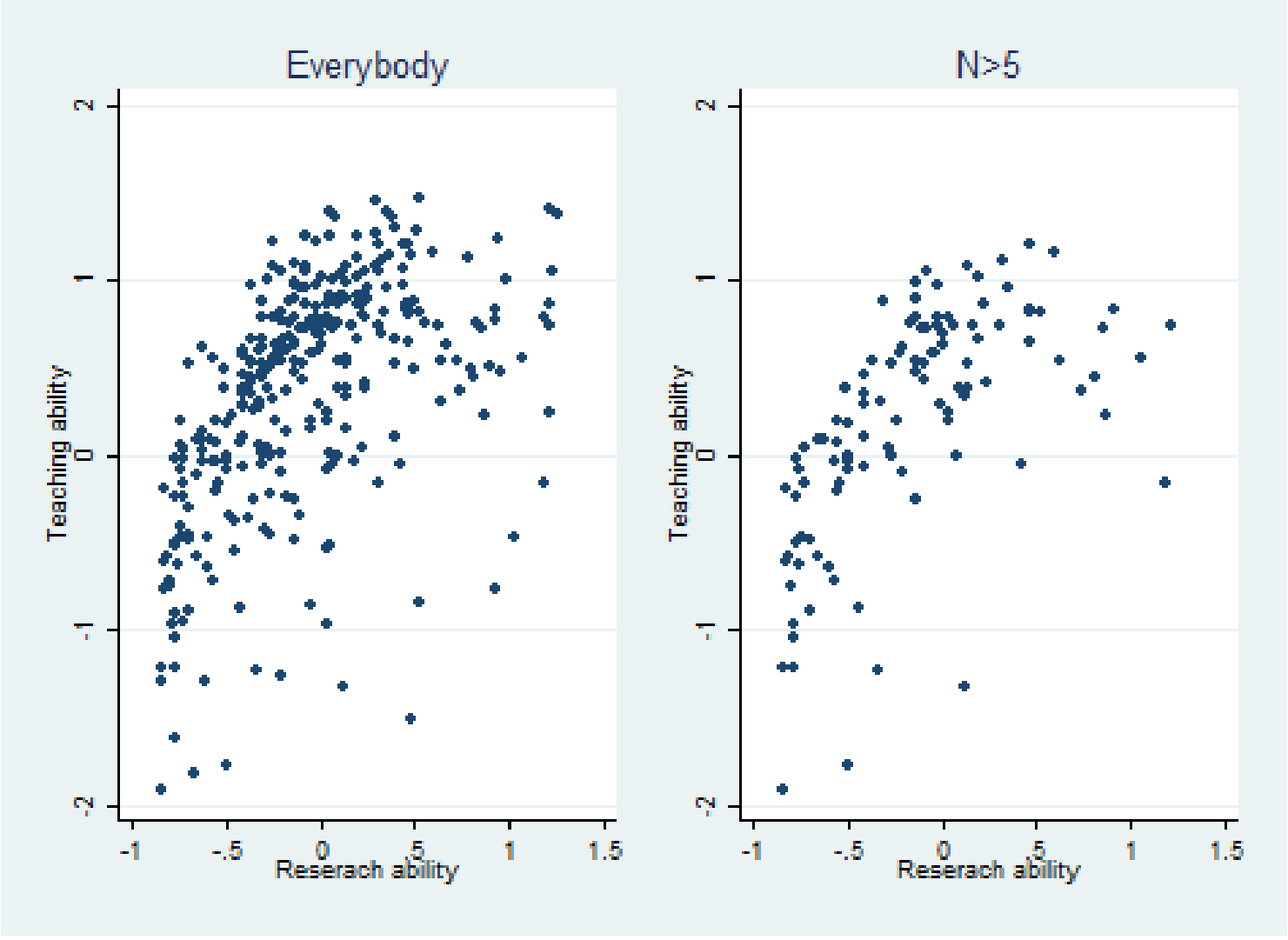
$$pub_{py} = \theta^r_p + \zeta_y + \zeta_2 Q_{py} + \eta_{py}$$

- For teaching:

$$\widehat{\alpha_{pcy}} = \theta^t_p + \delta_{cy} + \gamma_2 Q_{py} + \epsilon_{pcy}$$

- where:
  - Z_{pt} includes age, age squared and experience at Bocconi
  - δ^t_{cy} are course*year fe; δ^r_y are year fixed effects
  - θ^t_p and θ^r_p are professors fixed effects

Teaching and research skills are positively correlated



Regressions Heterogeneity effect by teachers' skills

Selection Effects

Regressions for those leaving/entering Bocconi

- How did the quality of researchers and teachers at Bocconi change after the policy?
- Only for internal (for external no exact record of year of entry)
- Separately for the years 2001–2005 and 2006–2011 (where $j = \{r, t\}$):

$$exit_p = \alpha_1^j \widehat{\theta_p^j} + \alpha_2^j X_p + \delta_e^j + \iota_p^j$$

$$entry_p = \beta_1^j \widehat{\theta_p^j} + \beta_2^j X_p + f^j(e_p) + \omega_p^j$$

- where:
  - $exit_p$ and $entry_p$ dummies if teacher p left/entered Bocconi
  - δ_e^j are year of entry fe; $f^j(e_p)$ polynomial year of entry
  - X_p= gender, coordinator
  - excluding those retired

Worst researchers are leaving Bocconi

Dep Variable:	1=exit pre 2006 [1]	1=exit post 2006 [2]	1=entry pre 2006 [3]	1=entry post 2006 [4]
research skills($\widehat{\theta_p^r}$)	-0.062	-0.109***	0.027	-0.039
	(0.087)	(0.038)	(0.049)	(0.053)
N	273	260	289	277
teaching skills ($\widehat{\theta_p^t}$)	0.123	-0.374***	0.113	-0.309**
	(0.119)	(0.136)	(0.089)	(0.141)
N	302	283	302	283

Excluding those exiting because retiring, omitted category=those staying. additional controls: dummies for year of entry, gender.

correlation pre-post fe

Robustness Checks

Other Robustness Checks and Extensions

- Alternative control groups(show research show teaching)
- Randomization(show)
- Robustness checks teaching (show)
- Grading and Teaching loads (show)
- Change in grade normalization (show)
- Non-compliance to class allocation (show)
- Student level regressions (show)
- Research Spillovers (show)

Conclusions

- Stronger research incentive =
  - Positive effect on research
  - Negative effect on teaching quality for undergrad students
  - Encouraging one more paper has an implicit cost of 0.3 standard deviation on teaching quality
  - Mostly generated by career concerns
  - Positive selection effects
- Research and teaching ability are positively correlated
- Acting on selection may potentially lead to large effects
- Room to minimize distortions by matching teachers with students who are less/no damaged

Thank you

Descriptives: students

Variable	Mean [1]	Std. Dev. [2]	Min. [3]	Max. [4]
1=female	0.469	0.499	0	1
year birth	1985	3.249	1954	1993
1=italian	0.973	0.163	0	1
1=from Milan	0.246	0.431	0	1
hs grade	0.899	0.103	0.6	1
exam grades	25.532	3.532	18	31
N		501189		

back

Results on Teaching: First step

Dependent variable: exam grade	
	All
	[1]
hs grade	-3.704***
	(0.225)
hs grade^2	4.159***
	(0.131)
1=female	-0.051***
	(0.003)
1=italian	0.142***
	(0.013)
1=from Milan	0.074***
	(0.003)
N	501132

back

Adding controls

	Dep. var= Teaching Fe			
	everybody	N>5	everybody	N>5
	[1]	[2]	[3]	[4]
Research Fe	0.672***	0.668***	0.583***	0.583***
	(0.046)	(0.068)	(0.043)	(0.065)
N	326	163	326	163
Controls	No	No	Yes	Yes

Additional controls: age at entry (linear and squared), gender.

back

Stronger effect for low ability researchers

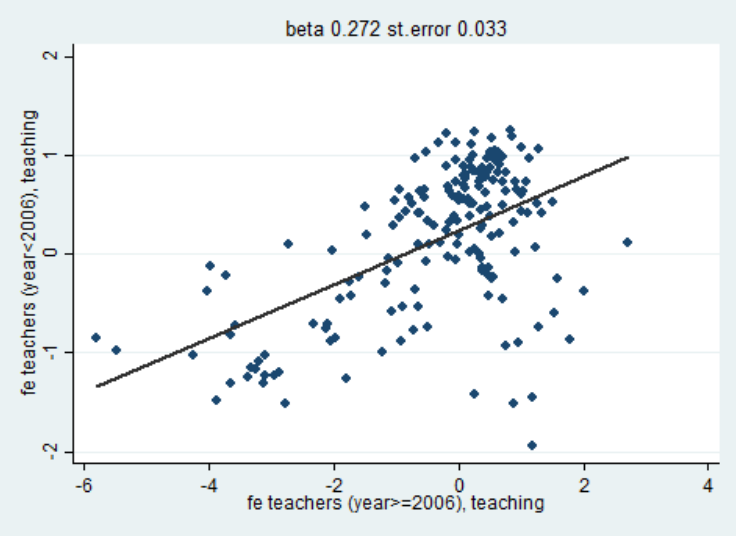
Dep. var	α_{pct} [1]	n pub [2]	n wp (google) [3]
int*post 06* ability q 1	-0.078***	0.138***	0.294**
	(0.029)	(0.041)	(0.120)
int*post 06* ability q 2	-0.020	0.162**	0.105
	(0.036)	(0.064)	(0.177)
int*post 06* ability q 3	-0.011	0.142	0.088
	(0.035)	(0.140)	(0.246)
N	3770	7005	6777

Control set: age, age squared, course*year and teacher fe. Se clustered by teacher.

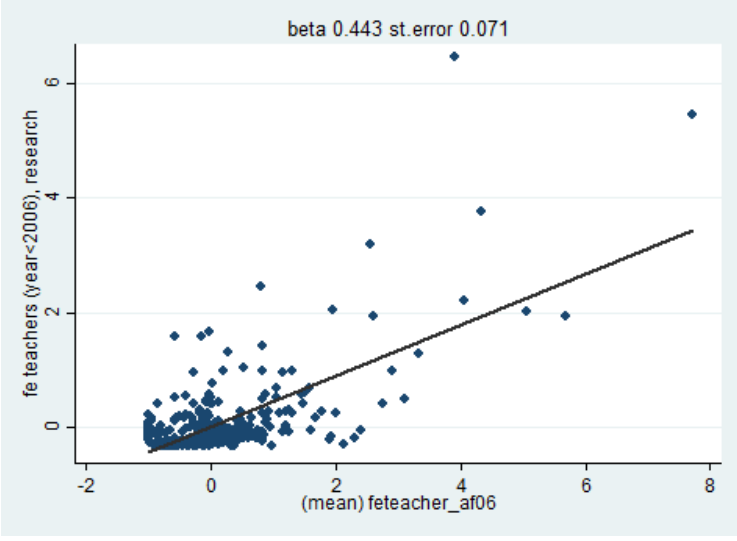
back

Fe estimated pre post 2006

Teaching ability



Research ability



back

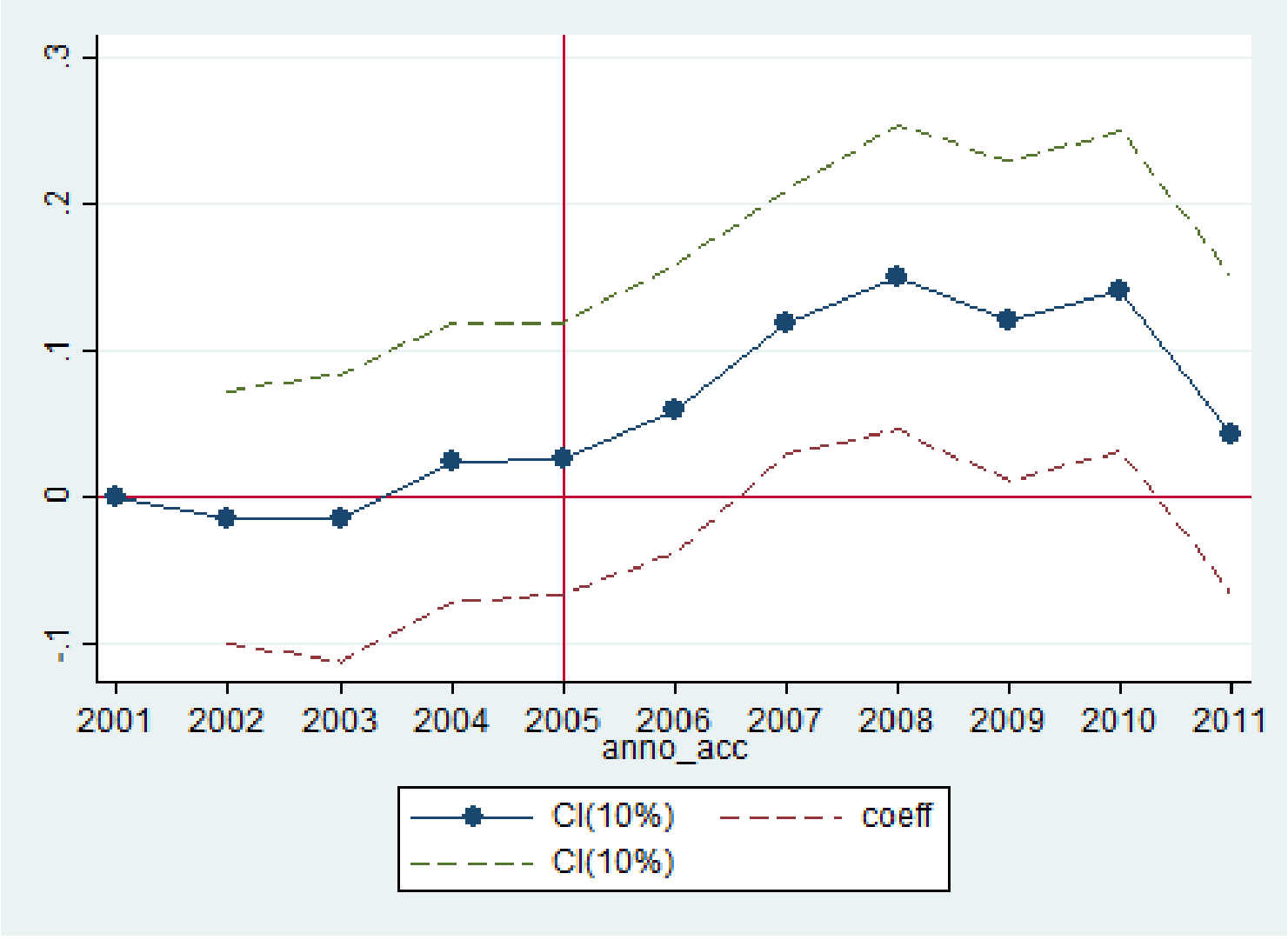
Alternative control groups for research

Dep. var	n pub				
Contr gr	External prof age groups		Bologna prof		
	< m age (43)	> m age (43)	All	Jun	Full
	[1]	[2]	[3]	[4]	[5]
internal*post06	0.170*	0.119			
	(0.096)	(0.124)			
bocconi*post06			0.162**		
			(0.064)		
jun bocc*post06				0.221***	
				(0.080)	
ord bocc*post06					0.051
					(0.107)
N	3119	2111	4497	3063	1434

Control set: age age squared. Se clustered by teacher.

descriptives Bologna prof

Parallel trends: Internal Bocconi-Bologna



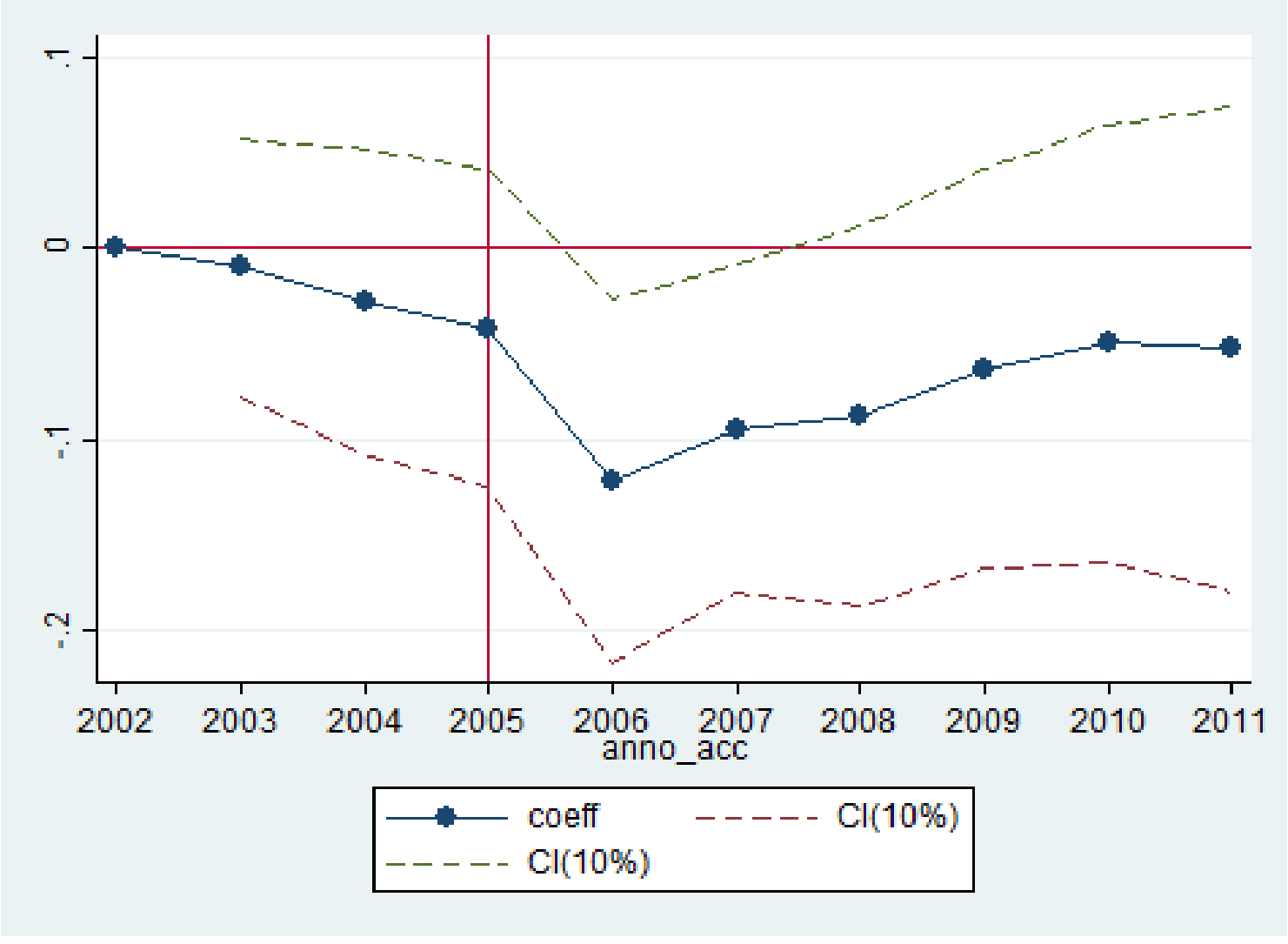
back

Alternative control groups for teaching

Dep Var: α_{pct}			
Contr gr:	External prof age groups		Prof just tenured
	$<$ m age (43)	$>$ m age (43)	
	[1]	[2]	[3]
internal*post06	-0.061*	-0.034	
	(0.032)	(0.029)	
no full pre05*post06			-0.042*
			(0.025)
N	1958	1931	2068
Teachers fe	Yes	Yes	Yes
Year*course*deg fe	Yes	Yes	Yes

Control set: age age squared. Se clustered by teacher.

Parallel trends: junior-full



back

Descriptives Bologna

	Bologna	Bocconi	diff
	Junior prof		
N pub	0.201	0.554	-0.354***
	(0.018)	(0.039)	()0.050)
	1221	1842	
	Senior prof		
N pub	0.280	0.495	-0.215***
	(0.030)	(0.046)	(0.053)
	792	642	

Bologna University is the most similar to Bocconi University in terms of dimension of the department and of quality of research as evaluated by the Italian Institute of University Research Evaluation (ANVUR) between 2004-2010. back to list

Randomization

	Av. final hs grade	Av. female	Av. from Mi	Sd final hs grade
internal	0.001	0.000	0.001	-0.002
	(0.001)	(0.000)	(0.003)	(0.002)
age	0.000	-0.000	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
female	0.001	0.000	0.000	-0.001
	(0.001)	(0.000)	(0.002)	(0.002)
coordinator	0.000	-0.001	0.003	-0.000
	(0.001)	(0.001)	(0.003)	(0.002)
N	3889	3889	3889	3889
course*year fe	Yes	Yes	Yes	Yes
F stat joint sign	0.75	0.95	0.39	1.58

back to list

Robustness Teaching

	no 09-10-11 [1]	also lecturers [2]	include switches [3]	weight by h. taught [4]
int05*post06	-0.037*	-0.034*		-0.035*
	(0.020)	(0.018)		(0.021)
lecturer*post06		-0.047		
		(0.042)		
int*post06			-0.027*	
			(0.016)	
N	2848	4201	3889	3889
Teachers fe	Yes	Yes	Yes	Yes

Robust standard errors clustered by teacher in parentheses. Additional controls: age and age squared of teachers, teacher experience in Bocconi class size. Column (1) excludes the years when teaching incentives were also in place; column (2) includes lecturers and specifies a different treatment effect for lecturers; column (3) includes switchers and teachers fixed effects; column (4) weights professors by number of teaching hours.

Back to list

Robustness Teaching 2

Dep var:	Grading α_{ptc}		1=course coordin	1=Num of taught h
	[1]	[2]	[3]	[4]
int*post 06	-0.045**	-0.042**	0.025	0.671
	(0.020)	(0.020)	(0.037)	(1.084)
int*post 06*obj	0.024			
	(0.047)			
int*post 06*math dep		0.017		
		(0.046)		
N	3889	3889	3889	2989
Teachers fe	Yes	Yes	Yes	Yes

Robust standard errors clustered by teacher in parentheses. Additional controls: age and age squared of teachers, dummies for teacher experience in Bocconi. A course is considered "objective" if the name includes the words *"math"*, *"stat"*, *"quantit"* . Math if the teacher belongs to the math and statistics departments. N of observations col 4 is at the teacher-year level (if a teacher teaches more than one courses n of teaching hours are summed)

Back to list

Robustness: suggested homogeneity?

	Av. av class grade	Sd av class grade
2001	26.008	0.436
2002	26.444	0.381
2003	26.372	0.396
2004	26.091	0.352
2005	25.723	0.389
2006	25.822	0.404
2007	25.778	0.412
2008	25.864	0.401
2009	25.958	0.483
2010	25.915	0.435
2011	25.800	0.445

Back to list

Non compliance to class allocation

Dep. var	log absences	
	[1]	[2]
internal*post2006	-0.007	
	(0.047)	
junior pr*post 2006		-0.026
		(0.049)
full pr*post 2006		0.045
		(0.058)
N	3369	3369

- use as dependent variable n students absent (proxy class reshuffling)
- De Giorgi et al. (2010): not correlated with quality teachers/courses.

Back to list

Diff in Diff for teaching, student level, heterogeneity

Student level regression:

$$grade_{ipcy} = \zeta_p + \zeta_{cy} + \beta_1(int_p * p06_y) + \beta_2(int_p * p06_y * HSgr_i) + \beta_3 W_{ipy} + v_{ipcy}$$

- $HSgr_i$ is high school grade of student i (my proxy for ability).
  3 categories: (i) high ability (omitted)=between 1 and 0.9; middle ability = between 0.8 and 0.9; low ability: below 0.8
- int_p=1 if professor internal in 2005 (before announcement)
- ζ_p are teacher fixed effects
- $p06_y$=1 if after academic year 2006
- ζ_{cy} are fixed effects for the interaction between time and courses
- W_{ipcy} are professor and students characteristics (age, gender, hs grade, whether Italian etc.)
- v_{pcy} is the error term

Student level regressions: heterogeneous effects

	Dependent variable: stud grade (std)			
	[1]	[2]	[3]	[4]
int*post06	-0.037***		0.002	
	(0.014)		(0.016)	
jun pr*post06		-0.045***		-0.005
		(0.016)		(0.017)
full pr*post06		-0.009		0.028
		(0.020)		(0.022)
int*post06*mid ability			-0.079***	
			(0.014)	
int*post06*low ability			-0.097***	
			(0.020)	
jun*post06*mid ability				-0.077***
				(0.015)
jun*post06*low ability				-0.100***
				(0.021)
ord*post06*mid ability				-0.086***
				(0.022)
ord*post06*low ability				-0.086**
				(0.036)
N	346628	346628	346628	346628
Teachers fe	Yes	Yes	Yes	Yes
Year*course*degree pr fe	Yes	Yes	Yes	Yes

Control set: teacher age, age sq, student gender, hs, whether Italian, whether from Milano. Se clustered by teacher.

Research spillovers

- Spillover effects coming from interactions with better researchers that operate differently for internal and external teachers may amplify or shrink the effect
- I look at 2 outcomes % of articles: (i) coauthored with other bocconi internal faculty members; (ii) coauthored with Bocconi new hires

Dep. var:	% articles co-authored with	
	internal	new hires
	[1]	[2]
Internal*post2006	-0.032	0.001
	(0.023)	(0.023)
N	5230	4699
Teachers fe	Yes	Yes

Back to list

Student level regressions: heterogeneous effects

	Dependent variable: stud grade (std)			
	[1]	[2]	[3]	[4]
int*post06	-0.037***		0.002	
	(0.014)		(0.016)	
jun pr*post06		-0.045***		-0.005
		(0.016)		(0.017)
full pr*post06		-0.009		0.028
		(0.020)		(0.022)
int*post06*mid ability			-0.079***	
			(0.014)	
int*post06*low ability			-0.097***	
			(0.020)	
jun*post06*mid ability				-0.077***
				(0.015)
jun*post06*low ability				-0.100***
				(0.021)
ord*post06*mid ability				-0.086***
				(0.022)
ord*post06*low ability				-0.086**
				(0.036)
N	346628	346628	346628	346628
Teachers fe	Yes	Yes	Yes	Yes
Year*course*degree pr fe	Yes	Yes	Yes	Yes

Control set: teacher age, age sq, student gender, hs, whether Italian, whether from Milano. Se clustered by teacher.