

Effetti desiderati e indesiderati della valutazione della ricerca: evidenze empiriche da analisi bibliometriche

Ciriaco Andrea D'Angelo¹, Giovanni Abramo², Flavia Di Costa^{1,3}

¹ Dipartimento di Ingegneria dell'Impresa, Università degli Studi di Roma "Tor Vergata"

² Laboratorio di Studi sulla Valutazione della Ricerca, Istituto di Analisi dei Sistemi e Informatica (IASI-CNR)

³ Università Telematica Mercatorum

PRIN (2017NKWYFC) - Roma, 22 giugno 2023

1. Contesto di riferimento e obiettivi delle ricerche condotte
2. Specifiche metodologiche delle analisi bibliometriche
3. Risultati
 - 3.1 La produttività di sistema e quella delle università
 - 3.2 Il reclutamento
 - 3.3 Le collaborazioni di ricerca internazionali
 - 3.4 Specializzazione vs diversificazione
 - 3.5 Comportamenti opportunistici: autocitazioni e “gift authorship”
4. Conclusioni

Contesto di riferimento e obiettivi



Obiettivi della valutazione:

- Aumentare la produttività di ricerca dei singoli e delle organizzazioni
- Incrementare l'internazionalizzazione delle attività di ricerca
- Migliorare l'efficacia del reclutamento

...

Domande di ricerca:

- Sono stati raggiunti?
- Si evidenziano effetti «indesiderati» dovuti a criteri non adeguati?
- Si rilevano comportamenti opportunistici?
- Si rilevano altri effetti degni di nota a livello di sistema?



Specifiche metodologiche delle analisi bibliometriche



- Unità di osservazione: singoli ricercatori/docenti delle università italiane nei settori «bibliometrici»
- Fonte dei dati: *Web of Science Core Collection* di Clarivate Analytics.
- Periodi di osservazione
- Indicatori bibliometrici
- Approcci di analisi controfattuale: prima-dopo, serie temporali interrotte, dati panel ...

Criticità:

- Attribuzione degli effetti osservati, presenza di covarianti ed effetti spuri
- Ritardo temporale nella manifestazione degli effetti complessivi
- Limiti tipici dell'approccio bibliometrico

Il campo di osservazione: un esempio

		2009-2012		2013-2016	
Area disciplinare	SSD	Università	Ricercatori	Università	Ricercatori
1 – Matematica e informatica	10	71	3136	74	2946
2 – Fisica	8	65	2191	66	2097
3 – Chimica	12	61	2856	63	2742
4 – Scienze della terra	12	49	1049	51	987
5 – Biologia	19	69	4730	72	4535
6 – Medicina	50	65	9843	67	8906
7 – Scienze agrarie e veterinarie	30	57	2968	56	2867
8 – Ingegneria civile	10	53	1550	60	1508
9 – Ingegneria industriale e dell'informazione	42	74	5045	72	5095
10 – Psicologia	10	68	1328	69	1347
13 – Economia e statistica	12	84	2788	86	2751
Totale	215	92	37484	94	35781

Livello di analisi:
MICRO

Risultati - La produttività* di sistema



2009-2012 vs 2013-2016

	N. di SSD	Con produttività decrescente	Decremento medio pesato	Con produttività crescente	Incremento medio pesato	Variazione media pesata
MAT	10	2	-6.4%	8	+35.0%	+30.5%
FIS	8	2	-3.4%	6	+17.9%	+12.7%
CHIM	12	2	-3.5%	10	+8.7%	+8.2%
TERRA	12	4	-7.5%	8	+44.6%	+26.9%
BIO	19	2	-13.2%	17	+16.3%	+14.1%
MED	50	14	-7.0%	36	+20.8%	+13.7%
AGR+VET	30	4	-11.9%	26	+52.0%	+44.7%
ING CIV	10	0	n.a.	10	+63.1%	+63.1%
ING IND-INF	42	4	-20.0%	38	+45.3%	+44.6%
PSI	10	2	-6.9%	8	+80.4%	+66.6%
ECO+STAT	12	1	-14.4%	11	+40.3%	+39.0%
Totale	215	37	-7.6%	178	+33.3%	+28.2%

Produttività: +28%

N. di pubblicaz.: +20%

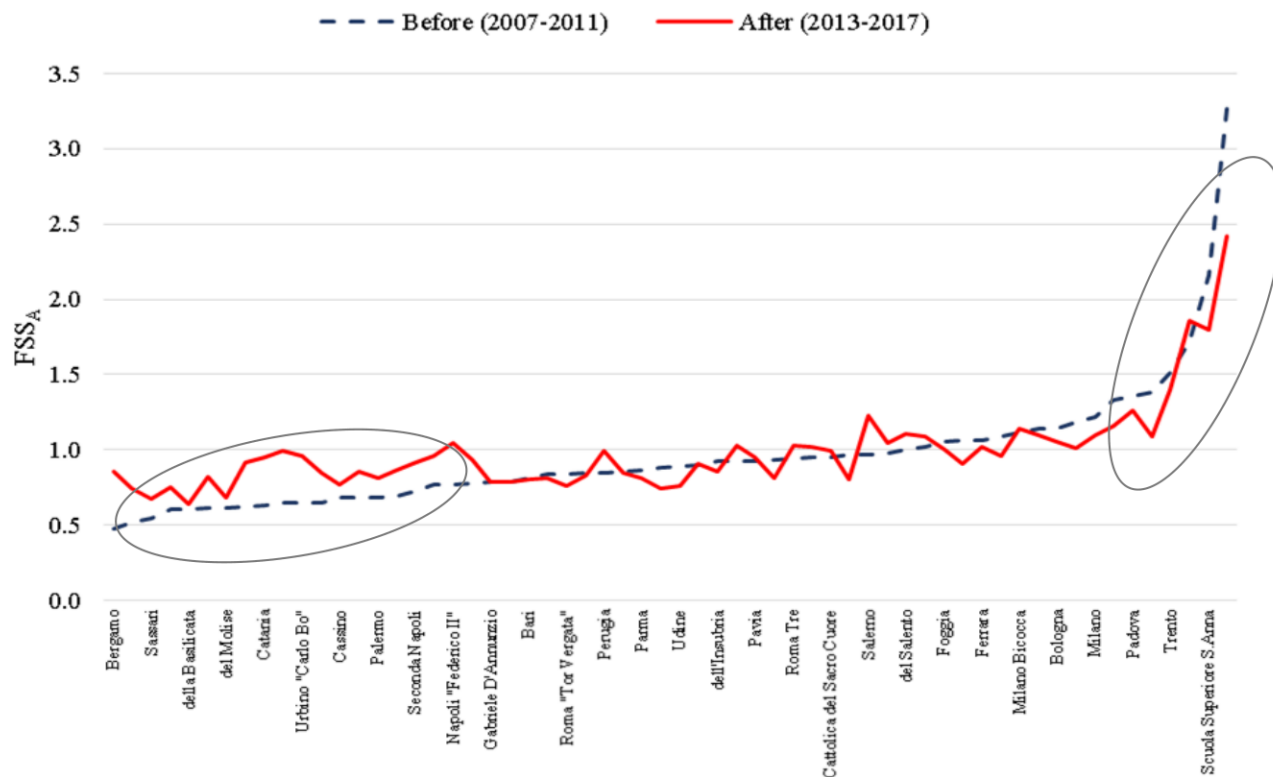
Impatto medio: +8%

* *Fractional Scientific Strength, indicatore di produttività aggregato*

Effetti desiderati e indesiderati della valutazione della ricerca: evidenze empiriche da analisi bibliometriche

PRIN 2017 (NKWYFC): Gli effetti della valutazione sulla ricerca accademica (Roma, 22 giugno 2023)

Risultati - La produttività delle università



Variazioni di performance
2013-2017 (VQR2) vs 2004-2010 (VQR1)

	N. di università	Di cui – migliorano	Di cui – peggiorano	Δ rank relativo medio
Nord	24	5 (21%)	15 (63%)	-14%
Centro	14	8 (57%)	6 (43%)	-2%
Sud	22	14 (64%)	8 (36%)	+4%

Lo studio evidenzia la “convergenza” delle performance “comparate” e, in particolare, l’incremento di produttività delle università del Sud.

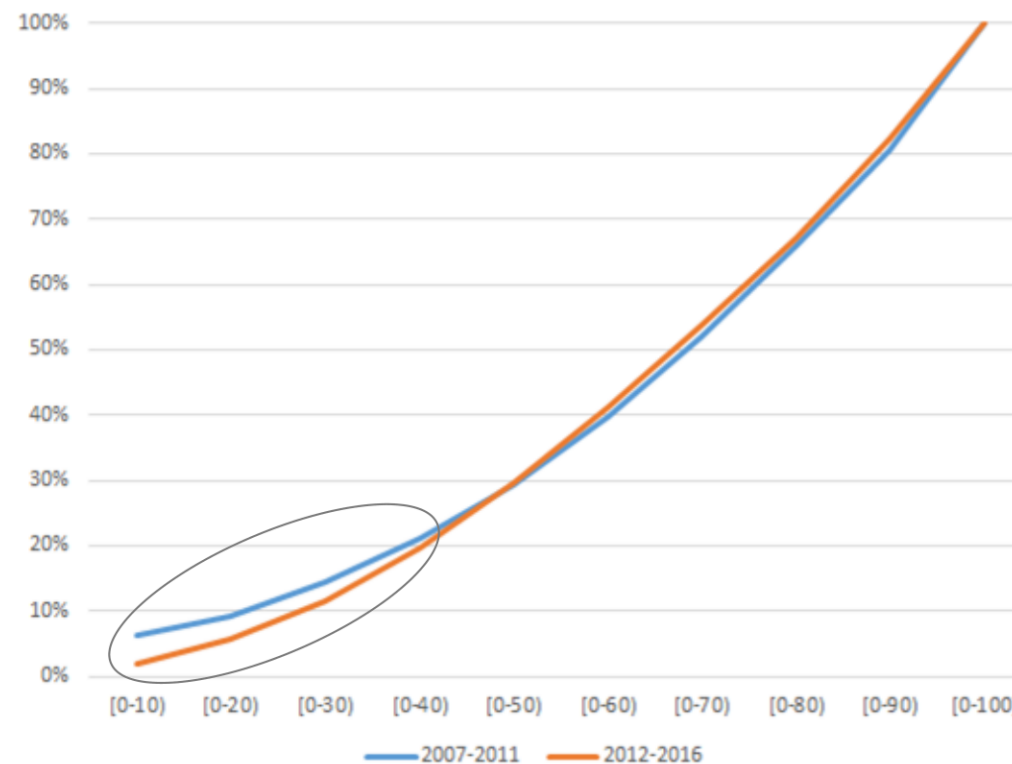
Risultati - Il reclutamento



	2007-2011	2012-2016
Reclutati	3741	7224
Produttività media (percentile)	63	64
Mediana	69	67
100 percentile (top)	3.0%	2.9%
[75;100)	38.2%	37.1%
[50;75)	29.5%	30.5%
[25;50)	17.5%	21.3%
(0;25)	6.6%	7.4%
Improduttivi (score nullo)	5.3%	0.8%

La % di improduttivi (FSS* nulla) è diminuita. Quella di reclutati sopra la mediana nazionale è invariata.

Distribuzione cumulativa di performance (percentile di FSS)



Risultati - Il reclutamento



I «tagli» hanno incentivato le promozioni interne, cresciute a discapito di nuovi ingressi e trasferimenti.

Non si evidenzia alcun miglioramento nella capacità di reclutamento delle università.

	2007-2011		2012-2016	
	%	Perc. medio	%	Perc. medio
Promozioni «esterne»	2.2	73.2	3.9	66.7 ↓
Promozioni «interne»	50.9	65.6	83.4 ↑	63.5
Trasferimenti	8.2	61.4	3.6 ↓	62.4
Nuovi ingressi	38.7	59.3	9.1 ↓	65.9 ↑
Media		63.0		63.8

	Tutti i reclutamenti		Solo promozioni «interne»	
	08-12	13-17	08-12	13-17
R1.1	5.87	2.66 ↓	5.35	2.21 ↓
R1.2	56.7%	64.0%	58.9%	63.7%
R2.1	1.62	1.45 ↓	1.71	1.39 ↓
R2.2	47.8%	51.5%	50.7%	50.7%

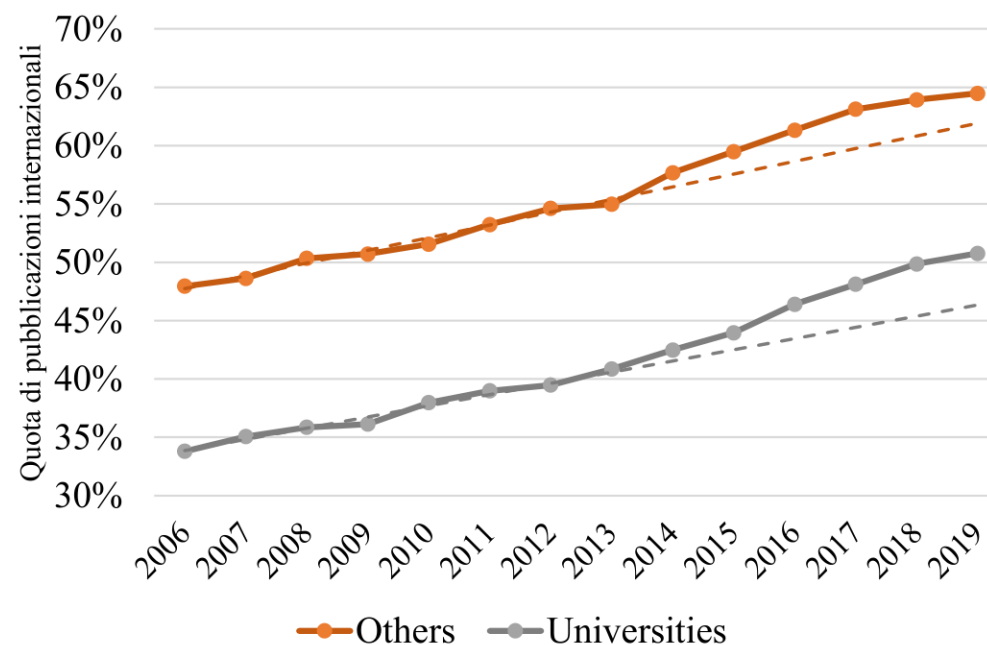
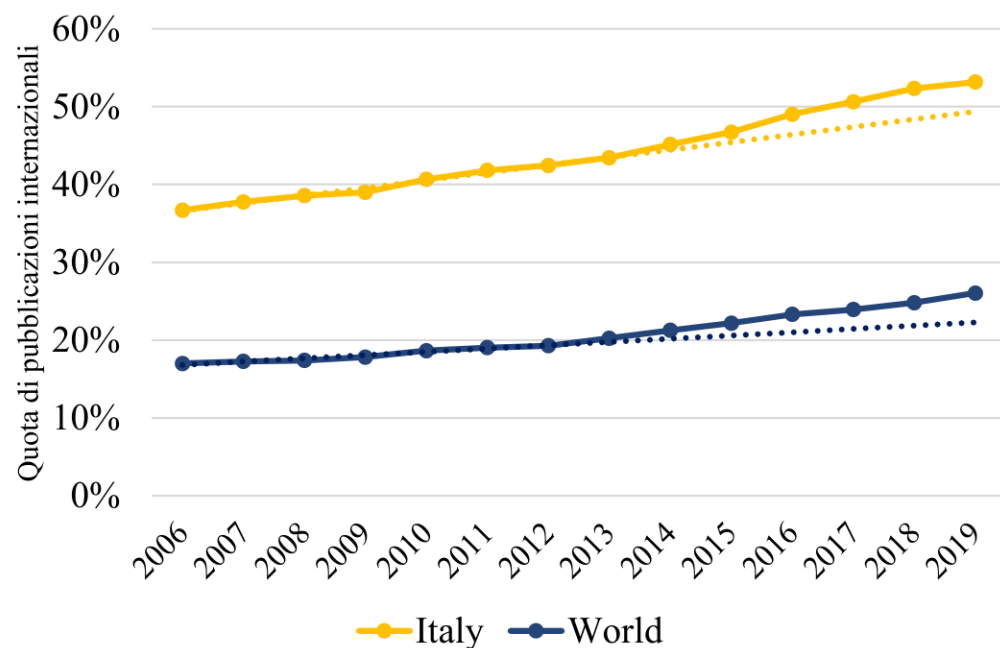
R1.1: Rapporto tra la produttività dei reclutati e quella media dei colleghi già in ruolo nell'università reclutante

R1.2: % di reclutati con produttività maggiore della media dei colleghi già in ruolo nell'università reclutante

R2.1: Rapporto tra la produttività dei reclutati e quella media di tutti i colleghi italiani già in ruolo

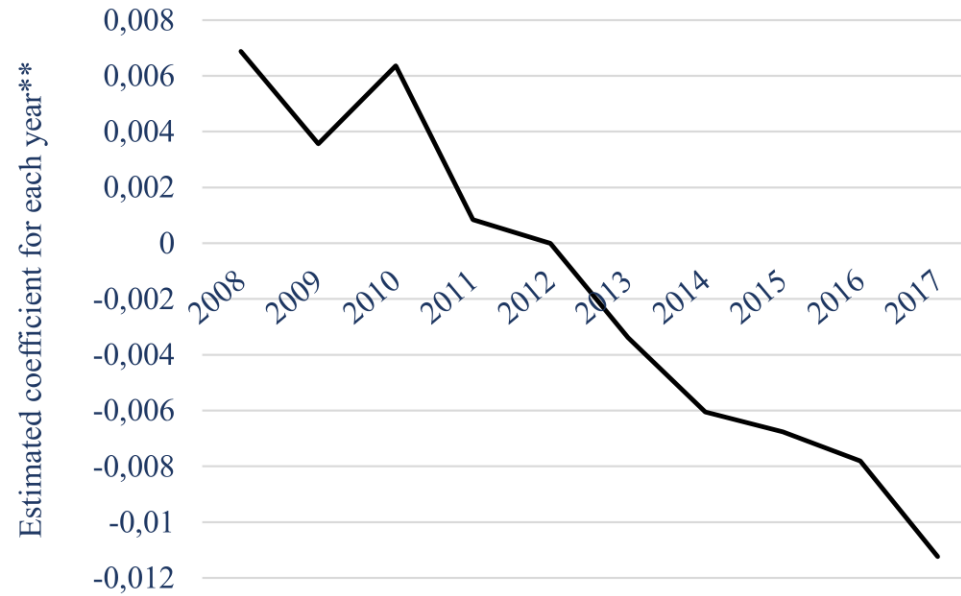
R2.2: % di reclutati con produttività maggiore della media dei colleghi italiani già in ruolo

Risultati - Le collaborazioni internazionali



- La crescita della quota di pubblicazioni italiane frutto di collaborazione internazionale (dal 37% al 53% nel periodo di osservazione) avviene in un contesto di trend crescente a livello globale.
- Discontinuità a partire dal 2013, evidente però anche per il resto del mondo.
- Le pubblicazioni delle università mostrano un tasso di internazionalizzazione inferiore ma con “discontinuità” più pronunciata.

Risultati - Specializzazione vs diversificazione



** 2012 come riferimento, controllo degli effetti fissi di area e delle caratteristiche personali

Significativa riduzione del livello medio di specializzazione dei docenti dopo l'introduzione dell'ASN, rafforzando un trend comunque presente anche prima.

Risultati – Le autocitazioni



Con l'introduzione dell'ASN, ricercatori e professori associati nei settori bibliometrici, per ottenere l'abilitazione, incrementano i loro tassi di autocitazione.

Autocitazioni per pubblicazione

	2008-2012	2013-2017
Totale (15 oss.)	1.74	2.00 (+15%)
Min: CHIMICA	2.80	3.06 (+9%)
Max: ING. CIVILE	1.07	1.61 (+50%)

Inferenza con controlli

	Effetto
Ordinari	+9.9%
Ricercatori e associati non abilitati	+15.0%
Ricercatori abilitati	+4.2%
Associati abilitati	+3.3%
Overall	+9.5%

Ordinari e, soprattutto, gli abilitati aumentano le autocitazioni in misura minore.

Dei quasi 15,000 soggetti indagati, il 41% mostra un tasso invariato. Il restante 59% presenta un'elevata variabilità, ma con pochissimi outlier: solo 22 soggetti aumentano in modo "anomalo" le loro autocitazioni.

L'effetto stimato (+9,5%, per altro "concentrato") è talmente modesto da poter essere trascurato.

Risultati – Le «gift authorship»

La VQR1 (2004-2010) prevedeva penalizzazioni per i «prodotti mancanti». Dopo la VQR, le università con alto tasso di improduttivi potrebbero aver incentivato «gift authorships», per «convertirli» fittiziamente. Tale prassi potrebbe essere associata al carattere segnatamente «intramurario» delle pubblicazioni di tali soggetti.

	“Nuovi produttivi”	Campione di controllo
Osservazioni	676	990
Pubb. medie per professore	4.1	13.9
Di cui internazionali (%)	15.3	30.6
Esclusivamente intramurarie (%)	32.8	25.7
Ad un solo autore (%)	2.6	1.5
Numero di co-autori	5.9	6.6

I dati sembrano apparentemente confermare l'ipotesi (32,8% vs 25,7%), ma altri (collaborazioni internazionali, pubb. ad un solo autore, etc.) sembrano smentirla.

Fra le università più efficaci nel “convertire” gli improduttivi, solo una mostra una quota di pubblicazioni intramurarie a firma dei nuovi produttivi maggiore di quella dei colleghi nel medesimo SSD sempre produttivi.

- Contrariamente a quanto indicato da parte della letteratura, la produttività di ricerca del sistema accademico italiano mostra un trend molto positivo, soprattutto in termini di numero di pubblicazioni, ma anche di impatto medio.
- La VQR ha certamente innescato un processo di convergenza della produttività di ricerca delle università: in termini relativi, quelle del sud sono cresciute di più di quelle del nord, malgrado la penalizzazione in termini di quote premiali di FFO.
- Per il reclutamento, l'ASN ha sì introdotto uno sbarramento alle candidature di soggetti improduttivi o scarsamente produttivi, ma non ha innalzato il livello di qualità medio complessivo. Le università non sembrano aver migliorato i processi di reclutamento: i vincoli di bilancio hanno incentivato le promozioni «interne» e vanificato gli effetti della ASN.

- Il livello di internazionalizzazione della ricerca italiana mostra un trend decisamente positivo, ma in linea con quello che accade a livello globale. La prima VQR pare non abbia avuto un impatto in tal senso (le successive hanno rimosso tale incentivo).
- L'ASN ha indotto una significativa riduzione del livello medio di specializzazione dell'agenda di ricerca dei docenti, rafforzando un trend già in atto.
- L'introduzione di indicatori e soglie citazionali nell'ASN, per i settori bibliometrici, ha determinato un aumento delle autocitazioni molto limitato e concentrato.
- La presenza di penalizzazioni per le università con molti «improduttivi» nella prima VQR, non sembra averle indotte ad adottare comportamenti opportunistici in termini di «gift-authoship».

- Abramo, G., & D'Angelo, C.A. (2023). Warnings of declining research productivity: does Italy buck the trend? *Working paper*.
- Abramo, G., D'Angelo, C.A. & Di Costa, F. (2023). The effect of bibliometric research performance assessment on the specialization vs diversification strategies of scientists. *19th ISSI Conference*, Bloomington, Indiana-US.
- Abramo, G., & D'Angelo, C.A. (2022). The impact of Italian performance-based research funding systems on the intensity of international research collaboration. *Research Evaluation*. DOI: 10.1093/reseval/rvac026
- Abramo, G., & D'Angelo, C.A. (2021). The different responses of Italian universities to introduction of performance-based research funding. *Research Evaluation*, 30(4), 514-528. DOI: 10.1093/reseval/rvab022.
- Abramo, G., D'Angelo, C.A., & Grilli, L. (2021). The effects of citation-based research evaluation schemes on self-citation behaviour. *Journal of Informetrics*, 15(4), 101204. DOI: 10.1016/j.joi.2021.101204.
- Abramo, G., & D'Angelo, C.A. (2020). Were the Italian policy reforms to contrast favoritism and foster effectiveness in faculty recruitment successful? *Science and Public Policy*, 47(5), 604-615. DOI: 10.1093/scipol/scaa048
- Abramo, G., D'Angelo, C.A., & Di Costa, F. (2019). When research assessment exercises leave room for opportunistic behavior by the subjects under evaluation. *Journal of Informetrics*, 13(3), 830-840. DOI: 10.1016/j.joi.2019.07.006

**Effetti desiderati e indesiderati della valutazione della
ricerca: evidenze empiriche da analisi bibliometriche**

PRIN 2017 (NKWYFC): Gli effetti della valutazione sulla ricerca accademica

Grazie dell'attenzione!

Ciriaco Andrea D'Angelo¹, Giovanni Abramo²
Laboratorio di Studi sulla Valutazione della Ricerca

¹ Dipartimento di Ingegneria dell'Impresa, Università degli Studi di Roma "Tor Vergata"
dangelo@dii.uniroma2.it

² Istituto di Analisi dei Sistemi e Informatica (IASI-CNR)
giovanni.abramo@iasi.cnr.it



Ricerca accademica e valutazione: quali implicazioni di policy?

Emanuela Reale - CNR IRCrES

IPOTESI

La valutazione configura un'idea di qualità diversa da quella delle comunità epistemiche e produce effetti differenziati influenzati dal contesto istituzionale e geografico, dai settori disciplinari, dallo stadio di carriera di ricerca (inizio carriera o senior) e da questioni di genere.

DOMANDE DI RICERCA

- Quali cambiamenti può generare la valutazione esterna della ricerca?
- Quale è la probabilità che tali effetti siano trasformativi?

- Come le differenti discipline interpretano e contribuiscono a definire i metodi di valutazione della ricerca

- Sono stati raggiunti gli obiettivi della valutazione?
- Si evidenziano effetti «indesiderati» dovuti a criteri non adeguati?
- Si rilevano comportamenti opportunistici?
- Si rilevano altri effetti degni di nota a livello di sistema?

DIFFERENTI APPROCCI DATI E METODI



Per approssimare il nesso di causalità tra valutazione e i suoi effetti:

- Approccio contributivo della valutazione rispetto ad altri fattori causativi
- Analisi longitudinale sui contributi delle comunità scientifiche interessate
- Approcci di analisi controfattuale

- Indagine nazionale su un campione probabilistico della popolazione docente
- Survey specifiche su ambiti disciplinari
- Analisi bibliometriche
- Interviste, storie di vita

Unità di osservazione: accademici delle università italiane

Differenti criticità

ma significativi punti di convergenza nei risultati osservati

Possibilità di fornire evidenze robuste che possano promuovere processi di riflessione sugli effetti della valutazione

Aumento generalizzato (di sistema, di università, individuale) in tutti i settori disciplinari

- Confermata anche dalle rilevazioni collegate alle percezioni individuali e alle narrative
- Ma una quota rilevante di soggetti non ritiene che la VQR abbia avuto un ruolo rilevante su questo cambiamento rispetto ad altri fattori (diversa la percezione sull'ASN)
- Più produttività non significa necessariamente più impatto scientifico

L'aumento è particolarmente accentuato

- Nelle università del Sud (maggiore margine di miglioramento) producendo quindi un fenomeno di convergenza nella produttività delle università italiane
- Tra i giovani ricercatori (più sensibili al publish or perish)

Diminuisce il numero di soggetti improduttivi

Risultato netto sul **cambiamento del modo con cui gli accademici delle scienze sociali e umane pubblicano**

- L'influenza che esercita la valutazione è più forte di altri fattori in particolare nelle SSH
- Questo cambiamento non è percepito come migliorativo della qualità della ricerca specialmente nei giovani ricercatori

In generale **la valutazione non è percepita come un fattore che ha effetti rilevanti** sulla qualità della ricerca

- Ma due settori disciplinari (Fisica/Economia) mostrano percezioni significative e di segno opposto
- L'ASN ha effetti significativi su giovani ricercatori e associati

L'aumento della produttività **non migliora la qualità del reclutamento** delle università

Multidimensionalità del concetto di qualità della ricerca e sua dipendenza dalle caratteristiche delle comunità epistemiche

- Limiti delle misurazioni basate su caratteristiche estrinseche

Aumento delle pubblicazioni su **riviste classificate in Fascia A**

- Disallineamento tra produttività e interesse (citazioni)

Ruolo delle case editrici

- Costruzione di logiche di accettazione/rigetto che influiscono sulle discipline (mainstream e interdisciplinarietà)

Ruolo degli Editor

- Custodi della disciplina
- Arbitri della carriera

AGENDA DI RICERCA E VALUTAZIONE



Riconoscimento da parte dei pari



Ricerca come beneficio per la società



Ricerca come passione personale



Bandi



Collaborazione scientifica

Riduzione del livello medio di **specializzazione** dell'agenda di ricerca (ASN) con evidenze significative rispetto al comportamento di giovani ricercatori e associati

- controllo con dati bibliometrici e in combinazione con le narrative

L'ASN fa emergere problemi connessi alle scelte di **ricercatori/ricercatrici** con almeno 1 figlio/figlia

La valutazione favorisce **comportamenti opportunistici** (mainstream vs innovazione) ma non aumenta significativamente le autocitazioni

DUNQUE



La valutazione ha raggiunto gli obiettivi per i quali è stata disegnata?

Quali sono le criticità più importanti tra quelle rilevate?

Come possiamo affrontarle?

Quanto è sostenibile questo sistema di valutazione in relazione ai cambiamenti in atto nella produzione di conoscenza scientifica?



Grazie dell'attenzione!