Decentralized academic selection mechanisms in Italy: opportunity or parochialism?

S. Verzillo
University of Milan

September 15, 2013

Abstract

This paper analyzes the impact of decentralizing academic recruitment mechanisms in Italy on professors observable research productivity following a policy reform introduced in 1998 (Berlinguers reform act). Is decentralization an opportunity to select higher quality researchers or is it a way to increase parochialism? To answer these questions we study the differences between individual productivity measures of researchers being hired before and after this reform in Italy. We focus our analysis on differences in associate professors research trends and mid-term outcome levels (research impact, productivity and notoriety as reported by ISI-web of Knowledge) six-years after being hired. Results suggest that an overall general worsening effect of this reform is not clearly identified due to a big heterogeneity of our research measures within scientific disciplines. Restring the world to "hard sciences" only, negative effects on research outcomes and trends are identified for researcher belonging to Math, Earth Sciences, Medicine and Veterinary disciplinary areas. We find both lower individual productivity and lower impact associated with decentralization.

Keywords: policy evaluation; coarsened exact matching; research productivity;

Acknowledgments: the author would like to thank D. Checchi, E. Rettore, S. Iacus, P.G. Lovaglio, R. Miniaci for their helpful suggestions and comments.

1 Introduction

In 1998, the Ministry of Education, University and Research (Miur) signed a new reform of the Italian Academy that introduced the possibility for

each university to locally recruit new researchers instead of filling their vacancies hiring professors selected by a national committee ("concorso nazionale" was substantially dismissed after twenty years). Decentralized academic recruitment has the advantage to meet immediate hiring needs of local departments with less burocratic load but could obviously be the source of differences in the quality of the selection processes. Economists have long argued about the role of incentives in influences individual performances but they rarely discussed about the impact of different recruitment mechanisms on individual productivity when the prevalent incentive for workers is a future promotion.
In particular foreknowing the future recruitment rules could influence the research publication strategies of young academics (particularly assistant and associate professors) to maximize their individual probability of being hired as associates or full professors under the new system. This study answers the following research questions:

- has decentralization had a significant impact on the prevailing level of research productivity of newly hired academics in Italy ?
- to what extent does the mechanism of recruitment provide incentives for researchers to reach better publication scores after selection ?

The aim of this study is to determine whether local recruitment has an impact not only on the "level of quality" of selected researchers ([6]) but also on their mid-term research outcomes ([10] and [18]).
We investigate this issue using an ad hoc dataset on Italian academics between 1991 and 2011. We especially focus on the impact of the 1998 decentralization reform of the Italian university system on research outcomes for candidates publishing in international journals, bearing in mind that all other aspects of the system remained unchanged over the last two decades (salary benefits, university funding mechanisms etc).
Data from the web version of ISI-Web of Knowledge (WoK) were collected to obtain standard and comparable bibliometric indicators of Italian researchers, while administrative records regarding their affiliation, academic position and disciplinary area are given by the National Ministry of Education, University and Research (henceforth Miur). From a theoretical viewpoint, it is possible that decentralization of recruitment mechanisms reduces the incentives for candidates to produce international research outcomes (conference papers, journal articles etc..) and to submit papers to higher-quality scientific journals, which usually implies longer publication times and lower acceptance rates. Local recruitment management could generate the expectation that less stringent research requirements will be applied focussing more on teaching experiences or administrative issues. This consideration would be most crucial for applicants to assistant professorships and for assistant professors applying for associate professorships rather than for newly appointed full professors. Indeed full professors productivity would not be driven by future careers incentives (this would be the case only considering the possibility of being hired by foreign universities or research centres) but mainly by academic reputation. Our results document that decentralization, controlling for academic discipline, has a negative effect on ISI research outcomes (especially on measures of research impact) in mid-term performances of hired people.

Despite of the increasing interest in literature on the role of connections in both academic promotions ([4] and [7]) and research publications ([14] ; [15]) the empirical evidence on academic changing due to recruitment reforms is relatively scarce. In the Italian case Perotti (2002; [21]) shows how the number of votes the external commissionaires under the local systems are completely unrelated with their scientific reputation while Durante et al. (2009; [9]) analyze the relationship between nepotism and the level of civic capital associared with decentralization. In particular they find an increase of nepotistic misbehaviours (measured comparing the pattern of family names in the recruiting university and among the candidates) associated with lower level of civic capital after the 1998 reform. The paper is structured as follows. Section 2 introduces the Italian Academic System. Section 3 briefly discusses from a theoretical viewpoint the incentives of local recruitment versus national. Sections 4 describes the data and the use of bibliometric indicators as research productivity measures. Section 5 introduces the "quasi-experimental" empirical strategy through Coarsened Exact Matching and the growth model setting. Section 6 presents main results while section 7 concludes.

2 The Italian Academic System

The Italian academic system is composed of 89 universities (28 private and 61 public) and 6 higher education institutions. The latter usually dispense only masters and PhD courses, being more research oriented than most of the other universities. Three out of the 61 public universities are polytechnics and eleven out of the 28 private institutions are distance-learning universities. The university system is divided into 372 sectors of discipline ("settore scientifico-disciplinare"), grouped into fourteen research areas, as designated by the Italian National University Council (CUN) . Sectors of discipline are categorized for homogeneity within each research area, and the selection of research candidates was conducted by recruitment commissions within each academic discipline under both national and local systems. Considering academic disciplines as the reference level of analysis in this study and always controlling for the academic sector (by including dummies for each "settore scientifico-disciplinare" in our models) ensure validity in accounting for heterogeneity of recruitment behaviours between disciplines.
In general the Italian university system is constrained by national regulations. Each professor working at an Italian university is categorized by a level of arrangement (full professor, associate professor and assistant professor) and by one out of 372 sectors of discipline. Each vacancy is coded in a standardized format, and each filled position becomes tenured after a review conducted three years after hiring. Each position is also associated to a school ("faculty") for teaching duties and to a department for research activity. Salaries in public universities are set by law and vary only by level of arrangement and seniority. Schools and departments are prevented from differentiating wages among professors, linking payment to research productivity and/or teaching loads.

As a consequence, in addition to celebrity and funds attraction, the strongest incentive to scientific productivity for individuals working in academia derives from expected promotion (being hired as associate professor and being promoted as full professor). Given the public nature of the employment contracts, university professors can only be hired through public competitions that should grant publicity of the vacancy, selection of the selecting committee based on objective criteria, transparency of the selection process. This may explain why it is crucial to study the different incentives designed by different selection procedures in terms of research productivity and quality levels. A reform in 1998 changed these procedures with respect to at least three dimensions:

1. level of selection (national or local), which mostly affects the number of competing applicants, but also the timing of the selection due to the heavier bureaucratic load associated with a nationwide competition;
2. selection of committees (in accordance to the cooptation attitude of academia, for most of the period under analysis the committees were elected out of professors of the same sector of discipline, with element of randomness introduced at some stage);
3. number of eligible applicants (each "concorso" declares a number of winners that are eligible to become professors, this number is usually equal to, but sometimes greater than, the number of available vacancies);

Since 1979, standardized competitions were held to hire assistants, associate and full professors, and until 1998, almost all academic recruitment was substantially centralized. Despite the legislative prescription of one concorso every two years, a three to four years interval occurred. National commissions of five members were chosen by lot within a pool of elected professors (from a pool of 15) belonging to the same discipline. Commissioners declared which of the candidates had the qualifications to be promoted to associate/full professorship. Eligibility was given to a number of candidates greater then the available positions (usually 20% higher) for each discipline. Universities with opening positions drew by multilateral bargaining between them from the list of eligible applicants to fulfil their vacancies ([5] Checchi shows some evidences from a single national selection procedure for associate professorship). Starting in 1999, recruitment procedures became entirely local, and each university could hold its own selection procedure (for assistants, associates and full professors). Local commissions were comprised of five members: one belonging to the institution itself -the internal commissioner- and the four others elected by the full set of Italian professors of that discipline.
In 2004, a new reform act established that the commissions members had to be drawn by lot in a pool of professors of three times the size of the local commission, elected by popular vote amongst the disciplines affiliates (a recent working paper investigating whether this reform decreased the relative weight of the internal examiner is Dal Maso et al [8]). The commissions initially declared three qualified candidates for each concorso, but moved to two between 2007 until 2008, and only one thereafter. In the following years, universities with open vacancies could hire any candidate who had obtained a qualification. Professors hired under the new policy

mechanism were engaged beginning in 2000, two years after the enactment of Berlinguers reform. Consequently, in the following empirical analysis we mark the beginning of decentralization that year.

3 Theoretical incentives

The majority of other factors that affecting research performance over this time has remained constant, setting the stage for a natural experiment for considering the effect of decentralization. For this reason decentralization of academic recruitment could be considered as an exogenous shock to recruitment rules that potentially impacted the subsequent career of selected professors. Notwithstanding the hiring changes, university funding mechanism remained totally disconnected from managerial behaviours. No salary incentives (or penalties) were provided to incentivize (or prevent) virtuous behaviours of the commissioners in selecting high (or low) quality candidates; student evaluations of teaching performance for new hired professors usually had no impact on the professional life of professors (although including aspects of these evaluations become compulsory after 2000). Due to strict privacy rules, the results of evaluation exercises were in most of the cases known only by each professor until recently; no evaluation mechanisms were established to assess recruitment procedures at department (or university) level by the central administration.
Thus, over the time period we consider, there was no private cost for opportunistic behaviour for (part of) the selecting committee, as well as no impact on institutional funding mechanisms, except for lower scientific reputation. In this context decentralizing academic recruitment could have mixed effects. On one hand, decentralization could improve productivity and efficiency for at least three reasons; first, local recruitment usually induces speedier selection procedures (national concorsi were held every 4-6 years) while the multilateral bargaining between winners and hosting universities could last one or two years under national mechanisms, second, they guarantee more certainty of fulfilling available vacancies when needs arise (both for research and more often according to teaching necessities), and third decentralization could lead to more competition among universities in attracting candidates (especially by institutional reputation). The better candidate could also be the one that particularly fits with institutional needs (in terms of research competences and experiences).
Moreover decentralization means less compromise with "The Academia" when it is dominated by few national prevailing "Schools". The Italian academia is certainly not so unfamiliar with such corporatism ([9]; [2]) and the existence of schools that could exert a direct influence on the selected candidates has to be considered as a possible problem (for a couple of studies on the role of connections in academic promotion in Spain and France see [4] and [7]).
On the other side, decentralized selections enable institution to favour individuals based on familiar, professional or political connections, independent of their experience or qualifications. Local processes also may lead to less competition with respect to national procedures. In addition, it is im-

portant to analyse the extent to which influence of the selection procedures induces behaviours of the potential candidates.
Before reforms, national "concorsi" were held less frequently and involved a larger number of interested applicants. This meant more competitors. In most circumstances, a greater number of competitors encourage more effort on the part of the candidates, providing incentives for individuals to maximize their probability of winning through performance. Furthermore, the number of peers under central selections was stable in the Italian context, meaning that opportunities for advancement could be considered equal over time for each "concorso". Decentralizing academic recruitment meant fewer competitors participate in "concorsi" both in the present, and future.
Geographical constraints are also important because, with a local system, there is an incentive to participate in "concorsi" that are relatively close to the candidates home area, rather than compete in all Italian "concorsi" for that year. If publishing more papers in impact journals increases an individuals probability of being selected (as should theoretically be the case), decentralizing academic procedures could provide fewer incentives for local candidates. Indeed the individual choice of putting less or more effort is not mainly driven only by the own candidate willingness to exert it but it could probably be influenced by also the selection year.
Then our research question is twofold: has decentralization of selection mechanisms improved (or worsened) mid-term research outcomes of winners? have local recruitment mechanisms incentivized (or discouraged) individual research productivity?

4 Data

Data for this paper were collected from two primary sources: MIUR Italian Ministry of Education, University and Research - and the bibliographic Web version of ISI WoK database (henceforth ISI).
ISI, powered by the Institute for Scientific Information and distributed by Thomson Reuters, has been the standard in the bibliometric field for the past 30 years and indexes more than 8.700 journals in the fields of arts, humanities, sciences and social sciences. Scopus published by Elsevier, indexes a greater number of journals (12.850, including 500 open access journals) within the medicine, technical and social sciences.
Alternative bibliometric sources are Scopus and Google Schoolar. Scopus (henceforth SCO) is significantly larger in size and covers more of the international literature, but it completely excludes the humanities ([17]). Google Scholar (henceforth GS) stands-today-as the main potential competitor of ISI and Scopus (particularly in light of the fact that it is the only one without commercial interests), but currently has outstanding information reliability problems.
However the literature documents the presence of high correlations between and among bibliometric measures obtained by Scopus and ISI ([3]). Thus, despite using a single source, we expected similar results of analysis using different bibliometric sources instead of ISI.

Table 1: Descriptive Statistics

Final database					Official statistics*					Delta (%)
Year	Full Prof.	Ass. Prof.	Assist. Prof.	Tot.	Year	Full Prof.	Ass. Prof.	Assist. Prof.	Tot.	
1991	11750	14042	15642	41434	1991	***	***	***	45248	0,08%
1992	11804	16746	14964	43514	1992	***	***	***	***	***
1993	11876	17084	15739	44699	1993	***	***	***	47839	0,06%
1994	13288	15915	16694	45897	1994	***	***	***	47824	0,04%
1995	14011	16313	18417	48741	1995	***	***	***	49098	0,0%
1996	13719	16093	19583	49395	1996	***	***	***	48560	-0,01%
1997	13399	15675	20105	49179	1997	13402	15619	20167	49187	0,0%
1998	13098	18108	18748	49954	1998	13103	18108	18745	49956	0,0%
1999	12905	18069	19815	50789	1999	12899	17863	19949	50711	-0,2%
2000	14411	16615	19200	50226	2000	14676	16973	19542	51191	1,9%
2001	16901	17879	20255	55035	2001	16418	17572	20011	54001	-1,9%
2002	18148	18504	21055	57707	2002	17571	18100	20714	56385	-2,3%
2003	17997	18115	20577	56689	2003	17388	17783	20371	55542	-2,1%
2004	18062	18094	21341	57497	2004	17469	17633	21149	56251	-2,2%
2005	19296	18982	22186	60464	2005	19147	18849	21904	59900	-0,9%
2006	19843	19084	23355	62282	2006	19676	18966	23099	61741	-0,9%
2007	19640	18776	23793	62209	2007	19623	18739	23560	61922	-0,5%
2008	18929	18253	25923	63105	2008	18932	18261	25569	62762	-0,5%
2009	17980	17630	25911	61521	2009	17878	17567	25434	60879	-1,1%
2010	15949	16967	25590	58506	2010	15834	16745	24784	57363	-2,0%

The main drawback we face concerns the disciplinary coverage specification of international research data with respect to the whole research domain. Researchers of some disciplines such as History and Literature in Italy usually publish on national journals only (usually with articles written in Italian). For these academics, little bibliometric information is available on ISI (and also on SCO or GS).
Distortion of data due to higher 'ISI exposure' of some disciplines in comparison to others can only be managed by disaggregating the analysis by discipline, which is the approach we employ. Our data include around 1.000.000 ISI products over the last 20 years. After filtering duplicates and incomplete records were deleted obtaining a consistent database of 963.181 scientific publications with at least one Italian author over the period 1991-2010. Information regarding academic positions, individual characteristics (age and gender), disciplinary areas and university affiliation are available online from 2000 to 2011 . We obtained data on academic careers before 2000 from Cineca, a MIUR agency which collects administrative data on personnel as well as on competition for professorship in Italy. These data have several known problems, often relating to the uniqueness of identifying codes of individuals, and missing data on academic disciplines over the first five years (1990-1995). After we corrected for these issues to the best of our ability, we found a 1.5% degree of imbalance with respect to the last available official statistics published by Miur (reported in table 1). These differences are likely due to a few rare categories of professors, such as newly hired associate and full professors attracted from abroad, fixed-contract new researchers positions and so on. However, such a small difference is unlikely to seriously bias our results, or be the cause of distortionary effects in our estimation procedures.

We employ a multi-step matching procedure to assign the corresponding author identifying codes to each research product in the bibliometric dataset. At the end we obtain a 91% of matched records. Indeed a 9% percent of ISI products is plausibly stored in the database reporting an Italian affiliation but with an author who is not included in the official faculty list provided by Miur, which may be plausible due to the presence of post-doctoral students, PhD candidates and individual researchers not included

Table 2: Percentage of professors with at least one paper on ISI by year

Year	No	Yes	Total
1991	86,03	13,97	100
1992	83,14	16,86	100
1993	80,94	19,06	100
1994	78,87	21,13	100
1995	76,35	23,65	100
1996	74,35	25,65	100
1997	72,51	27,49	100
1998	70,12	29,88	100
1999	68,15	31,85	100
2000	65,71	34,29	100
2001	62,24	37,76	100
2002	59,06	40,94	100
2003	57,19	42,81	100
2004	55,14	44,86	100
2005	52,07	47,93	100
2006	49,12	50,88	100
2007	46,84	53,16	100
2008	44,21	55,79	100
2009	42,32	57,68	100
2010	40,52	59,48	100
2011	38,25	61,75	100
Total	62,05	37,95	100

in Miur research and teaching faculty list.

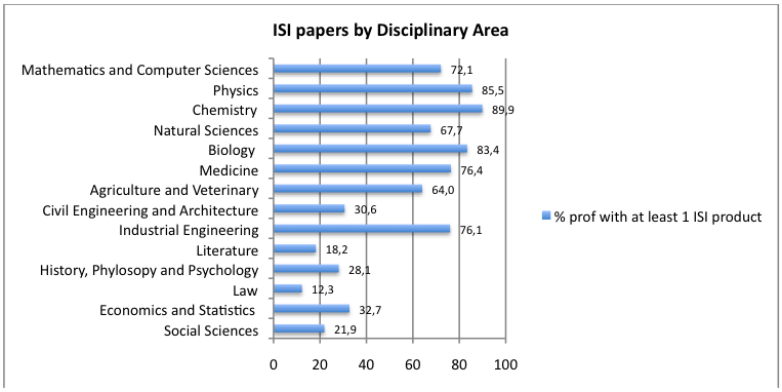
Our data show different numbers of individuals over time, the panel for each academic professor is unbalanced due to varying entry points into the administrative archives of Miur. Data vary depending on the year of selection, the year of the first published international paper, the persistence rate of publication on ISI of the own discipline. Discontinuities are also possible (and several are identified in our analysis), due to such considerations as working abroad, or that the individual enters unusually one year and ceases publishing after that point. These issues could have significant effects on the distribution of bibliometric indicators over time, and the challenges are more pronounced in some disciplines than in others. We arbitrarily decide a threshold in order to exclude disciplines with lower level of individuals with at least one product on ISI. Heterogeneity within academic research areas is highly effective in ISI studies, and figure (fig. 1) give us the idea of which of the standard disciplines overcome the 50% cut-off level of individuals with at least one record.

To face these problems we consider in this study especially the scientific area where this percentage is greater than 50% with respect to the different historical and individual nature of each discipline. Bibliometric indicators could be theoretically considered for all the scientific areas but a rate

Table 3: Fig.1 - Percentage of professors with at least one paper on ISI by Academic Position

Academic Position	**At least one ISI paper**		
	No	Yes	Total
Assistant Professor	50,98	49,02	100
Associate Professor	45,17	54,83	100
Full Professor	41,41	58,59	100
Total	46,32	53,68	100

greater than 50% guarantee a degree of reliability in our research exercise. A high sensitivity level among academic fields and aggregation levels is common in bibliometric studies. Thus we reduce the area on less heterogeneous disciplines this way and build the dataset longitudinally, considering year-of-selection as the (moving) starting point for all individuals considered.



Percentage of Professors with at least one ISI paper by disciplinary area

4.1 Bibliometric indicators

Three simple bibliometric measures are introduced. Firstly, as quantity of published research we calculate the individual cumulative frequency of ISI items extracted from the database at each year. This measure gives an idea of the quantity of papers on international journals published by the considered researcher up to each considered year.
Secondly, we calculate the cumulative average impact factor of each academic professor at each of the considered years. This measure could be considered as an individual "expected measure of impact".
The second measure is easily open to criticism. On one hand, one could argue that it is not correct to use journal impact factors to evaluate individual productivity and, on the other hand, that the impact factor could

not be considered as a proper measure of "quality". Indeed, impact factors are by definition the arithmetic average of citations of the journal papers in a given period. Referring the average impact factor to each article means that the underlying distribution of citations is assumed to be uniform. But this is empirically unproven, and in some respects refuted by observation. T The literature documents that the underlying distribution is strongly asymmetric to the right; few articles receive many citations, and most receive few ([23]; [1]). It is certainly the case however that the impact factor of a journal provides a priori information about the "expected number of citations" of a published article. In the absence of acceptable citations indexes, we consider the average impact factor of the journals were a researcher has published as a measure of the expected impact in terms of citations the researcher will obtain.
Secondly, it is also true that papers published in top journals with high reputation (high impact factor) have already been peer-reviewed by rigorous referees; this process should guarantee high quality standards of the published research. There is thus justification for thinking that a strong correlation exists between journal impact factors and research quality of its published papers.
Then a measure of network extent was extracted using the total number of citations. The significance of the network rate measure has to be widely discussed in literature because of its strong relationship with time. The greater is the number of years an article has been published, the higher is the number of citations due to time exposure on ISI. ISI citations are by definition updated to download time (April, 2012) and they could not be considered comparable measures across years without accounting for their time of exposure.
Descriptive statistics for assistant professors at year of selection evidences the wide differences of bibliometric indicators within academic disciplines. "Hard sciences" professors in Physics or Chemical Sciences at time of their selection had an average of 24 ISI papers, with 2.5 average impact factor, 15 average citations per paper, and more than 400 cumulative citations. On the other end of the spectrum, Arts and Humanities academics have on average less than 1 paper on ISI, with 0.33 impact factor, 1.7 average citations per paper, and 7 cumulative citations in their research careers.
Finally we take log-transformations of research outcomes (impact factor and number of papers on ISI) in our analysis to guarantee the normality of both variable distributions.

5 Empirical Strategy

The methodological approach we employ to evaluate the impact of different selection mechanisms caters to the specific research question we ask: Is there a causal effect of local (versus national) recruitment programs on the subsequent research productivity levels of selected academics ?
We focus on evaluating the effects of a shift to decentralized selection mechanisms in terms of subsequent research productivity average level of the

Table 4: Descriptive statistics of bibliometrics indicator by disciplinary area; Associate Professors at year of selection

Id	Discipline	Freq	N. paper	Imp. factor	citations (avg)	citations (cum)
0	Missing (1991-1995)	279	3,31	0,94	5,50	63,05
1	Math. and Computer Sciences	1.282	7,75	1,18	7,07	65,46
2	Physics	1.887	24,04	2,62	15,30	402,27
3	Chemistry	1.139	24,04	2,99	17,96	500,21
4	Natural Sciences	1.644	6,14	1,70	13,75	105,36
5	Biology	1.268	16,09	3,43	19,40	344,71
6	Medicine	948	19,47	3,73	14,81	361,47
7	Agriculture and Veterinary	376	6,33	1,38	9,98	90,35
8	Civil Eng. and Architecture	1.777	2,21	0,68	4,23	24,64
9	Industrial Eng.	563	10,73	1,43	8,68	107,67
10	Literature	1.471	0,70	0,33	1,82	8,89
11	History and Philosophy	964	1,49	0,72	3,74	23,24
12	Law	2.205	0,43	0,39	1,75	7,41
13	Economics and Statistics	558	1,24	0,66	4,08	14,96
14	Social Sciences	1.578	0,62	0,50	1,76	7,79

outcome and its time trend. The treatment status can be considered as the exposure of an individual to local selections instead of national ones. The problem is that we can observe almost one of these states for each individual of interest.
Indeed individuals who are exposed to local selection programs are by definition (due to a specific time constraint: they were selected after 2000) different from those who are exposed to national recruitment programs. These differences may invalidate the causal comparison of the impact of decentralization on future research productivity outcomes.
Recent studies in the econometric literature of program evaluation ([16]) and methodological research on causal inference ([22]) from observational studies (where investigators have no control over the treatment assignment) suggests the use of parametric methods, such as propensity score, or non-parametric strategies (as CEM) to accommodate general heterogeneity between two groups of individuals in estimating the treatment effects and to increase precision of the estimates. The treated (in our empirical application all the researchers selected with local mechanisms) and control (the selected with national mechanisms) groups may have significant differences in their observed covariates (scientific discipline and research productivity outcomes) that could lead to biased estimates of the selected effect. Our analysis implements a recent non-parametric method to obtain balanced treated and controls groups: the Coarsened Exact Matching (CEM) technique. ([12])

5.1 Coarsened Exact Matching

Coarsened exact matching is a matching method recently introduced by Iacus et al. [12] to improve causal inference controlling for the confounding influence of covariates in observational studies.
The time-dependent nature of our study, and the flexibility of CEM in estimating non-parametrically two balanced distributions of treated (locally selected professors) and controls (nationally selected) units, allows us to obtain desired counterfactuals for estimating the decentralization effect.
Balanced groups avoids having the researcher control for the heterogeneity while specifying the model, meaning that simple differences in means are

good estimates of the causal effect.
But usually finding a matching solution in empirical propensity applications does not guarantee good balance to all the selected covariates. Improving balance on most of them could leave the remainders imbalanced, often introducing more bias with respect to the initial distribution ([12]).
In addition to this, propensity score matching (and Mahalanobis distance methods) has the drawback of violating the congruence principle, which requires congruencies between data and analysis spaces metrics (the own metric of the two spaces is different). Parametric methods usually force covariates of the original data from a multi-dimensional original space in a new space defined by the propensity itself. Mielke and Berry ([19]) show how violating this principle produces less robust inferences. In comparison, coarsened exact matching meets the principle of not reducing the original data space, operating in the multidimensional variable space itself.

Applying CEM to our study means firstly to set variable-by-variable the non-overlapping intervals to coarsen original data about winners of selection procedures (before and after the reform) at the year of their selection. Then we match one-to-one each stratum treated and controls units after removing all the individuals (treated and controls) owning to zero controls strata.
Available covariates regarding bibliometric indicators and disciplinary area are coarsened according to reasonable assumptions: a 0.3 impact factor intervals, a one-to-one number of ISI publications and a 10 pages interval of cumulative number of pages written by the authors are settled as coarsening rules.
Missing data are treated as "missing as zeros" due to the particular nature of our data. Indeed, missing values of bibliometric indicators (impact factor, number of paper ISI, citations, sum of pages etc) reflect the absence of the author in the data and absence on ISI is equal to 0 international papers published, with 0 pages written, 0 citations received and 0 average impact factor. A real drawback of missing replaced with zeros could be represented by the equal "treatment" of an author with few ISI publications with zero impact factor and 0 citations and an author without ISI records. However, equal treatment of zero ISI publications or few records with no impact factor and no citations at the associate professorship level in our restricted word (hard sciences only) could be considered, without a significant loss of information, acceptable. The desirable output of this procedure is a sample of balanced treated and controls. For this case, we found 3.181 treated professors with one-to-one coarsened exactly matched controls over 5.292 potentially possible 1:1 couples.

The selected sample population is now composed of comparable subgroups of individuals (selected before and after the reform) with similar levels of bibliometric indicators (according to the coarsened intervals settled as before) and operating their research effort in the same disciplinary areas.

Table 5: Frequencies of Treated and Untreated units by CEM groups

Groups	Frequencies		Sample
Treated	total	12.646	3.181
	zero controls strata	6.096	
Untreated	total	5.292	3.181
	zero controls strata	1.727	
	matched		6.362

Table 6: Descriptive statistics of associates by treated and controls

treat=0					
Variable	**N**	**Min**	**Mean**	**Max**	**StdDev**
Impact factor	5.292	0	1,289	25,285	1,954
N of paper	5.292	0	5,002	133	10,156
Sum of written pages	5.292	0	35,726	923	67,93
treat=1					
Variable	**N**	**Min**	**Mean**	**Max**	**StdDev**
Impact factor	12.647	0	1,656	53,48	2,245
N of paper	12.647	0	9,559	231	16,334
Sum of written pages	12.647	0	253,151	923	5581

5.2 Multidimensional balancing

The following table (tab. 7) shows statistics for three of the four selected variables, impact factor, number of paper ISI and cumulative sum of written pages. Frequencies and descriptive statistics reported underline the differences between the two sub-populations (most of the difference is plausibly due to the different time horizon at which the two populations refer to).

Table 7 provides first evidence (univariate absolute difference in means) of balancing between CEM selected treatment and control groups in the overall sample. Mean and standard deviations of the two, equal-size, samples of units are relatively close from one to the next. Before and after reform associate professors have an average number of 1.2 papers published on ISI journals with an average impact factor of 0.45 and a number of written pages close to 10.

By construction, covariate descriptive statistics over the entire sample are almost equal to descriptive statistics of units in each of the selected disciplinary areas. However, despite being commonly use in observational studies (especially in propensity score studies), univariate distributions of means do not guarantee the absence of bias in estimating the treatment effect. Recent studies ([13]) looking at the multidimensional histograms of the two samples (for treated and controls) introduce methods to check for multivariate balancing of their empirical distributions. They propose

Table 7: Descriptive statistics of associates matched units by treated and controls

treat=0					
Variable	**N**	**Min**	**Mean**	**Max**	**StdDev**
Impact factor	3.181	0	0,459	10,61	0,998
N of paper	3.181	0	1,201	41	3,247
Sum of written pages	3.181	0	9,492	249	24,793
treat=1					
Variable	**N**	**Min**	**Mean**	**Max**	**StdDev**
Impact factor	3.181	0	0,457	10,55	0,995
N of paper	3.181	0	1,222	49	3,365
Sum of written pages	3.181	0	9,80	1255	27,496

Table 8: L1 matched and original population multivariate balance measures

L1 matched	L1 population
0,19	0,92

a measure of imbalance (L_1) that is the semi-sum of the absolute differences between relative frequencies of treated and controls for each identified strata in our case. L_1 for the entire population is close to 1 (highly unbalanced distribution of treated and controls). This means that a substantial number of cells in the multidimensional matrix have zero controls (or treated). Comparing the L1 of the matched population with the previous one provides evidence of the unbalanced reduction due to CEM. L_1 is equal to 0.19 after CEM, this means high rate of balancing between the populations of treated and controls (table 8).

We then plot parallel coordinates plot as a visualization method for detecting patterns of matched and unmatched units in a multivariate setting. Looking at the subsequent graphs (3a), it appears that the associate professor matched individuals are relatively well distributed between the considered dimensions; they belong to all the academic disciplines, produce a relatively low number of papers, with average citations and impact factor on ISI in the first bottom half of the distributions. So, despite of the fact that it could theoretically be the case that matched units belong to different points of the original treated and controls ability distributions (considered in its multidimensional bibliometric setting), matched units seem to be effectively comparable for our research purposes.

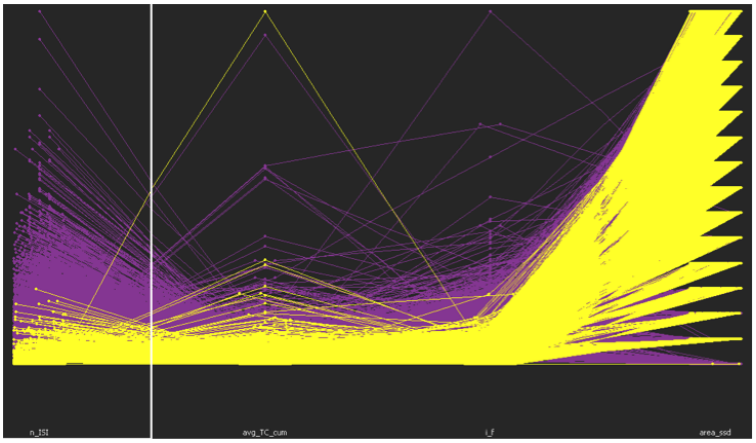


Fig. 3a - Associates Professors parallel coordinates plot across included covariates

We then fit both a standard treatment model and a growth longitudinal model ([20]; [11]) through the individual careers of treated and controls that allow us to identify a negative effect of decentralization on research slopes (impact and quantity) of selected individuals, for both associate and full professors. A growth model is specifically designed for exploring longitudinal structure of the data and, despite of what happens in commonly specified generalized linear models where estimations are carried out forcing the pattern of covariances (or correlations) to be constant across time, the growth model allows all subjects in each group to change differently over time. This is a more realistic assumption especially thinking of the heterogeneity of research patterns across individuals in our setting. First, let us consider a model of the changes in logarithmic transformation of research output y_i across time as a function of treatment group, time and the interaction of treatment and time:

$$\log y_{it} = \beta_0 + \beta_1 * time_{it} + \beta_2 * treat_i + \beta_3 * (treat_i * time_{it}) + \beta_4 * X_{it} + v_{0i} + v_{1i} * time_{it} + \epsilon_{it}$$

where $\log y_{it}$ is the log-transformation of research output (impact factor, n of paper etc...). β_0 is the mean of the dependent variable of all the selected individuals at time zero (sixth year). The term $time_{it}$ is a time variable with values from -5 (at time of selection) to 0 (at the sixth year after selection) for each individual i. We consider a six years period after selection according to both the empirical average period of the associate careers before being appointed as full professors and the time lag due to review processes between submission and publication on international journals.
So, in terms of its representation, this model could be divided into a within-subject model:

$$\log y_{it} = b_{0i} + b_{1i} * time_{it} + {}_{it}$$

and the between-subjects model:

$$b_{0i} = \beta_0 + \beta_2 * treat_i + v_{0i}$$

Table 9: Growth model estimations - Associate Professors

Variable		Overall		Pre-94 vs. After-2000	
Parameter	treat	log impact factor	log nof paper	log impact factor	log nof paper
intercept		0,302*** [0,035]	2,197*** [0,066]	0,552*** [0,044]	1,98*** [0,072]
time		-0,026** [0,010]	0,126*** [0,020]	-0,011 [0,016]	0,21*** [0,026]
treat	0	0,077*** [0,017]	-0,051 [0,034]	0,137*** [0,027]	-0,201*** [0,043]
treat	1	.	.	.	.
time(treat)	0	0,081** [0,011]	0,0199*** [0,020]	0,008 [0,027]	0,035 [0,043]
time(treat)	1	.	.	.	.
treat a vs b intercept		-0,077*** [0,0017]	0,051 [0,034]	-0,137*** [0,027]	0,201*** [0,043]
treat a vs b slope		-0,081*** [0,011]	-0,0199*** [0,020]	-0,008 [0,027]	-0,035 [0,043]
Pr $\geq$ *ChiQuadr*		$\leq$ .0001	$\leq$ .0001	$\leq$ .0001	$\leq$ .0001
Obs.		6.362	6.362	5.111	5.111

*significant at 10%; ** significant at 5%; *** significant at 1%, standard errors in brackets; region, gender, age, calendar year dummies and scientific discipline controls included

$$b_{1i} = \beta_1 + \beta_3 * treat_i + v_{1i}$$

With this econometric characterization of treatment and time effects, we can interpret the parameters as follows: β_0 is the average of $\log y_i$ at time 0 (sixth year after selection) for the untreated group (nationally selected individuals, where treat=0); β_1 is the average trend across time for the untreated group (treat=0); β_2 is the average difference in $\log y_i$ at the sixth-year after selection between centrally and nationally selected professors; β_3 represents the average difference in trend lines between treated and untreated.

Furthermore this regression model allows each individual to deviate from the owning group trend line in terms of final intercept (v_{0i}) and time-trend across time (v_{1i}). Time trend is assumed to be linear (instead of a square root trend specification that probably be more realistic when considering the whole career of a researcher) accordingly to the idea of a constant growth rate of publication of associates in their first six years after selection.

We also included in our models individual controls such as age, gender, institution and academic sector ("settore scientifico-disciplinare") to keep into account for the presence of researchers heterogeneity. A set of calendar year dummies is also included to avoid the possible bias due to both measurement error in bibliometric outcomes (discipline coverage and total number of considered journals vary according to the years in the last two decades with a clear increasing trend) and differences in research publication behaviour of the Italian academia.

6 Empirical Results

Significant negative effects regarding local selections on both final outcomes level (β_2) and slope differences (β_3) are estimated for impact factor outcomes in both career steps. Local selection negative effects are statistically significant only for slope differences with the quantity measure of international research.

Newly hired associate professors have fewer incentives to produce high quality papers in the subsequent years. The productivity outcome high-

Table 10: Standard and Growth model estimations by discipline - Associate Professors

Academic Discipline		Regression model			Growth model		
		Estimate	Std.Errors	Obs.	Estimate	Std.Errors	Obs.
Math and Computer Sciences	treat	-0,13***	0,042	537	-0,15***	0,034	537
	time(treat)	-0,013	0,014		-0,014***	0,015	
Physics	treat	0,22	0,133	151	0,31	0,17	151
	time(treat)	-0,057	0,042		-0,036	0,12	
Chemistry	treat	0,03	0,035	348	0,021	0,049	348
	time(treat)	-0,0039	0,011		-0,024	0,042	
Earth Sciences	treat	-0,22***	0,099	270	-0,17***	0,11	270
	time(treat)	-0,049	0,035		-0,038	0,09	
Biology	treat	0,04	0,014	620	0,085	0,071	620
	time(treat)	-0,015	0,061		-0,024	0,068	
Medicine	treat	-0,11***	0,044	1703	-0,12***	0,045	1703
	time(treat)	-0,015	0,015		-0,025	0,043	
Agricultural and Veterinary	treat	-0,19**	0,013	1371	-0,20**	0,051	1371
	time(treat)	-0,037***	0,051		0,0143	0,05	
Industrial and Information Eng.	treat	-0,028	0,029	2728	-0,011	0,029	2728
	time(treat)	0,003	0,011		0,007	0,028	

* significant at 10%; ** significant at 5%; *** significant at 1%, standard errors in brackets; region, gender, age and calendar year controls included

lights a negative difference in slopes (better for untreated) and no significant difference in the final quantity level. The effects are robust on pre-1994 vs. after-2000 professors considered as controls of the incentive schema due to the possibility of mixed strategies for people selected as associates after 1994 but before 2000 with national procedures, and applying as full professors with the local mechanism. The model results by each of the selected disciplines (table 10) shows the negative effects of local recruitment process over Math and Computer Science, Earth Sciences, Medicine and Agricultural and Veterinary in terms of final impact factor level (sixth year after selection) of associate professors. The slope is negative and significant only in the Match and Computer Science area. Both estimates of the standard regression model and growth model are basically consistent.

7 Conclusions

In this paper we analyze the role of the recruitment system on research incentives of selected associate professors. We focus on the period between 1991 and 2010, when the system of centralized competitions was reformed (1998) into a local recruitment system actually still in place.
We find negative effects of decentralization on individual research productivity (both in terms of slopes and six-year levels). Indeed lower incentives for publishing on international top-level journals are associated with the decentralization reform. In a nutshell, the main results of this paper are the following:

a) we found negative effects of decentralization on both final outputs levels and slopes for impact factors of associates (and also for full professors, see the appendix). The results were statistically significant and negative only for slopes the number of paper on ISI;

b) Differentiating by research areas, the negative incentives of local selections six years after selection research levels are statistically significant in the: Math, Earth sciences, Medicine, Veterinary and Agricultural Sciences;

c) Differentiating by research areas, the negative incentives of local selections on slopes are significant only in Math and Agricultural and Veterinary areas ;

d) Effects are robust on pre-1994 vs. after-2000 professors (used as controls of the incentive schema due to the possibility of mixed incentives for people selected as associates after 1994 but before 2000 with national procedures, and applying as full professors with the local mechanism);

However credible results are available for "hard sciences" only, due to the higher exposition of these disciplines onto the international research area (collected in ISI). Arts and Humanities and Social Sciences could not be tested due to the inconsistence of available data; We also have no data on the relative importance of teaching and research as the two main dimensions of academic recruitment. The effort of candidates is usually divided on both these activities in the years before each concorso, and it is likely the case that most universities in the last decade judged candidates with respect of both their level of research outcomes (quantity and impact) and teaching. These identified problems probably arise from the scarcity of competition between universities in Italy. An insufficient mobility rate of professors within the country (typical of local mechanisms), the predominance of institutional needs (both in research and teaching), and the preeminence of a small number of academic "schools" with respect to others, all have had a negative impact on research growth paths of local recruited researchers. No penalties were associated with collusive behaviors, and the lack of competitiveness between institutions is endemic in the Italian university system. This is probably due to a "false-autonomy" of universities where salary levels and teaching loads are centrally regulated (by MIUR).

Table 11: Growth model estimations by discipline - Full Professors

Parameter	**treat**	**log impact factor**	**log n of papers**
time		0,014*** [0,0044]	0,159*** [0,0066]
treat	0	0,129*** [0,0465]	0,014 [0,087]
treat	1	.	.
time(treat)	0	0,014** [0,0063]	0,051*** [0,0094]
time(treat)	1	.	.
treat a vs b intercept		-0,129*** [0,0465]	-0,014 [0,087]
treat a vs b slope		-0,014** [0,0063]	-0,051*** [0,0094]
Pr $\geq$ *ChiQuadr*		$\leq$ .0001	$\leq$ .0001
Obs.		3.090	2.763

* significant at 10%; ** significant at 5%; *** significant at 1%, standard errors in brackets; discipline, region, gender, age and calendar year controls included

8 Appendix

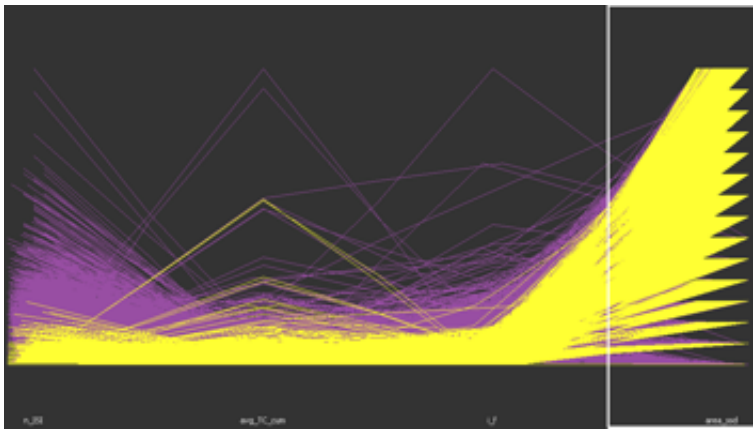


Fig. 3a - Associates Professors parallel coordinates plot across included covariates

References

[1] Adler R., J. Ewing J., Taylor P., (2009) Citation statistics (with discussion), Statistical Science, 24, 1-28;

[2] Allesina, S. (2011), Measuring nepotism through shared last names: the case of Italian academia, PLoS one 6(8), e21160;

[3] Archambault E., Campbell D., Gingras Y., Larivire V. (2009) Comparing Bibliometric Statistics Obtained from the Web of Science and Scopus, Journal of the American Society for Informatin Science and Technology, 60(7);

[4] Zinovyeva, N. and Bagues, M. F., (2012) "The Role of Connections in Academic Promotions," IZA Discussion Papers 6821, Institute for the Study of Labor (IZA);

[5] Checchi, D. (1999), Tenure. an appraisal of a national selection process for associate professorship, Giornale degli Economisti 58(2), 137181;

[6] Checchi, D., Verzillo S. (2013) "Selecting University Professors in Italy: Much 'Ado' about Nothing¿', Poster presented at the 25th annual Conference of the European Association of Labour Economists;

[7] Combes, P., Linnemer, L. and Visser, M. (2008) "Publish or peer-rich? The role of skills and networks in hiring economic professors", Labour Economics 15:423-441;

[8] Dal Maso, C., Rettore, E., Rocco, L., (2013) "Procedures vs. Incentives: the Case of the University Promotion System in Italy", Working paper presented at the EIEF workshop "Problems and Solutions for the Evaluation of Performance in the Public Administration", Roma, September, 13 - 2013;

[9] Durante, R., Labartino G., Perotti R., and Tabellini G. (2009), Academic dynasties, Bocconi University Working Paper;

[10] Fox, M. F. (1983), Publication Productivity Among Scientists: A Critical Review, Social Studies of Science, 13, 283-305;

[11] Howell, D.C., (2008) Mixed models for repeated (longitudinal) data;

[12] Iacus, S. M., King G., and Porro G., (2009) CEM: Coarsened Exact Matching Software, Journal of Statistical Software 30(9), http://gking.harvard.edu/cem;

[13] Iacus, S. M., King G., and Porro G., (2011) ;

[14] Laband, D. N. and Piette M. J. (1994), Favoritism versus Search for Good Papers: Empirical Evidence Regarding the Behavior of Journal Editors, Journal of Political Economy, Vol. 102(1), pp. 194-203;

[15] Li, D. (2011), Information, Bias, and Efficiency in Expert Evaluation: Evidence from the NIH, mimeo, MIT;

[16] Imbens, G., and Wooldridge, J.M., (2008) Recent developments in the Econometric of Program Evaluation, IZA DP No. 3640;

[17] Klavan and Boyak, (2007) Is there a Convergent Structure of Science? A Comparison of Maps using the ISI and Scopus Databases1, 11th International Conference of the International Society for Scientometrics and Informetrics meeting in Madrid, Spain, June 2007;

[18] Levin S., Stephan P. (1991), Research Productivity Over The Life Cycle: Evidence for Academic Scientists, The American Economic Review, Vol. 81, Issue 1, 114-132;

[19] Mielke and Berry (2007), Permutation methods: A distance function approach, New York: Springer;

[20] Overall, J. E., Ahn, C., Shivakumar, C., Kalburgi, Y. (1999) Problematic formulations of SAS Proc.Mixed models for repeated measurements, Journal of Biopharmaceutical Statistics, 9, 189-216;

[21] Perotti R. (2002), The Italian University System: Rules vs. Incentives, paper presented at the First Conference on Monitoring Italy, ISAE, Rome;

[22] Rosenbaum, P. R., and Rubin, D. B., (1983) The central role of the propensity score in observational studies for causal effects, Biometrika, Vol. 70, 41-55;

[23] Seglen P.O. (1997), Why the impact factor should not be used for evaluating research British Journal of Medicine, 314, 498-502;