Identifying potential pitfalls in the quantitative appraisal system for scientific careers

UNIVERSITY of HOUSTON

Experiencing Ethics Seminar

December 3, 2012

Institutions Markets Technologies IMT INSTITUTE FOR ADVANCED STUDIES LUCCA

Alexander M. Petersen

IMT Institute for Advanced Studies, Lucca Italy

Ethics in the appraisal of Scientific Careers

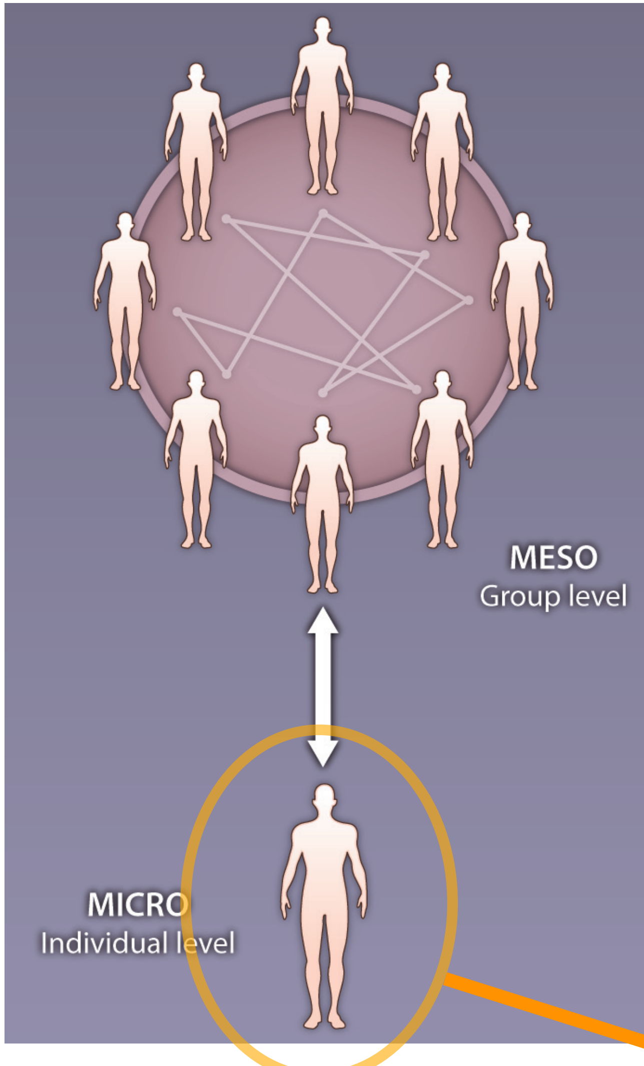
- **Competition ("fairness"):**
  - strategizing / extreme behavior, e.g. scientific fraud
  - CED (cognitive enhancing drugs)
  - free-riding + "tragedy of the commons"
- **Careers:** predicting future career achievement using incomplete information and poorly understood/ designed achievement measures
- **Funding:**
  - financial incentives & who should subsidize early career risk
  - how to attribute / appraise / reward achievement, especially in the case of extremely large team projects

Outline

- **Science:** growth and emergent complexity
- **Citations:** proxy measure for scientific impact:
  - comparison across discipline
  - comparison across time / role of output inflation (baseball analogy!)
- **Careers:** measured by appraisal of the publication/ impact portfolio:
  - h-index + empirical regularities + criticisms
  - career longevity
  - Competition: cognizant enhancing drugs (CED) + Is academia becoming more like a professional sport?
  - **New results:** complexity of career predictability and co-evolution of the scientific production function

Evolution of Science

An "atomic" view of Science as a Multi-level system



K. Börner, et al. A multi-level systems perspective for the science of team science. Sci. Transl. Med. 2, 49cm24 (2010).

Interactions mediated by social "forces":

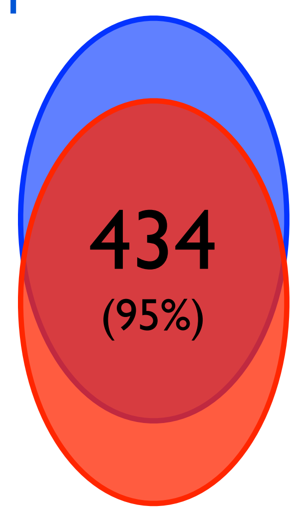
- Collaboration (attractive)
- Competition (repulsive)
- Knowledge (an "exchange particle")

Watson-Crick strategy:

* **Michael Stuart Brown**
* **Joseph L. Goldstein**

Recipients of the 1985 Nobel Prize in Physiology or Medicine for describing the regulation of cholesterol metabolism.

451 publications



458 publications

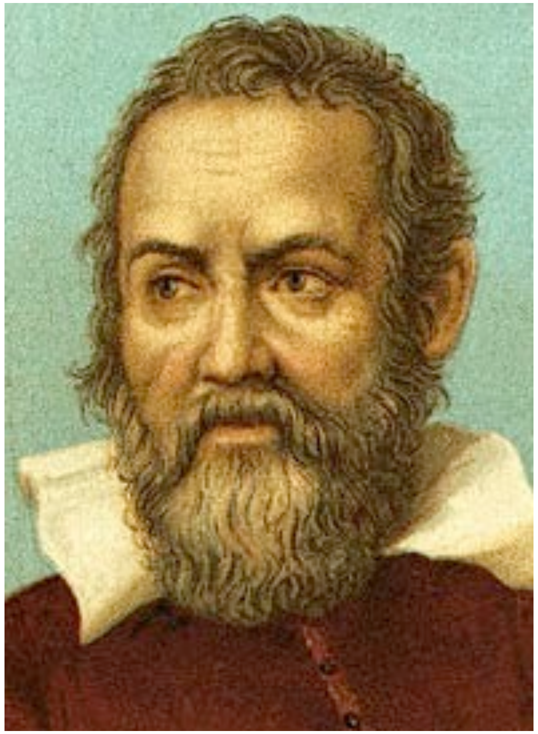
Solo-artist strategy:

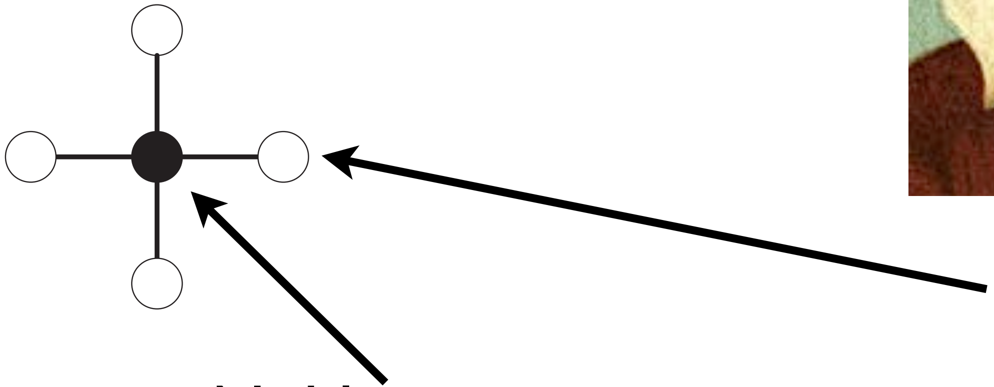
* **Marilyn Kozak** (also cell biologist)

$N = 70$, $N_{solo} = 59$

Evolution of Science: "In the beginning..."

[Social networks in science:]()
serve as the backbone for reputation signaling used to overcome the asymmetric information problem
⇒ reputation tournaments

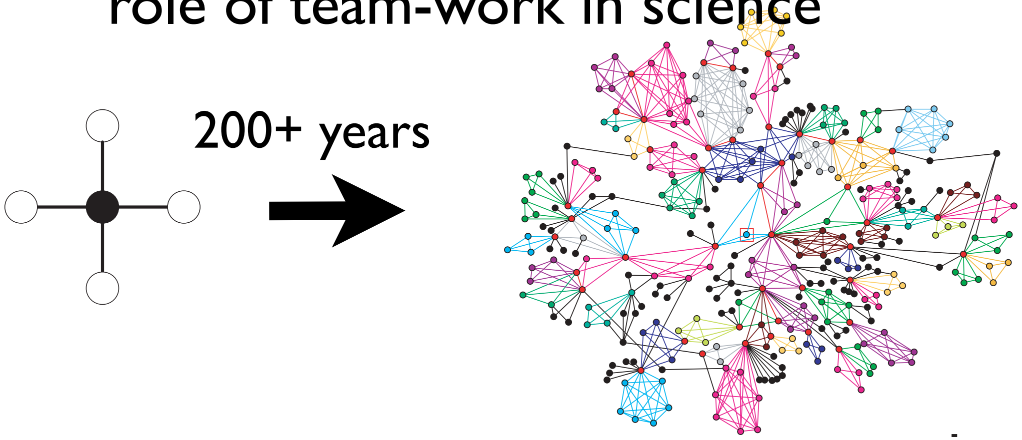




Paul A. David. *The Historical Origins of 'Open Science': An essay on patronage, reputation, and common agency contracting in the scientific revolution.* Capitalism and Society 3(2): Article 5 (2008).

Emerging trends in Science

- emergence of small-world collaboration networks with the increasing role of team-work in science



G. Palla, A.-L. Barabasi, T. Vicsek. Quantifying social group evolution. Nature 446, 664-667 (2007)

S. Wuchty, B. F. Jones, B. Uzzi. The increasing dominance of teams in production of knowledge. Science 316, 1036-9 (2007)

- organizational shifts in the business structure of research universities

- shifts away from tenure towards shorter-term contracts + bottle neck in the number of tenure-track positions available

Chait RP, ed. The Questions of Tenure. (Harvard University Press, Cambridge USA, 2002).

- redefining the role of teaching -vs- research faculty

- shifts in the competitive aspects of science, universities, and scientists: reputation tournaments in omnipresent competition arenas

… articles on the academic labor market

280 | NATURE | VOL 472 | 21 APRIL 2011

RETHINKING PHDS

Fix it, overhaul it or skip it completely — institutions and individuals are taking innovative approaches to postgraduate science training.

Chait RP, ed. The Questions of Tenure. (Harvard University Press, Cambridge USA, 2002).

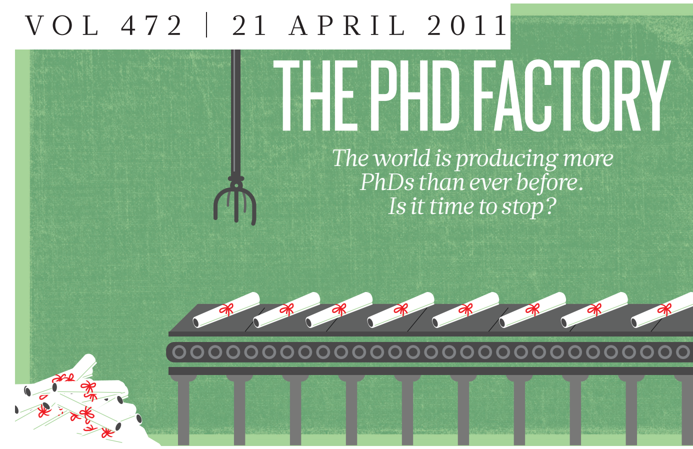
ACADEMIA

The changing face of tenure

Although still highly desirable, tenure is not as prevalent as it was in some places — and that may not be a bad thing.

NOVEMBER 2010 | VOL 468 | NATURE | 123

NATURE | VOL 472 | 21 APRIL 2011



Issues & Perspectives

Academia's Crooked Money Trail

Science Careers From the journal Science

10.1126/science.caredit.a1200001

Undergraduates also carry an increasing share of the load, she adds: Their tuition, often paid with student loans, rises as more funds go to research. Their teachers, meanwhile, increasingly are cut-rate adjuncts rather than the famous professors the recruiting brochures boast about.

Education bubble!

UNITED STATES

Mid-career crunch

Some senior scientists feel neglected by the National Institutes of Health's grant formula.

17 MARCH 2011 | VOL 471 | NATURE | 399

… on inefficiencies…

Clean up the waste

Fixing inefficiencies at academic institutions will strengthen — not jeopardize — teaching and research, says **Thomas Marty**.

5 APRIL 2012 | VOL 484 | NATURE | 27

NATURE | NEWS

Funding uncertainty strands Spain's young scientists

Delayed decisions disrupt international collaborations.

Lucas Laursen

06 March 2012

“Consider the financial calculations that encourage universities to hire a series of postdocs rather than staff scientists. Postdocs earn around half to two-thirds of a staff scientist’s salary….. and are temporary, so can be let go when budgets decline. But, in reality, postdocs are not cheap: substantial resources - both their own and society’s - have been invested in training them. If a postdoc doesn’t get a [permanent] research job, taxpayers do not get a return on their investment. Neither does the postdoc….”

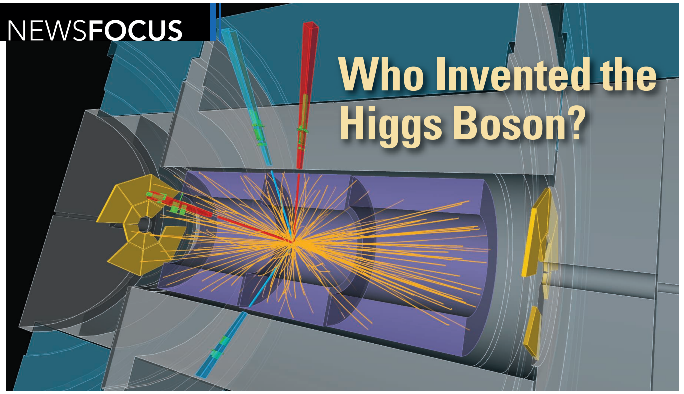
Perverse incentives

Counterproductive financial incentives divert time and resources from the scientific enterprise. We should spend the money more wisely, says **Paula Stephan**.

5 APRIL 2012 | VOL 484 | NATURE | 29

… increasing team size & the reward system in science

NEWS**FOCUS**



Five living theorists have claims to having dreamed up the most famous subatomic particle in physics. But what did they really do?

Kingdom. Others question whether the advance was a big enough step beyond previous work to merit science's biggest prize.

14 SEPTEMBER 2012 VOL 337 **SCIENCE** www.sciencemag.org

"50-way tie for the Nobel Prize"

www.sciencemag.org **SCIENCE** VOL 336 6 APRIL 2012

Published by AAAS

9 DECEMBER 2011 VOL 334 **SCIENCE**

CITATION IMPACT

Saudi Universities Offer Cash In Exchange for Academic Prestige

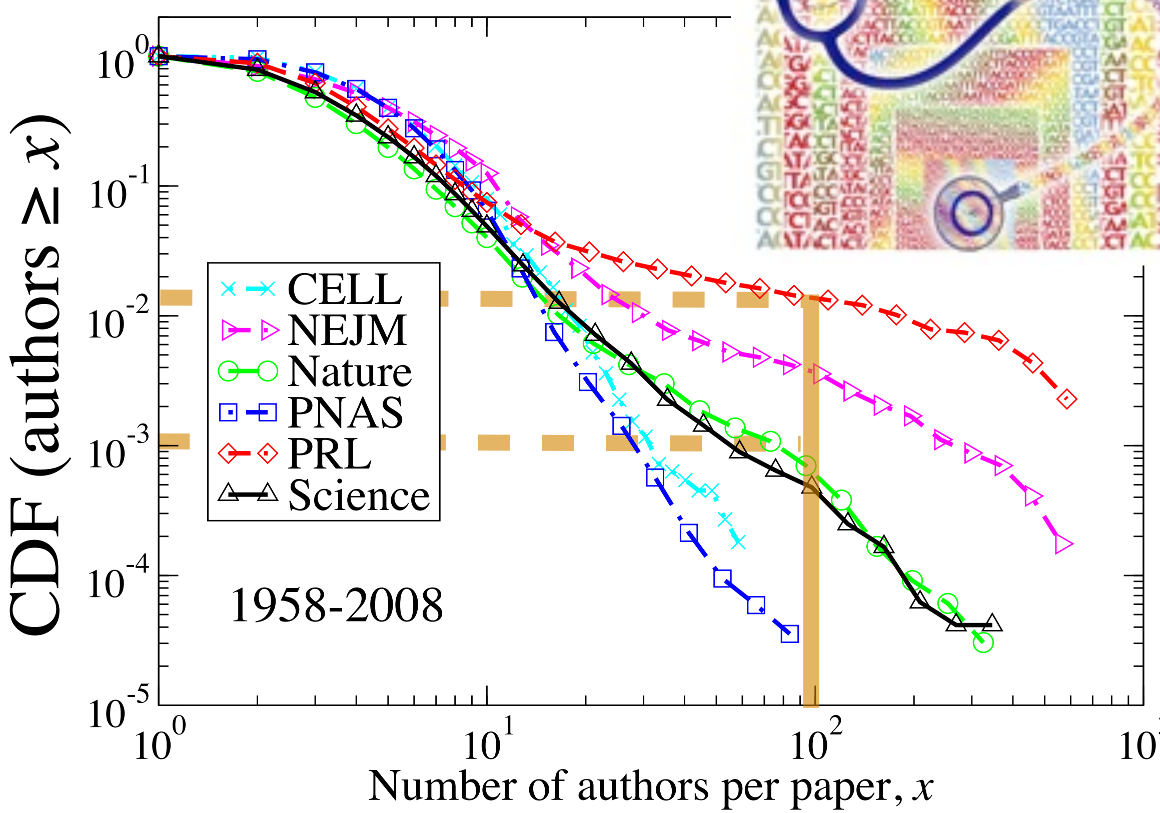
Two Saudi institutions are aggressively acquiring the affiliations of overseas scientists with an eye to gaining visibility in research journals

genome.gov

National Human Genome Research Institute

National Institutes of Health

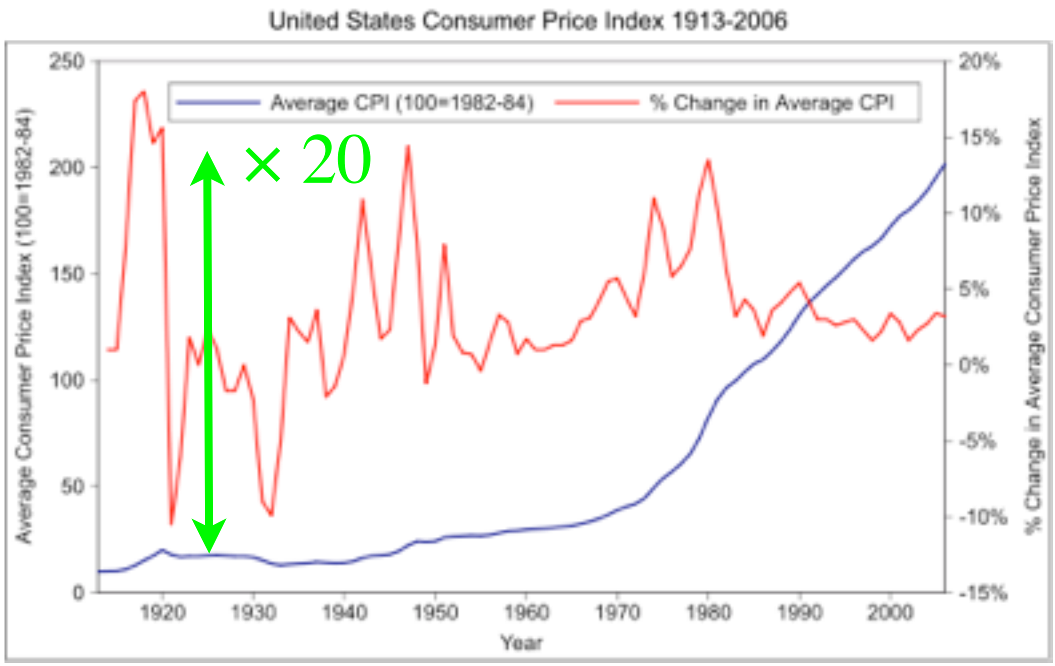
ENCODE



||

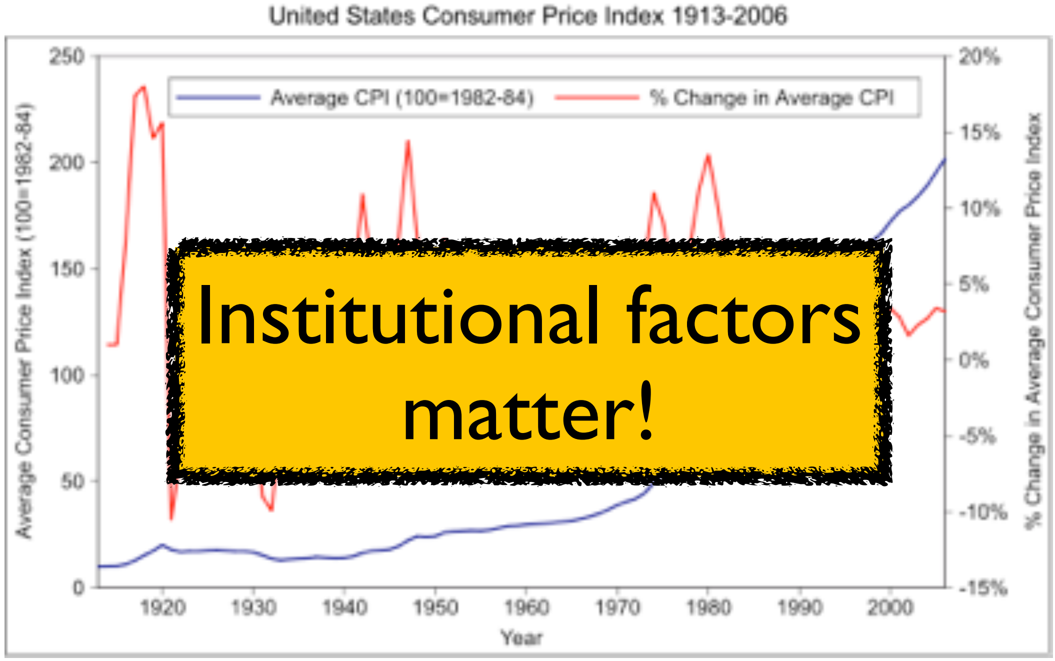
.... a short baseball interlude...

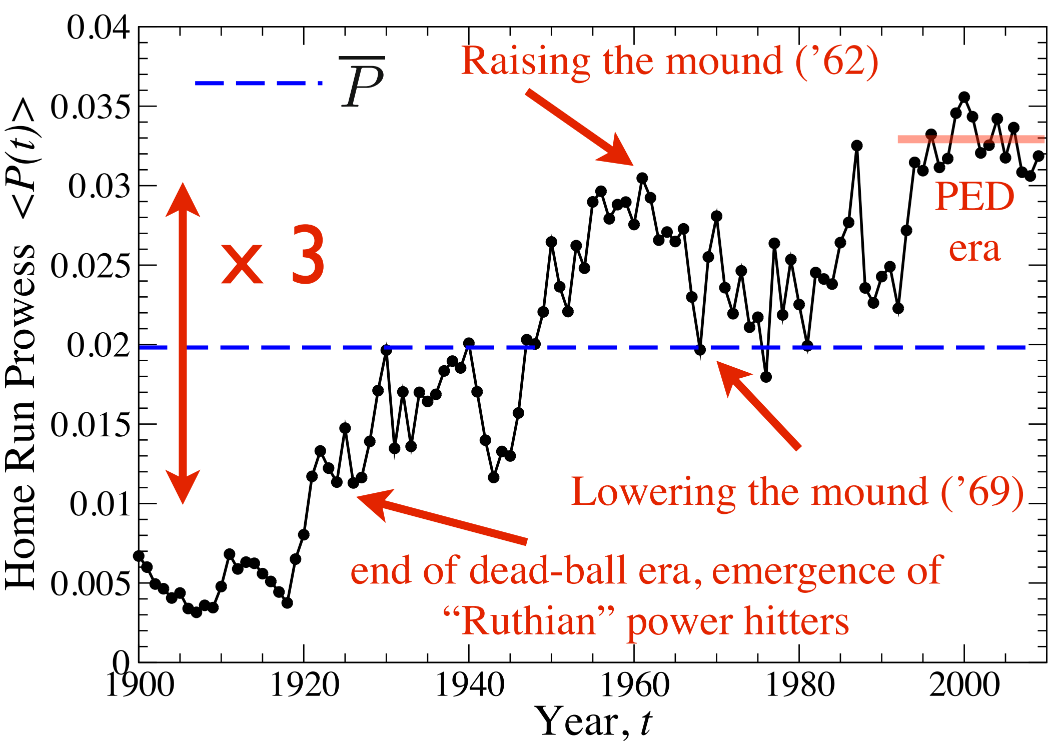
Accounting for Inflation



Just as the price of a candy bar has increased by a factor of ~ 20 over the last 100 years (roughly 3% inflation rate),

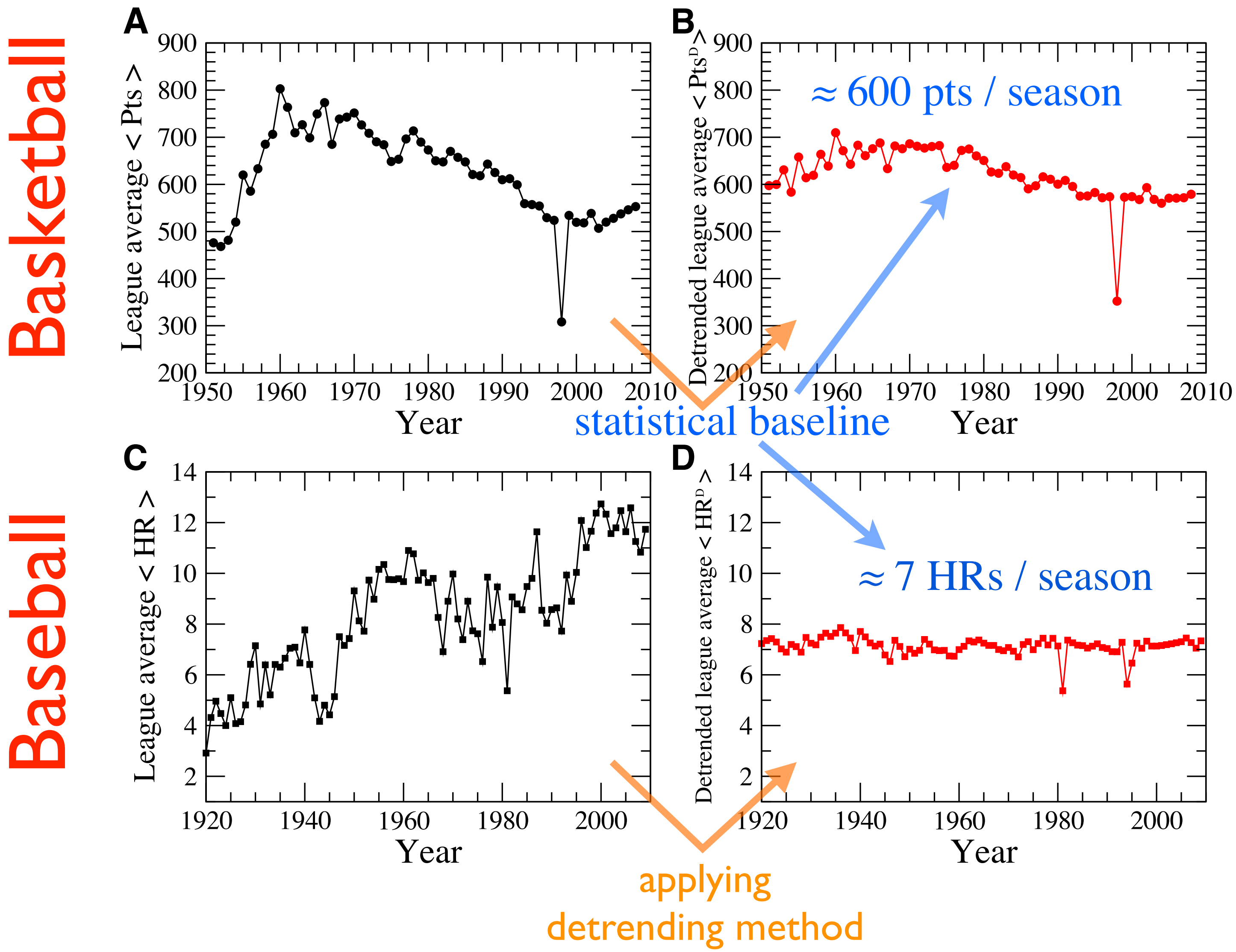
Accounting for Inflation





Just as the price of a candy bar has increased by a factor of ~ 20 over the last 100 years (roughly 3% inflation rate), the home run hitting ability of players has also increased by a significant factor over the same period

Accounting for socio-technological factors that underly achievement



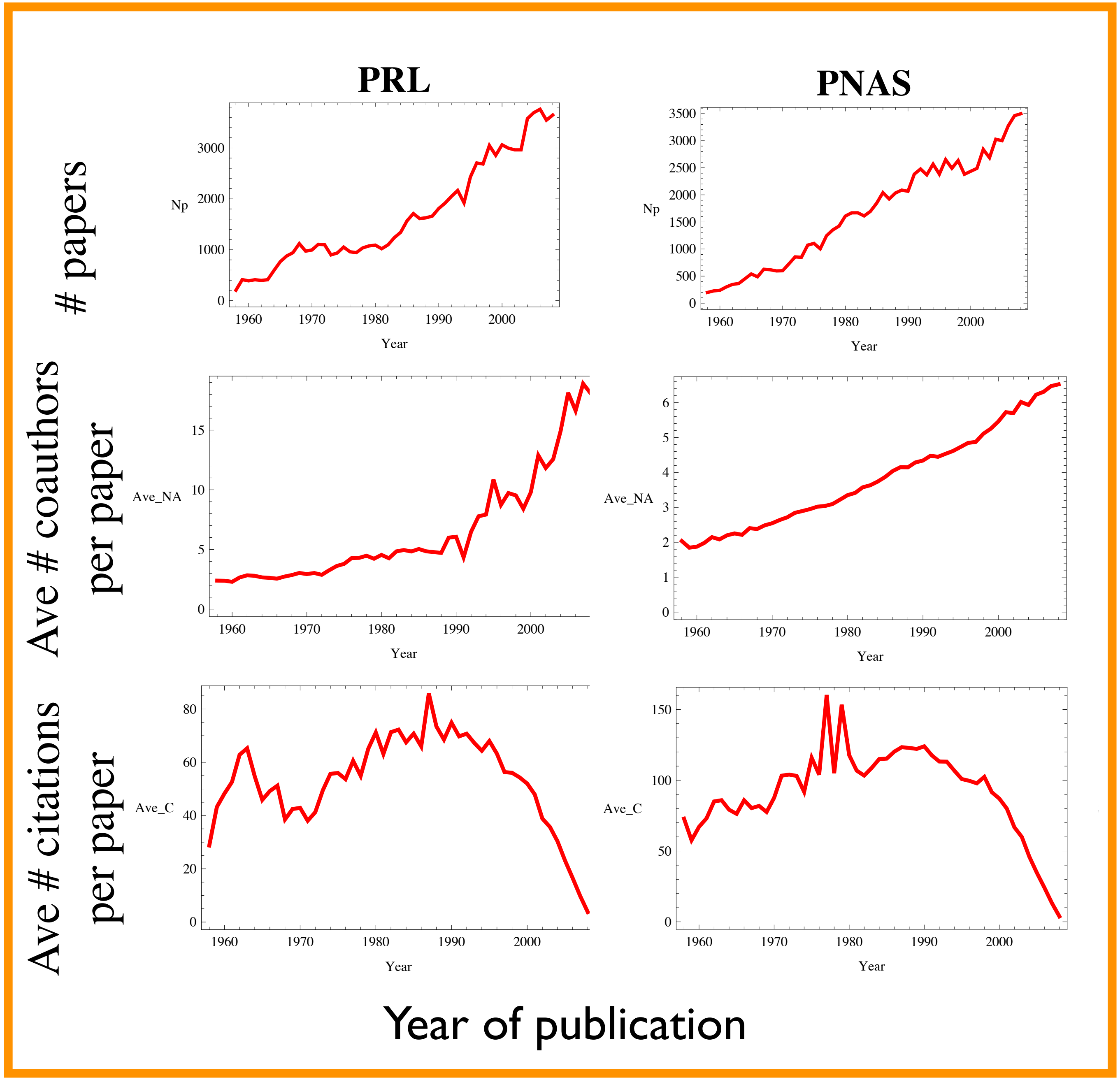
Quantitative measures for success are important for comparing both individual and group accomplishments, often achieved in different time periods.

However, the evolutionary nature of competition results in a non-stationary rate of success, that makes comparing accomplishments across time statistically biased.

A. M. Petersen, O. Penner, H. E. Stanley.
Methods for detrending success metrics to account for inflationary and deflationary factors
Eur. Phys. J. B 79, 67-78 (2011).
Pre-print title: Detrending career statistics in professional Baseball: accounting for the Steroids Era and beyond

Fighting output inflation in Science:

towards bibliometrics that account for time-dependent productivity/impact factors

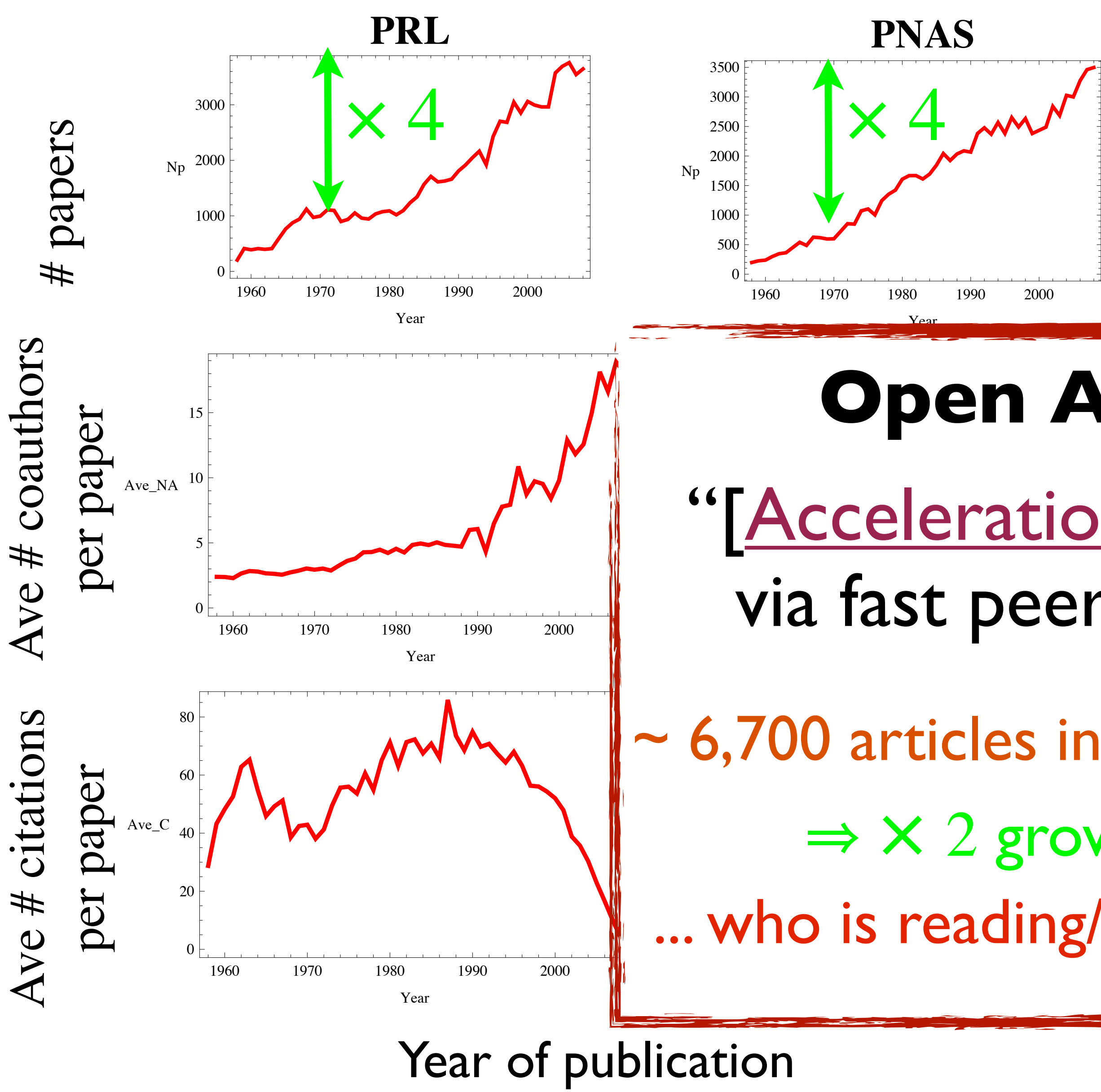


productivity inflation and population growth

growth of team science

time-dependent citation rate depends strongly on the (sub) discipline, the era, and the age of the paper

Fighting output inflation in Science: towards bibliometrics that account for time-dependent productivity/impact factors



productivity inflation and population growth

Open Access Journals

"[Acceleration of scientific progress via fast peer-review/publication]"

PLoS One:

~ 6,700 articles in 2010 and ~ 14,000 in 2011

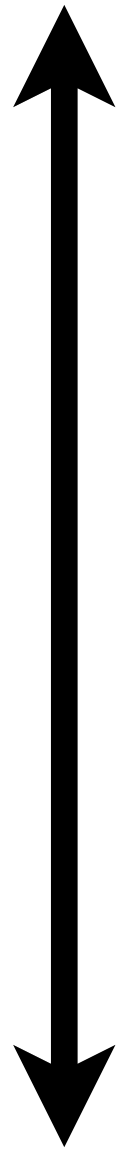
⇒ × 2 growth in one year alone!

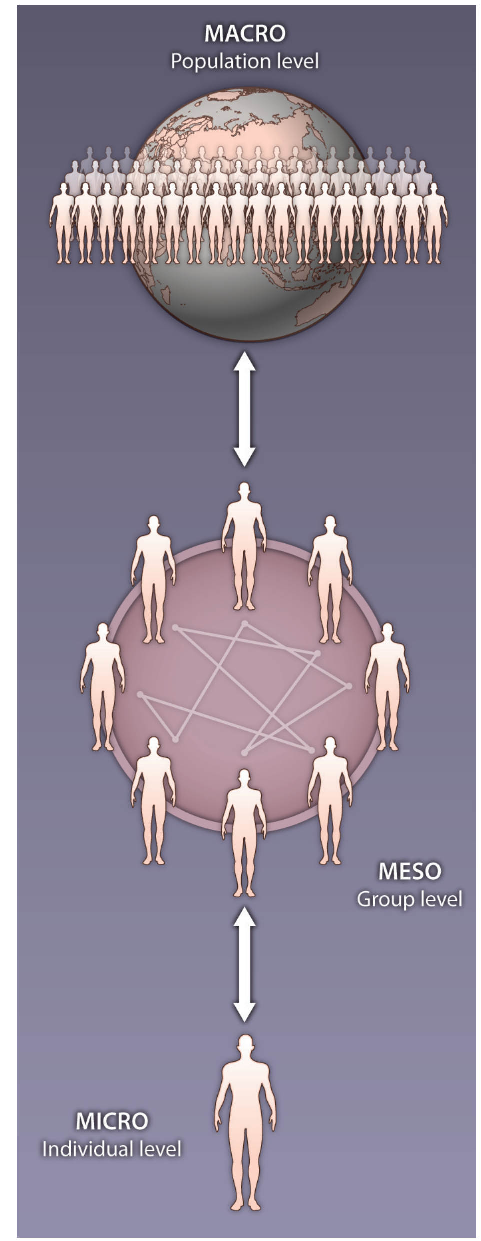
... who is reading/refereeing all these papers??

Question: how to measure scientific output and impact at various scales while accounting for systemic heterogeneity

K. Börner, et al. A multi-level systems perspective for the science of team science. Sci. Transl. Med. 2, 49cm24 (2010).

* Country
* Institution
* Lab / Team
* Individual
* Paper





Quantifying impact and productivity in science

"Math-letes"

|||

Using "big-data" to better understand academic careers

Phys. Rev. Lett. 42, 673–676 (1979)

Scaling Theory of Localization: Absence of Quantum Diffusion in Two Dimensions

Abstract | References | Citing Articles (2,099) | Page

Download: PDF (622 kB) Buy this article **Export:** BibTeX or EndNote (RIS)

E. Abrahams
Serin Physics Laboratory, Rutgers University, Piscataway, New Jersey 08854

P. W. Anderson*, D. C. Licciardello, and T. V. Ramakrishnan†
Joseph Henry Laboratories of Physics, Princeton University, Princeton, New Jersey 08540

Received 7 December 1978; published in the issue dated 5 March 1979

Arguments are presented that the $T=0$ conductance G of a disordered electronic system depends on its length scale L in a universal manner. Asymptotic forms are obtained for the scaling function $\beta(G)=d\ln G/d\ln L$, valid for both $G \ll G_c \simeq e^2/\hbar$ and $G \gg G_c$. In three dimensions, G_c is an unstable fixed point. In two dimensions, there is no true metallic behavior; the conductance crosses over smoothly from logarithmic or slower to exponential decrease with L.

For example, P.W. Anderson:
n = 64 articles
published in *PRL* over this
51-year period

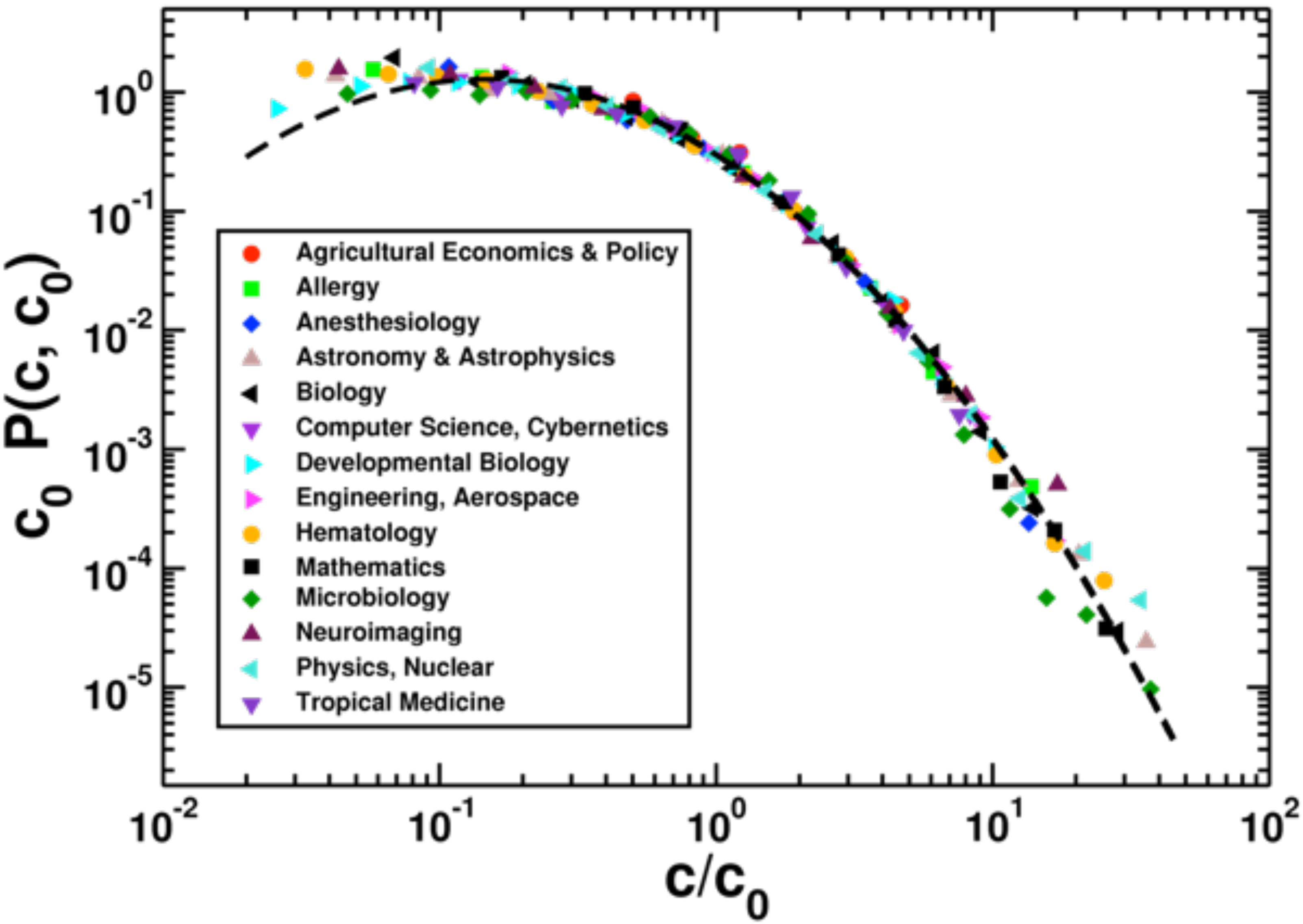
TABLE I. Summary of data set size for each journal. Total number N of unique (but possibly degenerate) name identifications.

Journal	Years	Articles	Authors, N
CELL	1974–2008	53290	31918
NEJM	1958–2008	17088	66834
Nature	1958–2008	65709	130596
PNAS	1958–2008	84520	182761
PRL	1958–2008	85316	112660
Science	1958–2008	48169	109519

"Methods for measuring the citations and productivity of scientists across time and discipline" A. M. Petersen, F. Wang, H. E. Stanley. Phys. Rev. E, ***81*** *(2010) 036114*

a) Paper arena

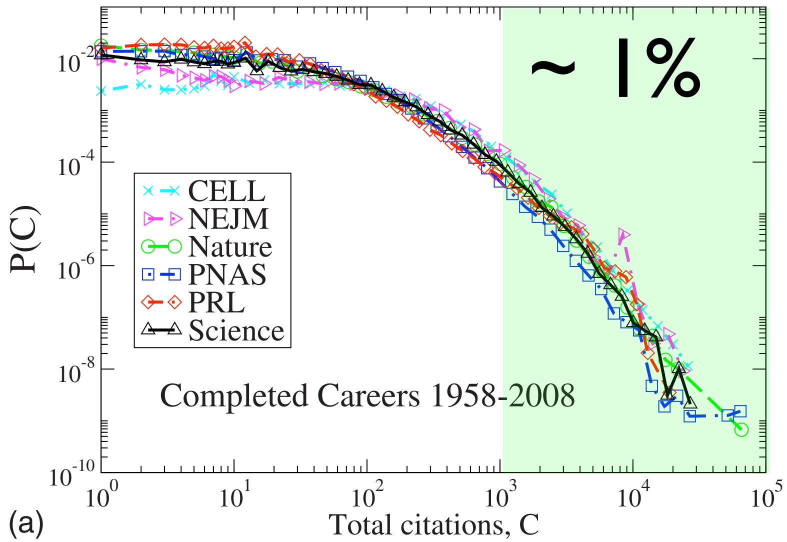
Log-normal citation distribution w/in journal (subfield) accounting for (paper) age cohorts

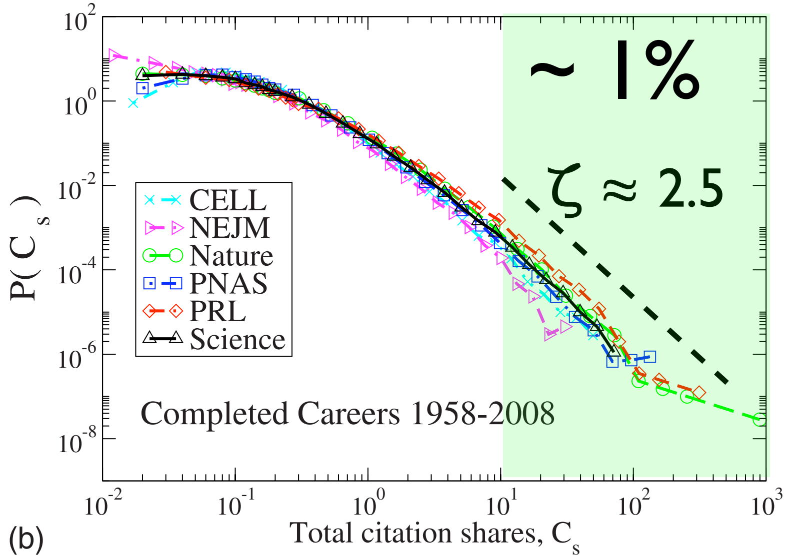


Radicchi F, Fortunato S, Castellano C. Proc Natl Acad Sci USA 105, 17268–17272 (2008)

b) Career arena: "The 1%"

heavy-tailed citation distribution w/in career





Each author has n articles in a given journal j.

Each article i, published in year y, can be quantified by the number of citations c_i it has received at the time of data extraction.

(May, 2009)

Two possible ways to measure citations:

(i) Total citations:

$$C = \sum_{i=1}^{n} c_i .$$

(ii) Total citations ``shares":

$$C_s = \sum_{i=1}^{n} \frac{1}{a_i} \frac{c_i(y)}{\langle c(y) \rangle} .$$

Empirical evidence for the Matthew "rich-get-richer" effect in Science

Gospel of St. Matthew: "For to all those who have, more will be given." Still true 2000 years later!!

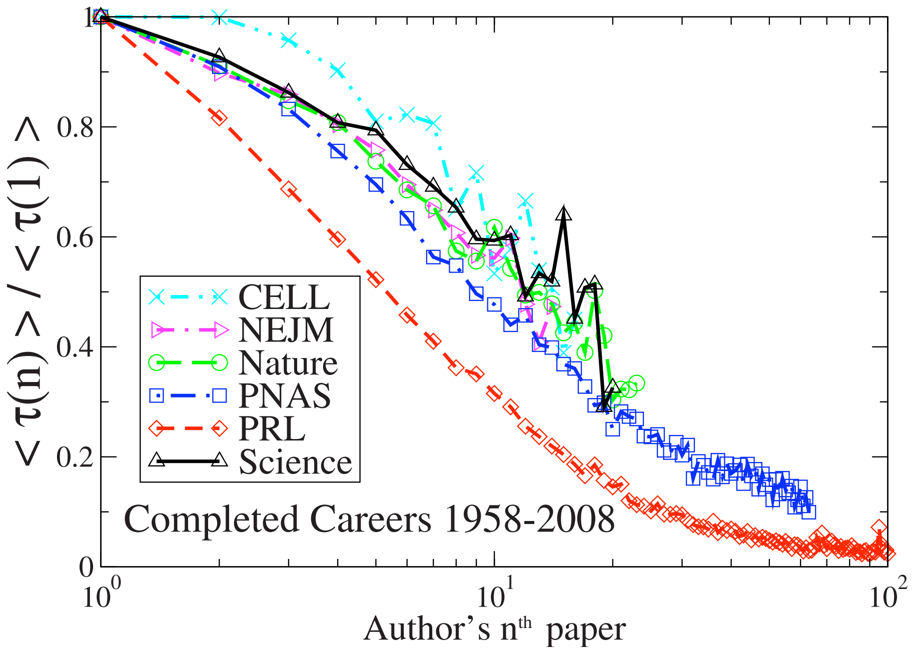
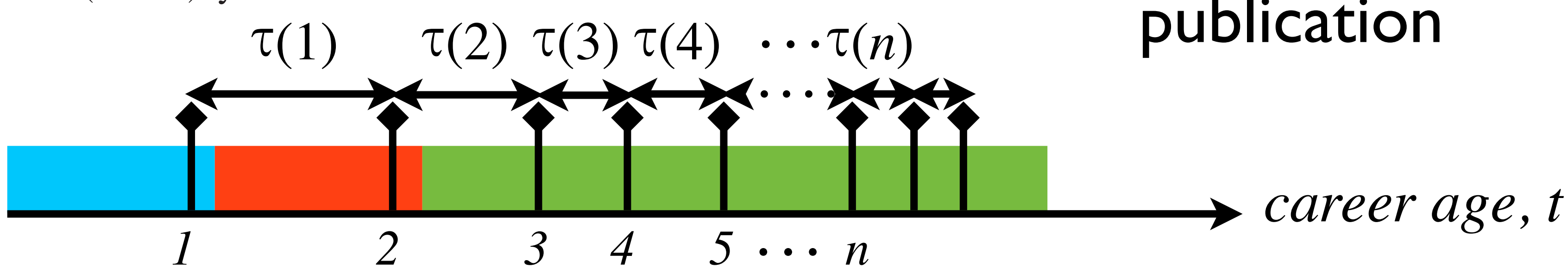


FIG. 7. (Color online) A decreasing waiting time $\tau(n)$ between publications in a given journal suggests that a longer publication career (larger n) facilitates future publications, as predicted by the Matthew effect. We plot $\langle\tau(n)\rangle/\langle\tau(1)\rangle$, the average waiting time $\langle\tau(n)\rangle$ between paper n and paper $n+1$, rescaled by the average waiting time between the first and second publication, $\langle\tau(1)\rangle$. The values of $\langle\tau(1)\rangle$ are 2.2 (*CELL*, *PRL*), 3.0 (*Nature*, *PNAS*, *Science*), and 3.5 (*NEJM*) years.



- For a given journal: the waiting time $\tau(n)$ is the number of years between an author's paper n and paper n+1
- A decreasing $\langle\tau(n)\rangle$ indicates that it becomes "easier" to publish in a journal with each successive publication

Persistence vs Uncertainty

Can a quantitative picture of career dynamics shed light on the saying: *"publish or perish"* ?

Longitudinal career data:

Set A: 100 most-cited physicists, average h-index $\langle h \rangle$ = 61 ± 21

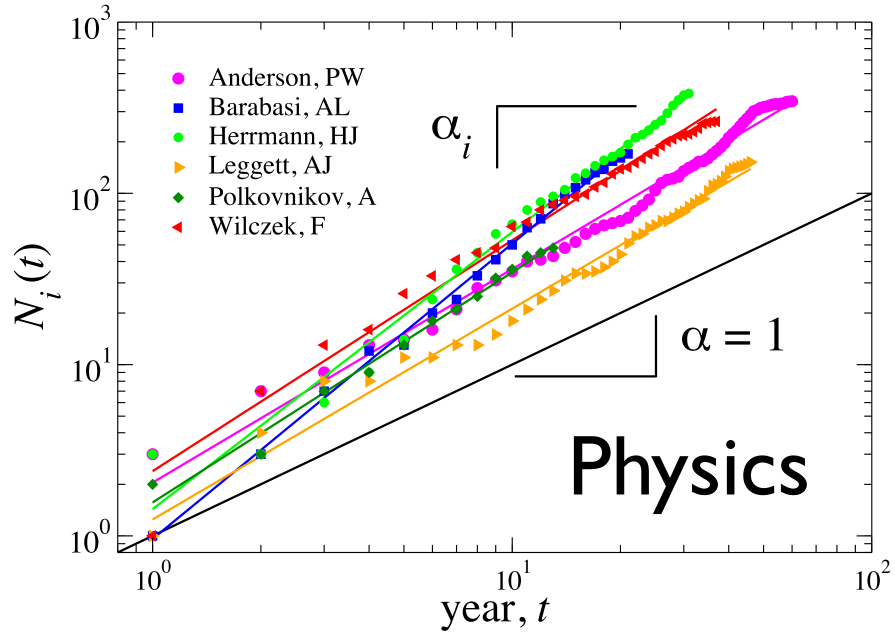
Set B: 100 additional highly-prolific physicists, $\langle h \rangle$ = 44 ± 15

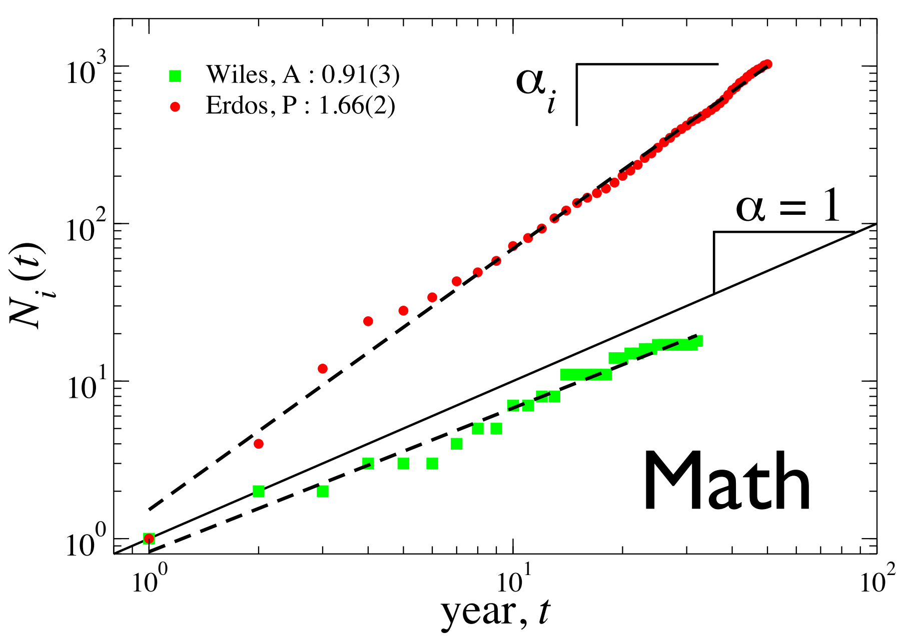
Set C: 100 current assistant professors from 50 US physics depts., $\langle h \rangle$ = 15 ± 7

Set D: 100 most-cited cell biologists, $\langle h \rangle$ = 98 ± 35

Set E: 50 highly-cited mathematicians, $\langle h \rangle$ = 20 ± 10

The career trajectory in science:
a tale of knowledge, collaboration, and reputation spillovers





Annual production of individual i

$n_i(t)$ number of publications in year t

Cumulative production, a proxy for career reputation

$$N_i(t) \equiv \sum_{t'=1}^{t} n_i(t')$$

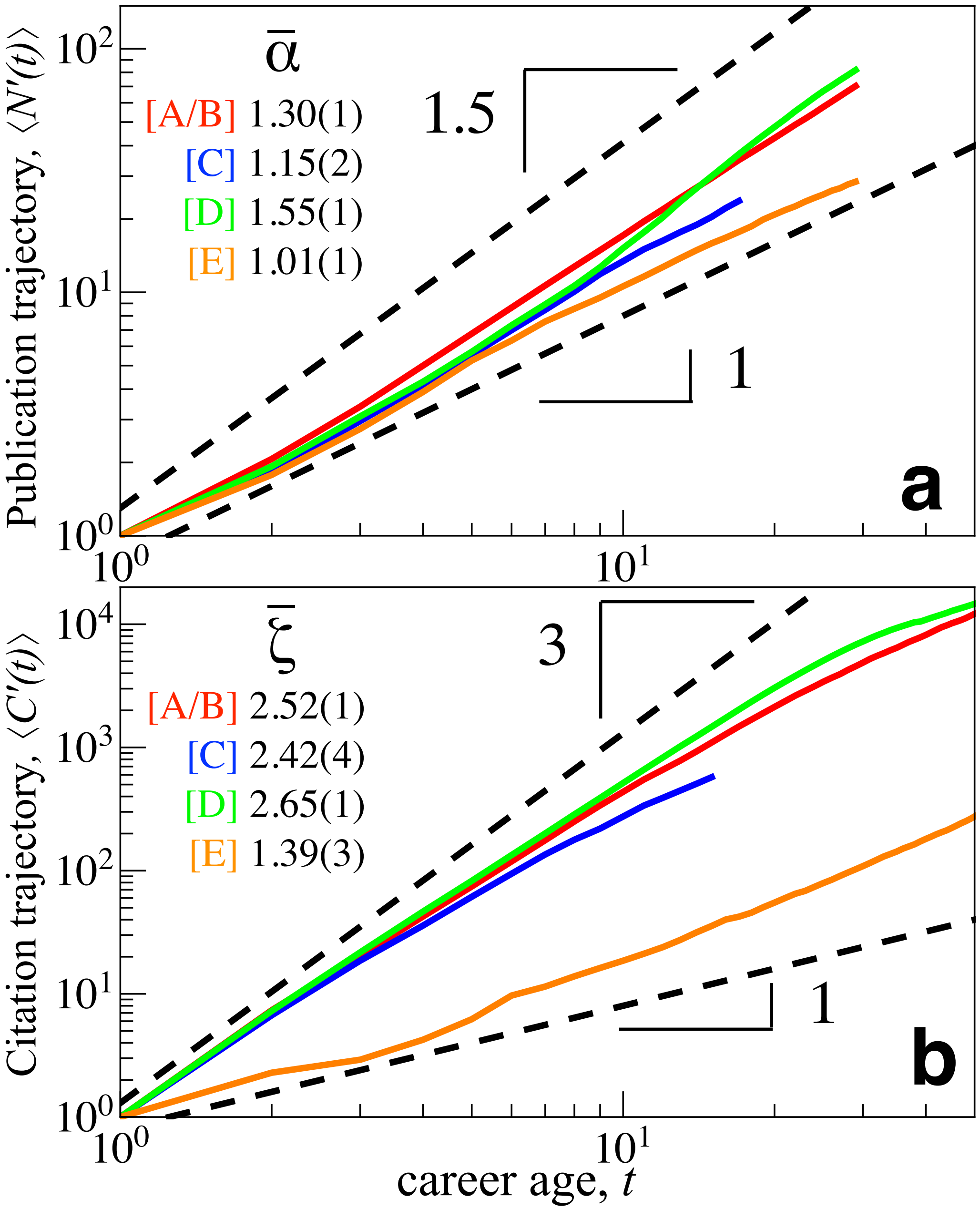
$$\approx A_i t^{\alpha_i}$$ ← for many prolific careers!

knowledge, reputation, and collaboration spillovers contribute to the increasing returns across the academic career

Cumulative advantage: Successful leaders become "attractors" of new opportunities

"Persistence and Uncertainty in the Academic Career," A. M. Petersen, M. Riccaboni, H. E. Stanley, F. Pammolli. Proc. Natl. Acad. Sci. USA 109, 5213-5218 (2012).

Are there characteristic career growth patterns?



(normalized) production trajectory $\langle N'(t) \rangle \sim t^{\bar{\alpha}}$

(normalized) cumulative citation trajectory $\langle C'(t) \rangle \sim t^{\bar{\zeta}}$

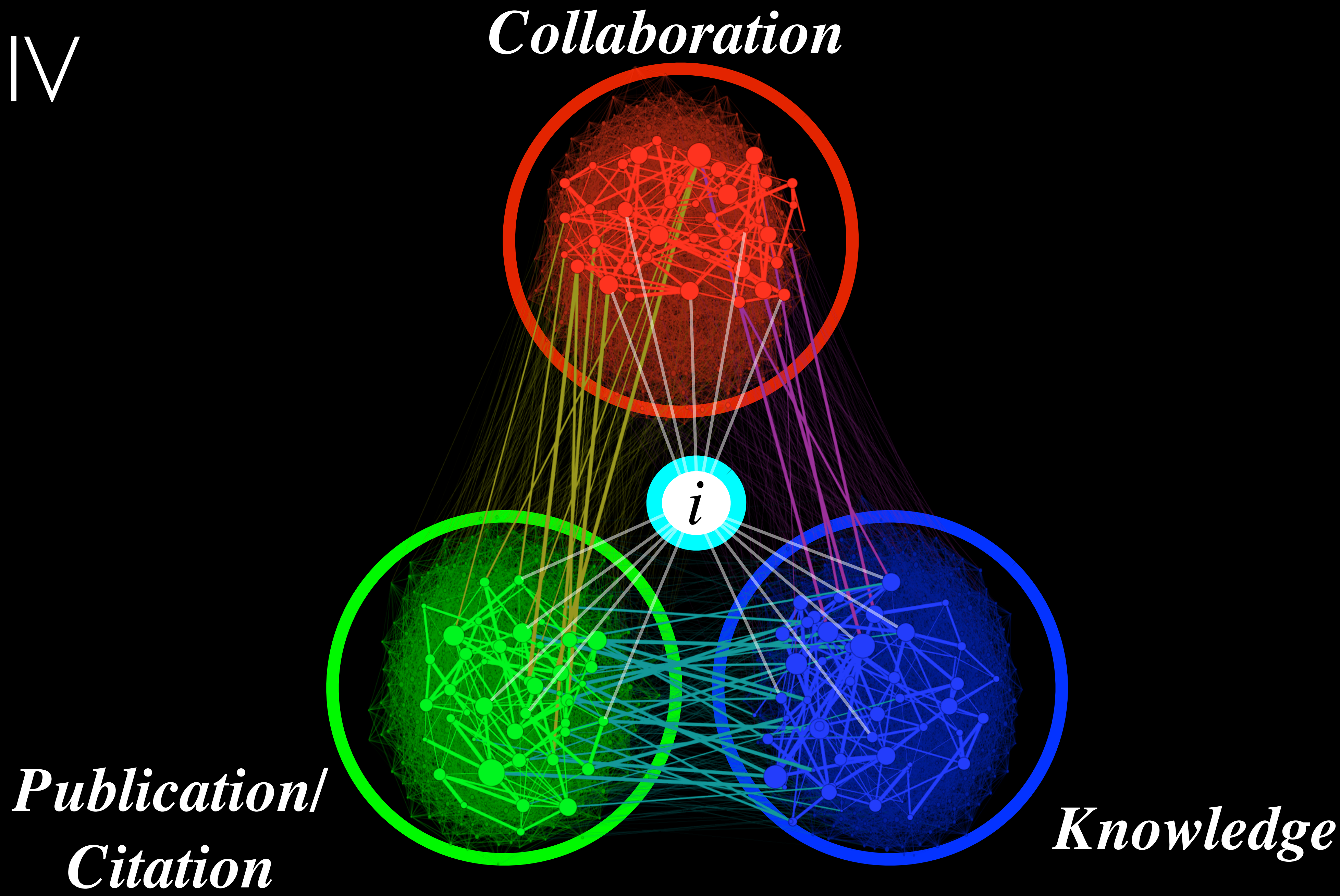
$\zeta > \alpha > 1 \Rightarrow$ *increasing returns*

Open questions:

How strong are the reputation spillovers which manifest in a tipping point whereby careers become “attractors” of new opportunities instead of “pursuers”

Co-Evolution of the Scientific Production Function

IV



A

Complexity

- coevolutionary system
- behavioral components

The holy grail:

a comprehensive disambiguated career portal for the entire scientific labor force

Online user-input repositories

proprietary: researcherid.com

RESEARCHERID
THOMSON REUTERS

integrated

non-proprietary: orcid.org

- also integrates grant/funding info

ORCID

Potential problems:

- honor system
- require constant updating
- might only serve to reinforce the "rich-get-richer" effect in science
- How to quantify the other productivity outputs associated with an academic career (teaching, community outreach, press/online coverage altmetrics , etc.)

Physiological/Behavioral components of games

High competition levels can make careers vulnerable to early career negative production shocks (ie stress, burn-out, productivity lulls, etc.)

Achievement-oriented systems: incentives for cut-throat “zero-sum” behavior, i.e. use of performance/cognitive enhancing drugs, blatant cheating and falsification



Sudden career termination in science due to ethical scandals

Jan Hendrik Schön Scandal (2001)

On October 31, 2002, *Science* withdrew eight papers written by Schön
On December 20, 2002, *Physical Review* withdrew six papers
On March 5, 2003, *Nature* withdrew seven papers

Diederik Alexander Stapel Scandal (2011)

Social psychologist made up data for at least 30 publications according to preliminary investigation, which is still ongoing.

Hisashi Moriguchi Scandal (2012)

“Transplant of induced pluripotent stem cells to treat heart failure probably never happened.... He is affiliated with University of Tokyo but not with Massachusetts General Hospital nor with Harvard Medical School. The study did not receive Institutional Review Board approval.” nature.com

Cognizant Enhancement Drugs (CED)

Professor's little helper

The use of cognitive-enhancing drugs by both ill and healthy individuals raises ethical questions that should not be ignored, argue **Barbara Sahakian** and **Sharon Morein-Zamir**.

NATURE|Vol 450|20/27 December 2007

"Is it cheating to use cognitive-enhancing drugs?.... How would you react if you knew your colleagues — or your students — were taking cognitive enhancers?... we know that a number of our scientific colleagues ... already use modafinil [Modiodal, Provigil] to counteract the effects of jetlag, to enhance productivity or mental energy, or to deal with demanding and important intellectual challenges..."

NATURE|Vol 452|10 April 2008

Poll results: look who's doping

In January, *Nature* launched an informal survey into readers' use of cognition-enhancing drugs. **Brendan Maher** has waded through the results and found large-scale use and a mix of attitudes towards the drugs.

"One in five respondents said they had used drugs for non-medical reasons to stimulate their focus, concentration or memory. Use did not differ greatly across age-groups..., which will surprise some. "

Towards responsible use of cognitive-enhancing drugs by the healthy

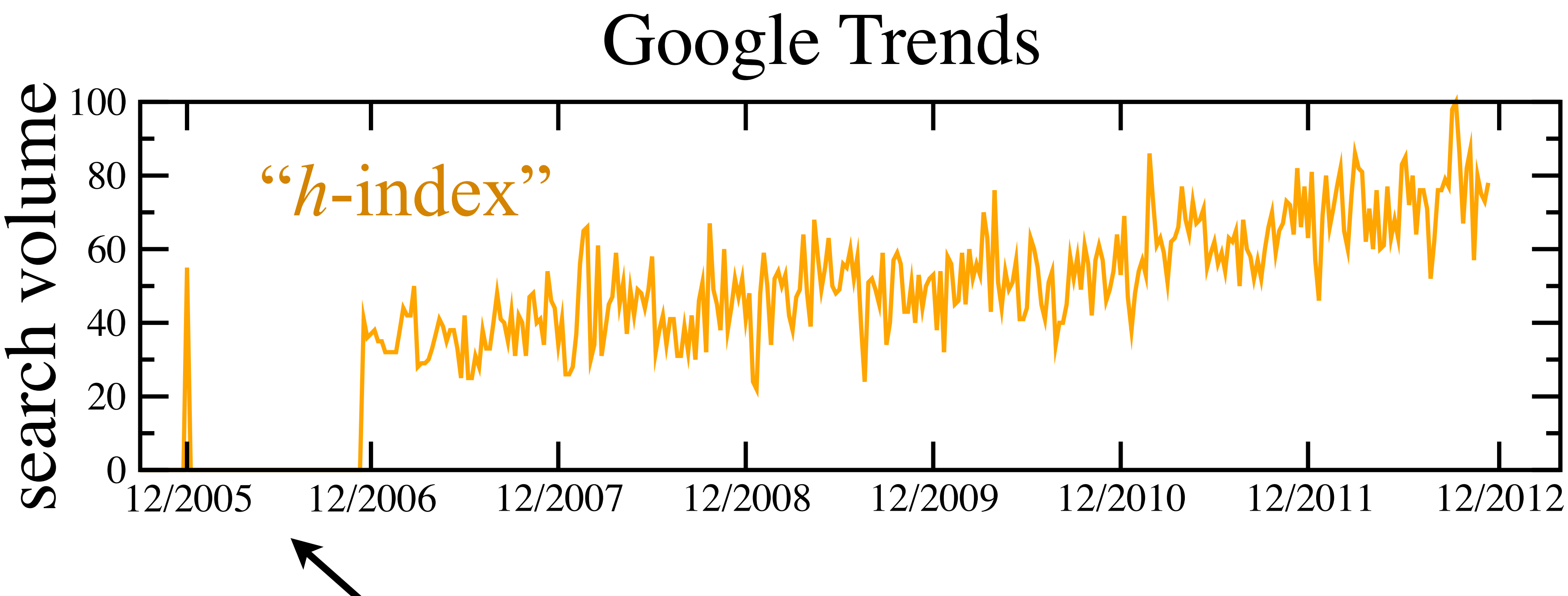
Society must respond to the growing demand for cognitive enhancement. That response must start by rejecting the idea that 'enhancement' is a dirty word, argue **Henry Greely and colleagues**.

"...one survey estimated that almost 7% of students in US universities have used prescription stimulants [Adderall and Ritalin] in this way, and that on some campuses, up to 25% of students had used them in the past year. These students are early adopters of a trend that is likely to grow, and indications suggest that they're not alone."

NATURE|Vol 456|11 December 2008

B

Deconstructing the *h*-index



* ~ Nov. 2005

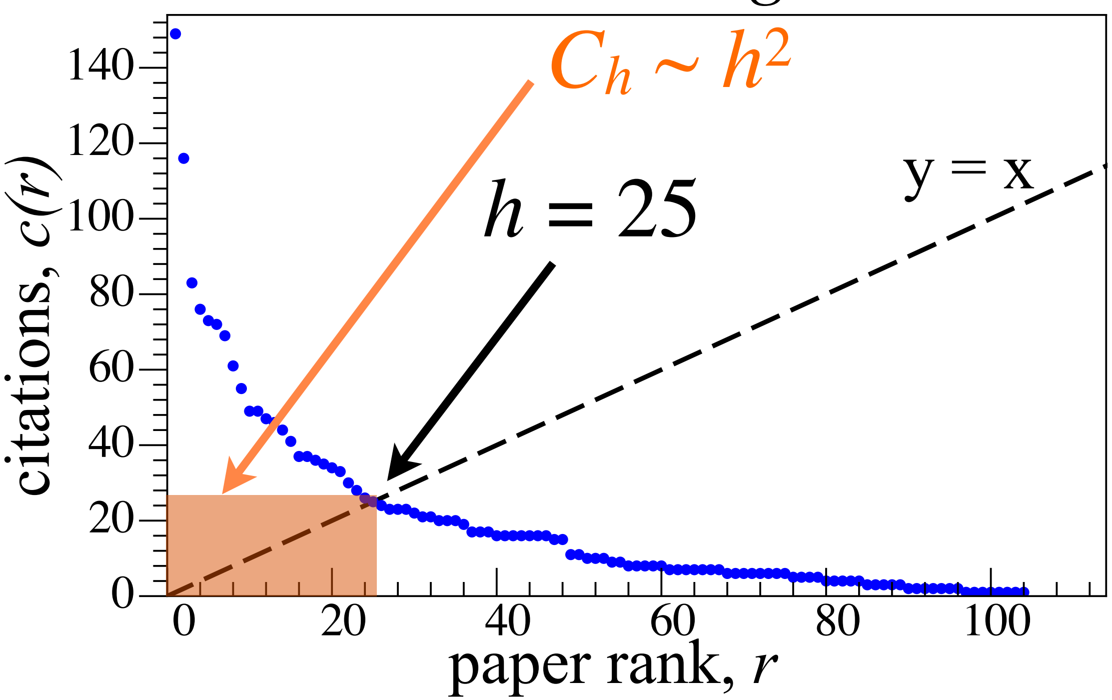
J. E. Hirsch,
"An index to quantify an individual's scientific research output".
Proc. Natl. Acad. Sci. USA ***102****, 16569- 16572 (2005).*
Google Citations = 2,653

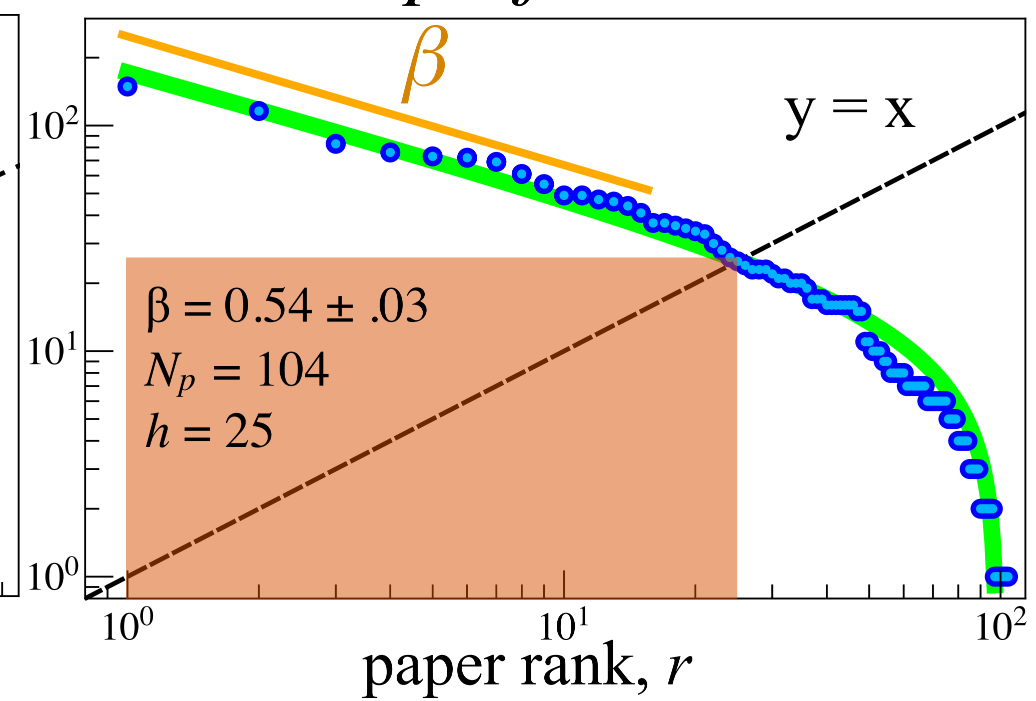
Quantifying scientific achievement: productivity -vs- impact

"A scientist has index *h* if *h* of his or her N_p papers have at least *h* citations each and the other (N_p-h) papers have ≤ *h* citations each."

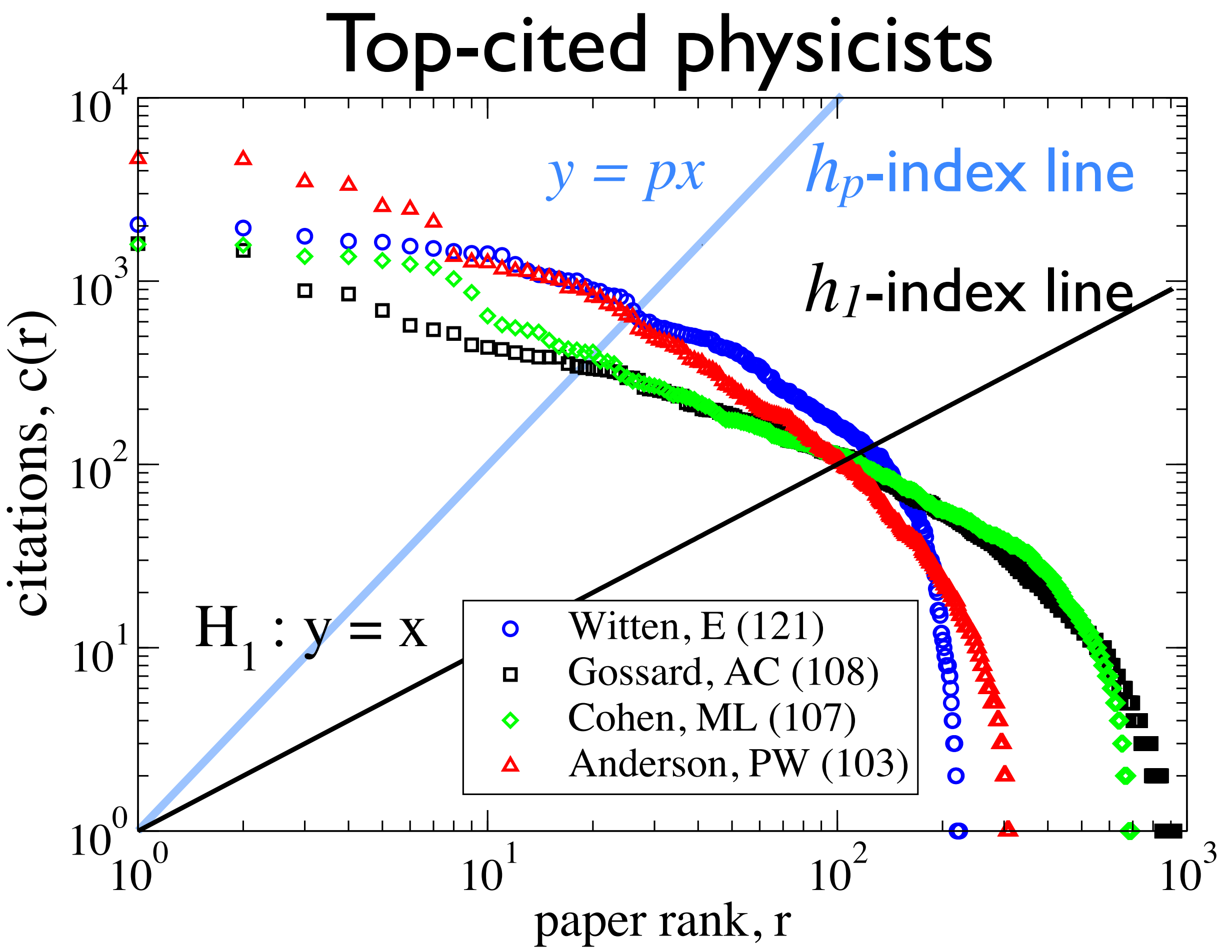
J. E. Hirsch, "An index to quantify an individual's scientific research output". PNAS **102**, 16569- 16572 (2005).

Prof. Ioannis Pavlidis, U. Houston
Google Scholar Citations profile





Regularities in the rank-citation profile $c_i(r)$



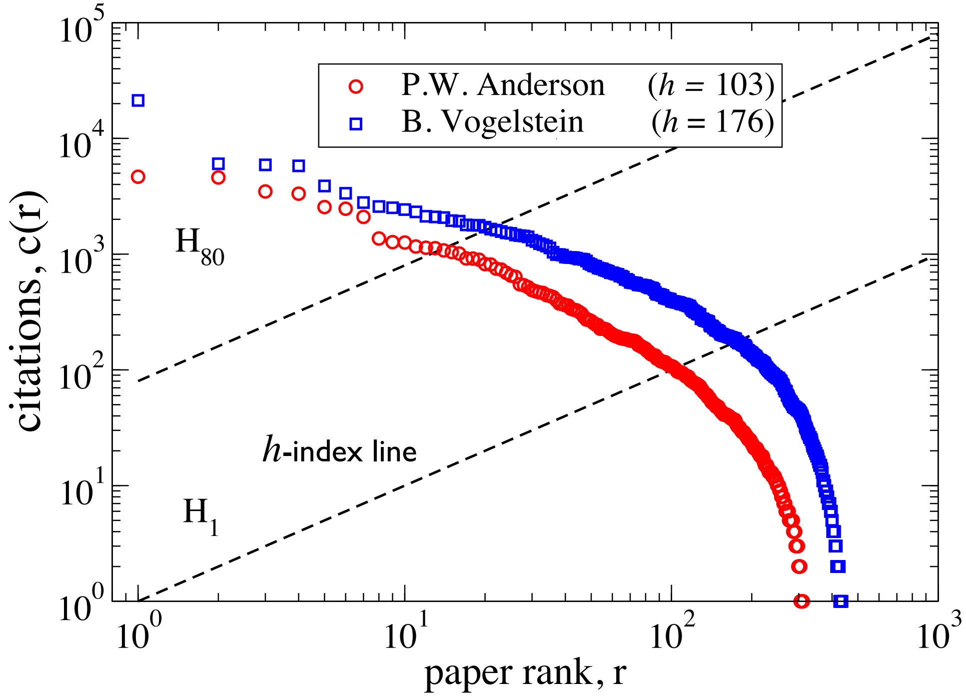
$c_i(r)$ is the rank-ordered (Zipf) citation distribution of the N papers published by individual i in his/her entire career

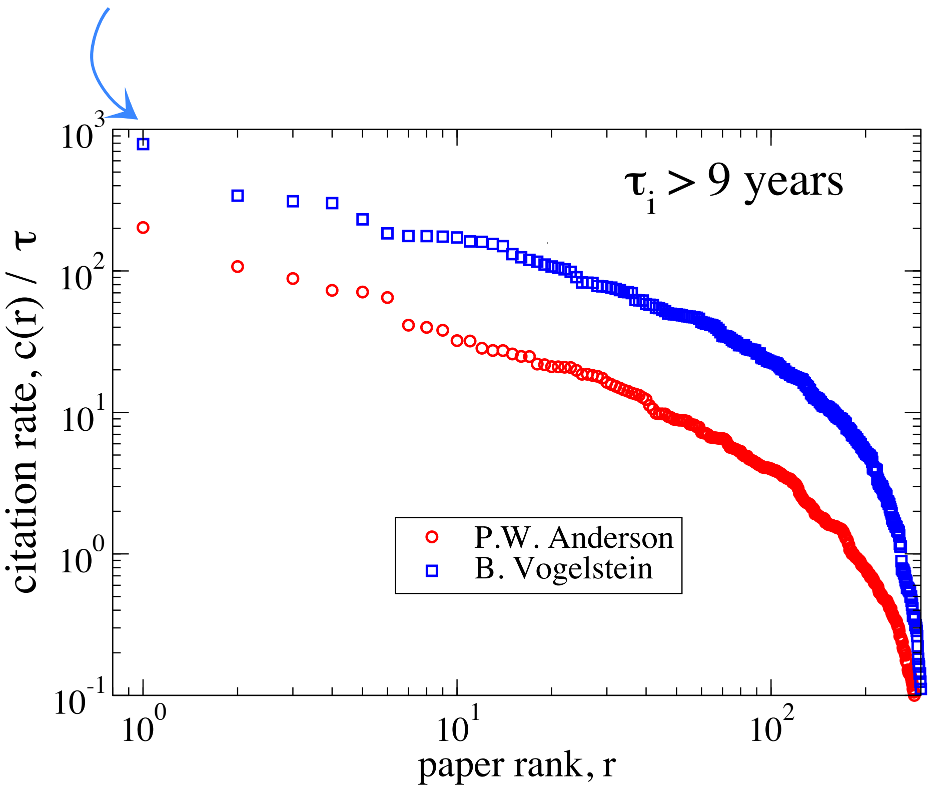
A. M. Petersen, H. E. Stanley, S. Succi. "Statistical regularities in the rank-citation profile of scientists." Scientific Reports 1, 181 (2011).

Interestingly, even the very top scientists have a significant number of papers that go relatively un-cited.

Accounting for the time-dependence of $c_i(r)$

FEINBERG, A, VOGELSTEIN, B.
A TECHNIQUE FOR RADIOLABELING DNA RESTRICTION ENDONUCLEASE FRAGMENTS TO HIGH SPECIFIC ACTIVITY. Anal. Biochem. (1983) Citations [ISI] = 21,270 in 2010 (ave. rate 760 /year)



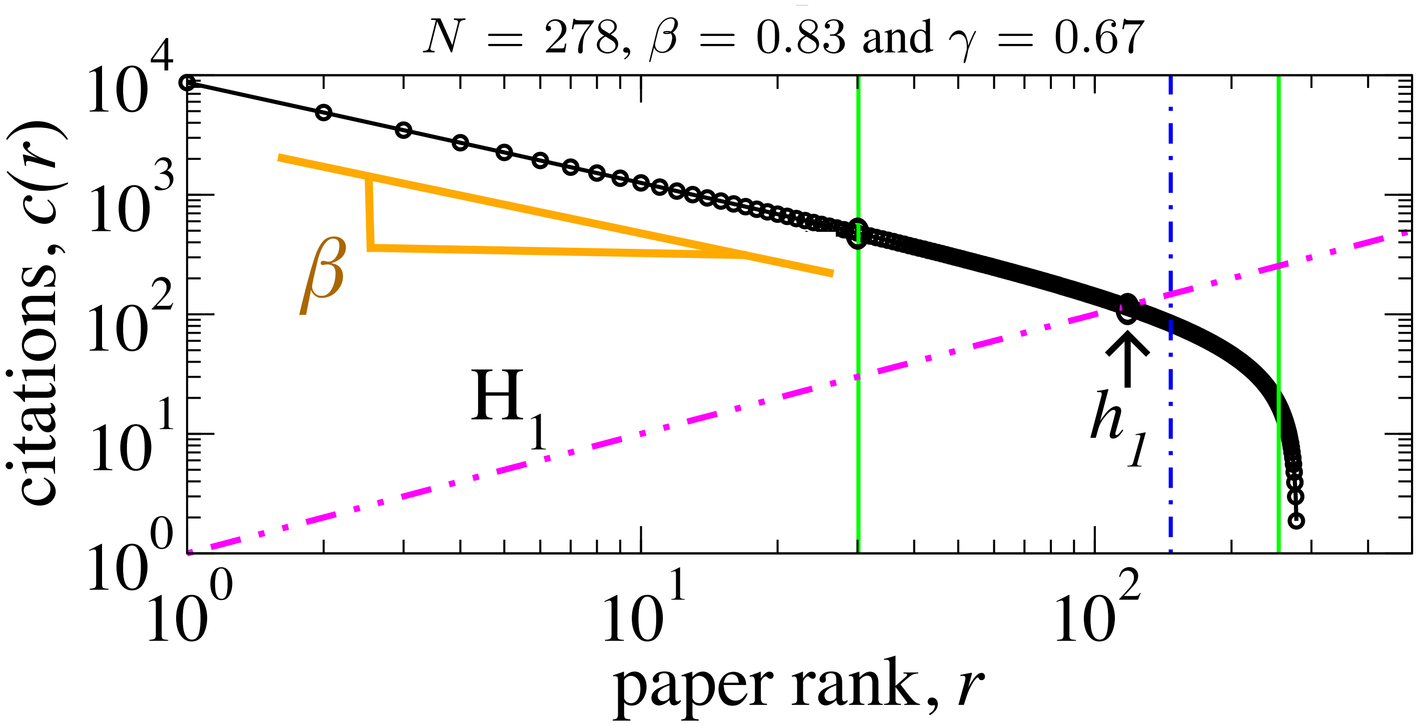


τ_i is the lifetime of the paper at the time of download

Even the average citation rate $\dot{c}_i(r) \equiv c_i(r)/\tau$

is heavily tailed!

The Discrete Generalized Beta Distribution (DGBD) model for $c_i(r)$



$$c(r) \equiv A r^{-\beta} (N + 1 - r)^{\gamma} \; .$$

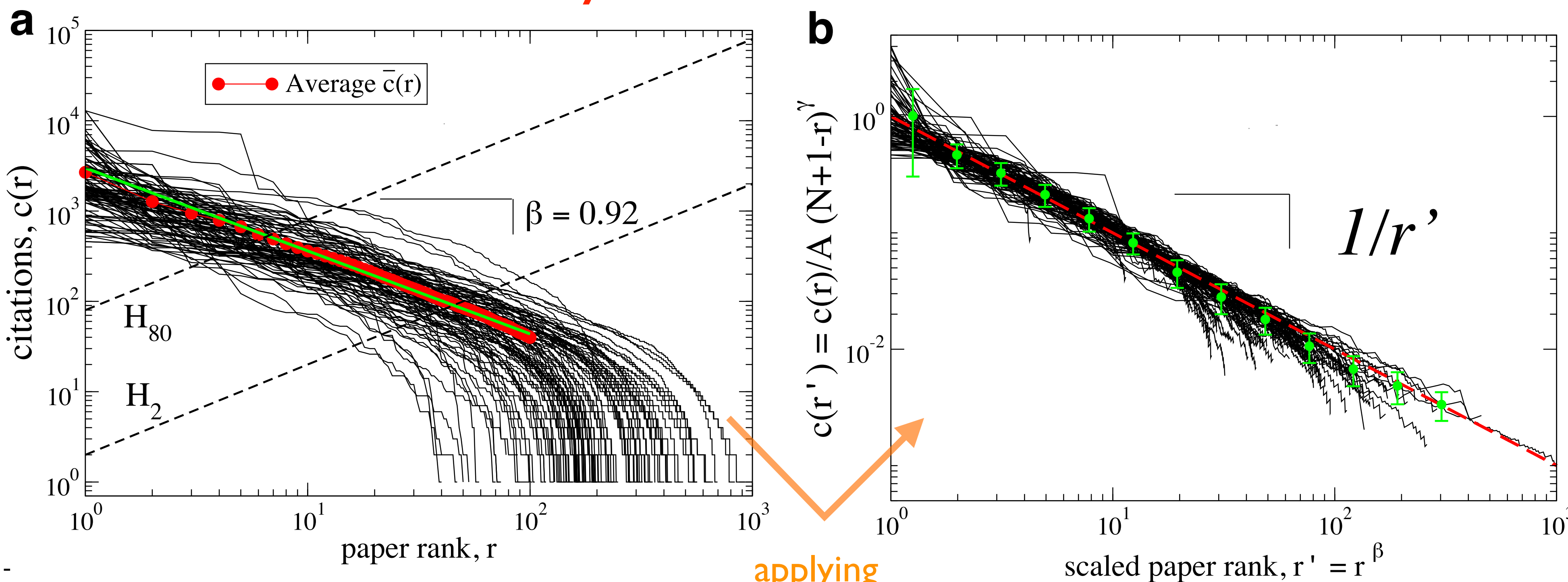
N_i = # of publications
β_i = scaling slope of top papers
γ_i = truncation scaling of less-cited papers
C_i = total citations from all papers

Scaling relation between C, h, and β

$$C \sim h^{1+\beta}$$

⇒ Hence, knowing both the h-index and C is ≈ redundant

A comparison of $c_i(r)$ for the top-100 "champions" of Physical Review Letters



applying
scaling transformation

Set A: average h-index $<h>$ = 61 ± 21

Discrete Generalized
Beta Distribution(DGBD):

$$c(r) \equiv A r^{-\beta} (N + 1 - r)^{\gamma} \, .$$

Average values of the DGBD model parameters:

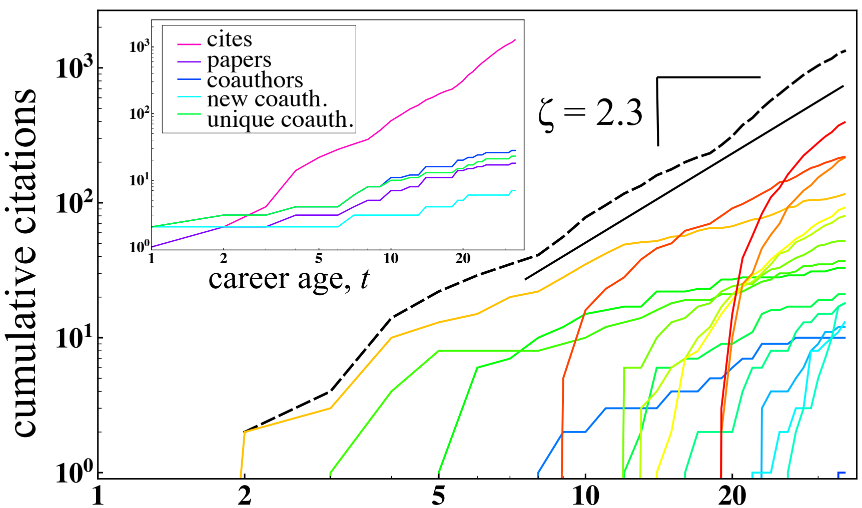
$<\beta> = 0.83 \pm 0.23$ and $<\gamma> = 0.67 \pm 0.19$

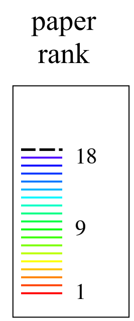
"Don't throw the *important career data* out with the bathwater"

towards comprehensive publication, impact, and collaboration profiles

Andrew Wiles

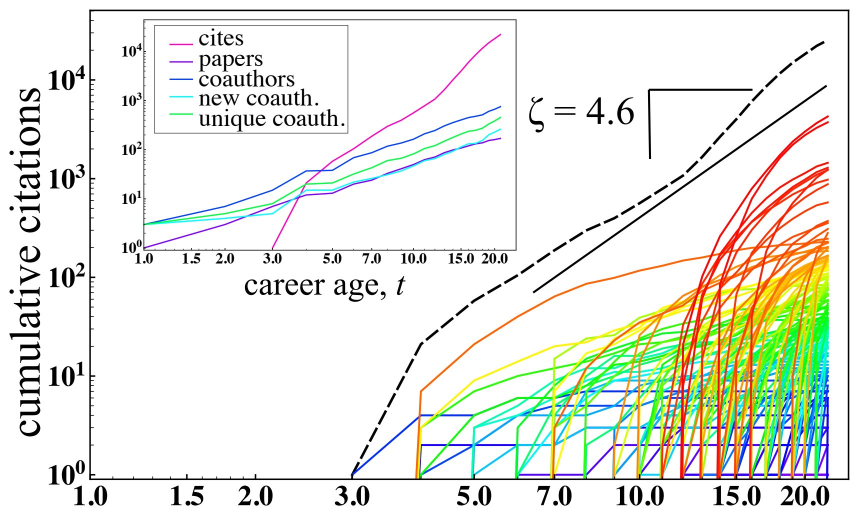
Wiles, Andrew, Modular elliptic curves and Fermat's Last Theorem. *Annals of Mathematics* **141**, 443–551 (1995).

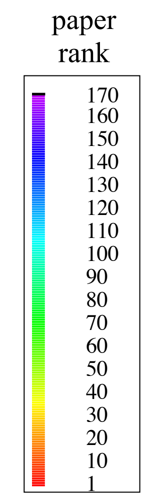




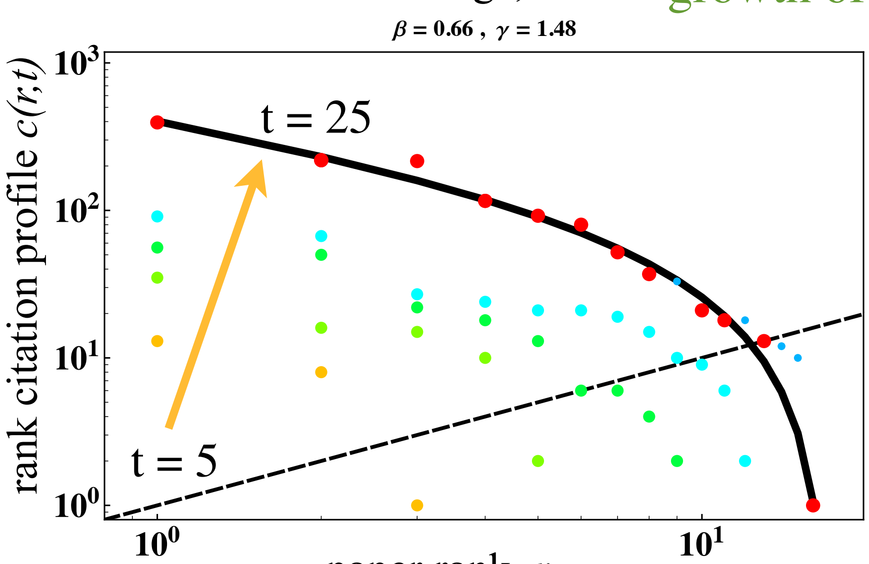
Albert-Laszlo Barabasi

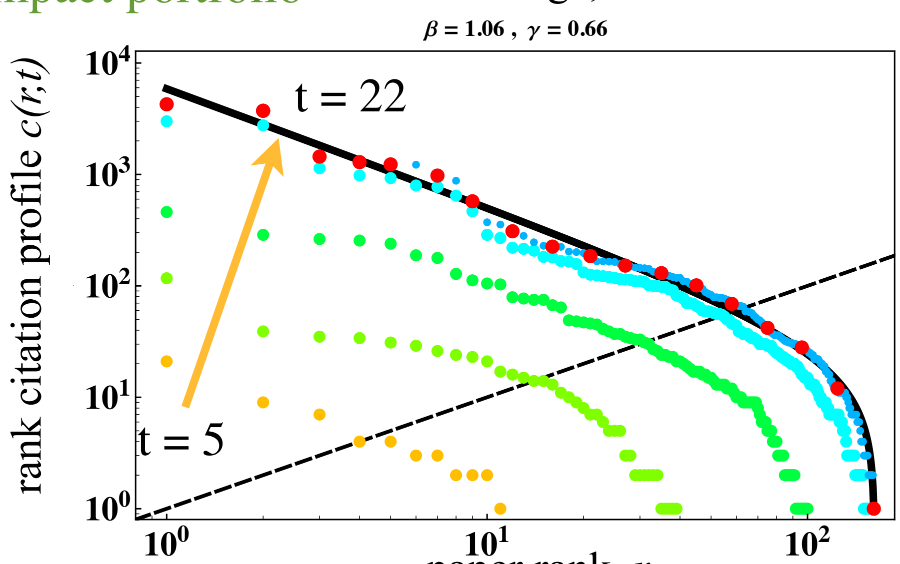
R. Albert, H. Jeong, A.-L. Barabási, Diameter of the world wide web. *Nature* **401**, 130-131 (1999).





growth of the impact portfolio



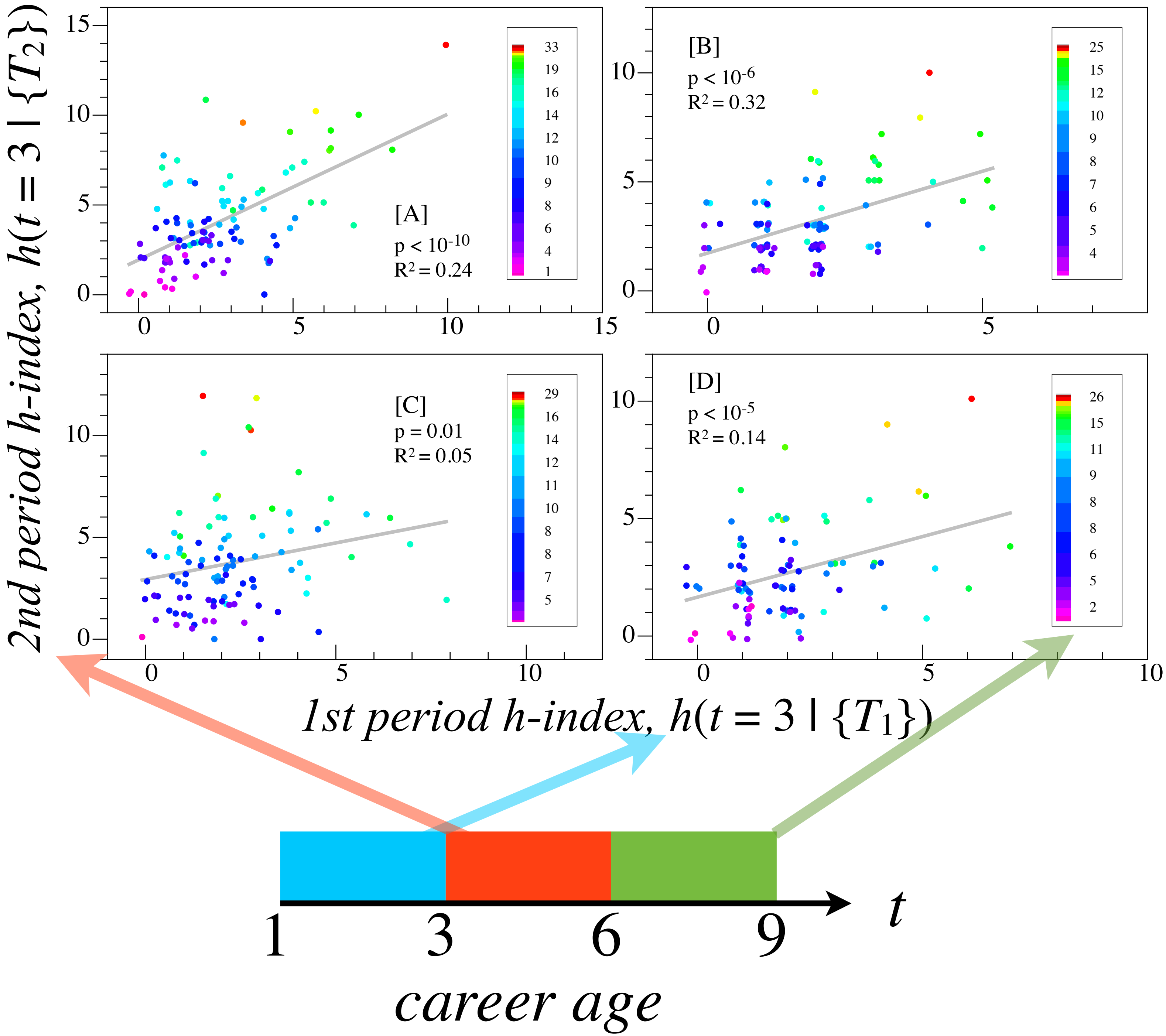


emergence of the DGBD profile

C

Forecasting careers?

Difficulty of predicting impact of future papers



Predicting scientific success

Daniel E. Acuna, **Stefano Allesina** and **Konrad P. Kording** present a formula to estimate the future *h*-index of life scientists.

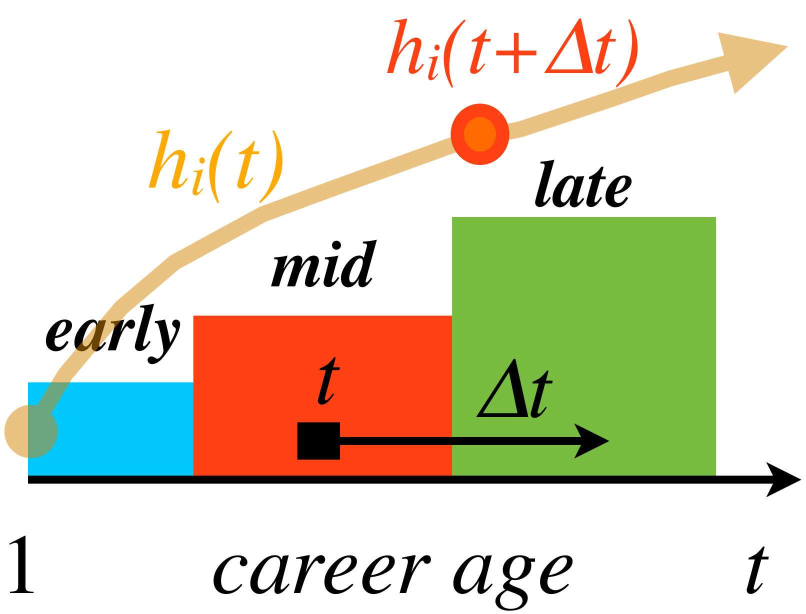
13 SEPTEMBER 2012 | VOL 489 | NATURE | 201

Major Flaws!

1) aggregating across different career ages

2) h-index is non-decreasing ⇒

R^2 will be artificially large



METRICS

Predict your future h*-index*

These are approximate equations for predicting the *h*-index of neuroscientists in the future. They are probably reasonably precise for life scientists, but likely to be less meaningful for the other sciences. Try it for yourself online at go.nature.com/z4rroc.

- **Predicting next year** ($R^2=0.92$):

$$h_{+1}=0.76+0.37\sqrt{n}+0.97h-0.07y+0.02j+0.03q$$

- **Predicting 5 years into the future** ($R^2=0.67$):

$$h_{+5}=4+1.58\sqrt{n}+0.86h-0.35y+0.06j+0.2q$$

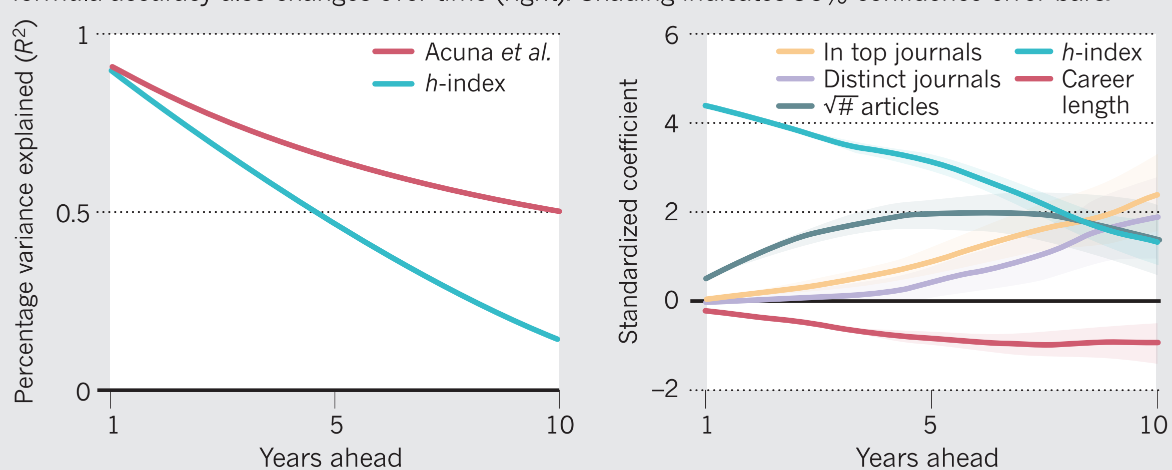
- **Predicting 10 years into the future** ($R^2=0.48$):

$$h_{+10}=8.73+1.33\sqrt{n}+0.48h-0.41y+0.52j+0.82q$$

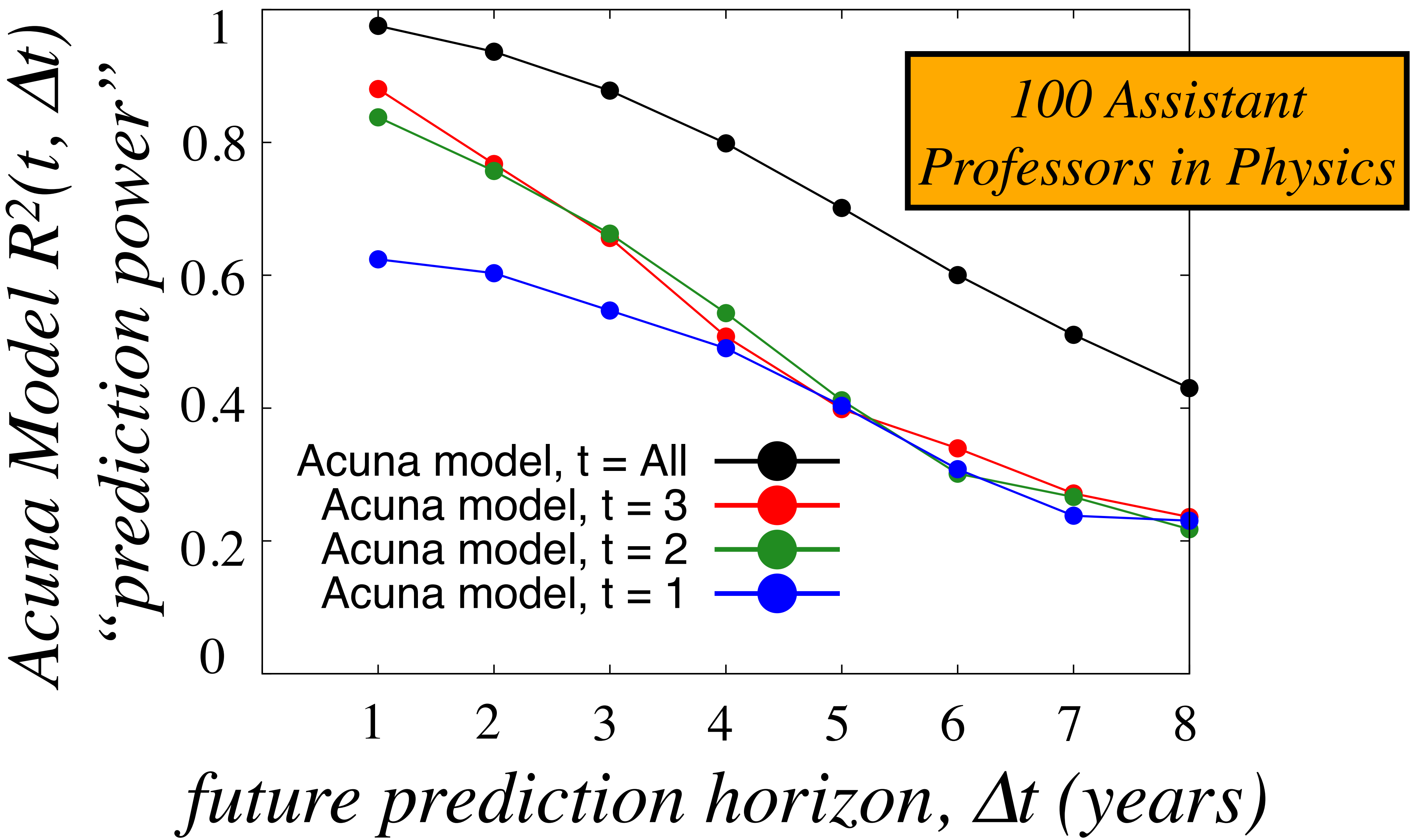
Key: *n*, number of articles written; *h*, current *h*-index; *y*, years since publishing first article; *j*, number of distinct journals published in; *q*, number of articles in *Nature*, *Science*, *Nature Neuroscience*, *Proceedings of the National Academy of Sciences* and *Neuron*.

PATHS TO SUCCESS

The accuracy of future *h*-index prediction decreases over time, but the Acuna *et al*. formula predicts future *h*-index better than does current *h*-index alone (left). The contribution of each factor to the formula accuracy also changes over time (right). Shading indicates 95% confidence error bars.



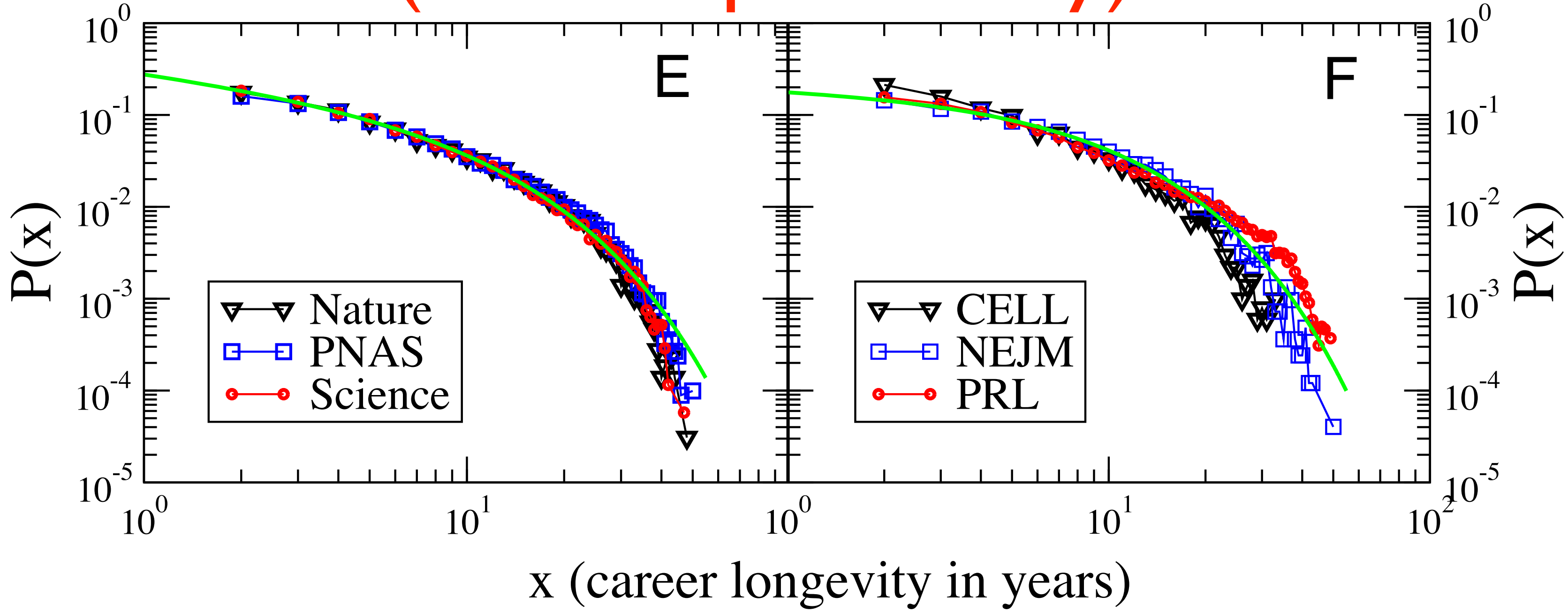
$h_i(t+\Delta t \mid t)$: predicting growth *conditional* on (early) career age t is much more difficult



D

Career longevity:
uncertainty associated with career hazards

Difficulty in predicting career longevity (survival probability)

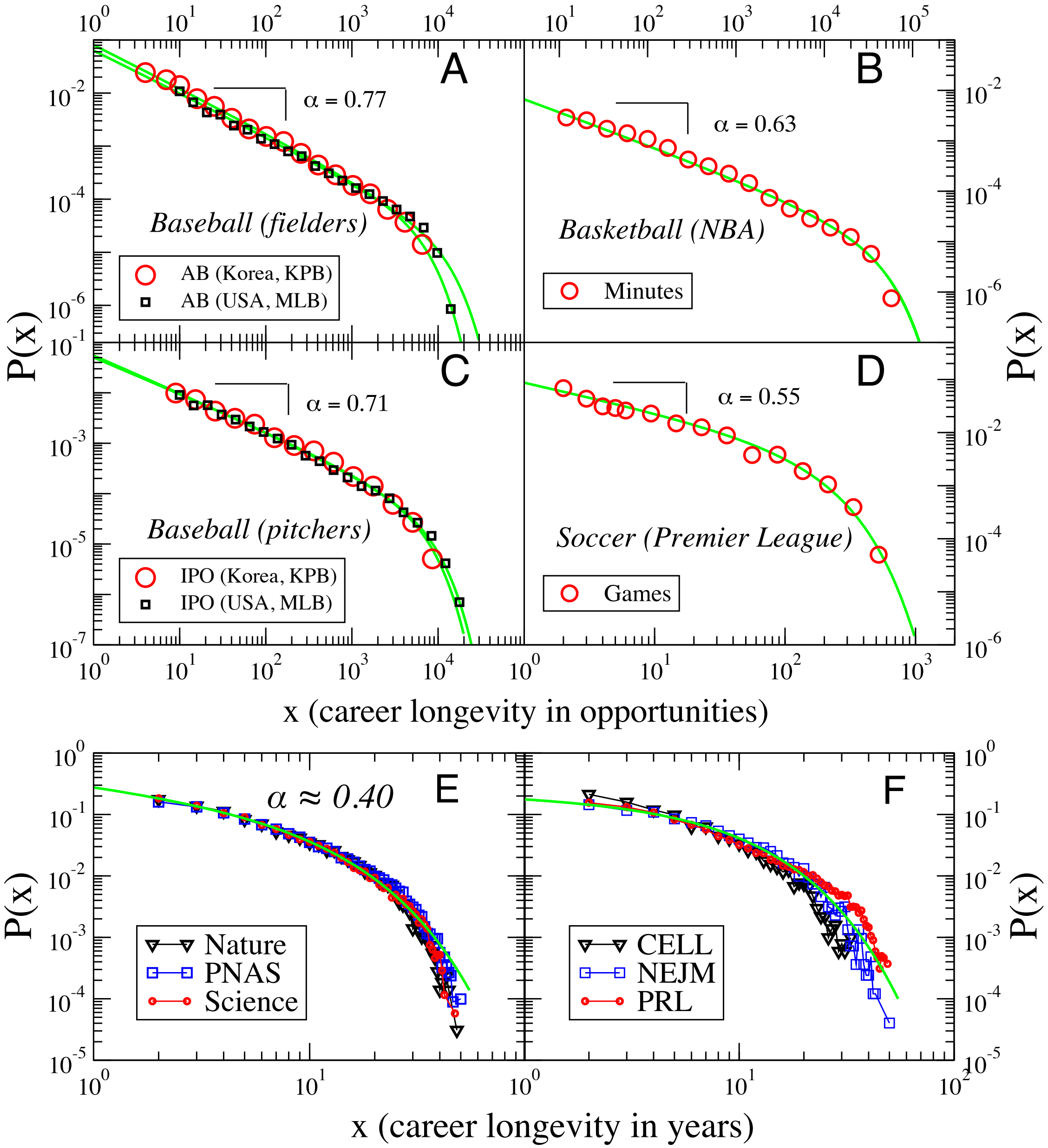


high-impact journals are competitive arenas

- Each author *i* has *n* articles in a given journal *j*. As a proxy for career longevity in academia, we define the journal longevity *x* as the number of years separating his/her first and last publication in journal *j* :

$$x_{i,j} = y_{i,j}(f) - y_{i,j}(0) + 1$$

Empirical longevity distributions in sports and academica



*"Quantitative and empirical demonstration of the Matthew effect in a study of career longevity." A. M. Petersen, W.-S. Jung, J.-S. Yang, H. E. Stanley. Proc. Natl. Acad. Sci. USA **108**, 18-23 (2011).*

Major League Baseball

- 130+ years of player statistics, ~ 15,000 careers

``One-hit wonders"

- 3% of all fielders finish their career with ONE at-bat!
- 3% of all pitchers finish their career with less than one inning pitched!

``Iron horses"

- Lou Gehrig (the Iron Horse): NY Yankees (1923-1939)
- Played in 2,130 consecutive games in 15 seasons! 8001 career at-bats!
- Career & life stunted by the fatal neuromuscular disease, amyotrophic lateral sclerosis (ALS), aka Lou Gehrig's Disease

Survival (longevity) analysis of tenure track assistant professors

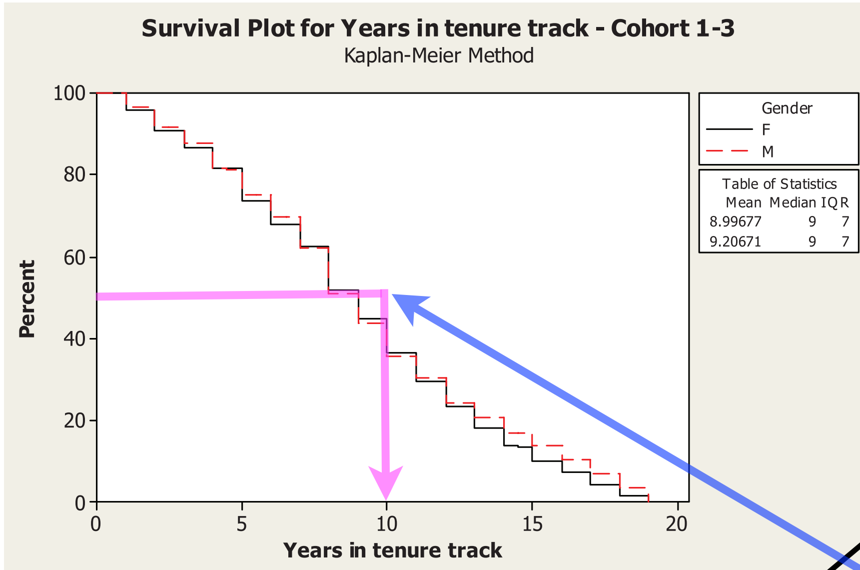


Fig. 1. Nonparametric survival curve for faculty who entered between 1990 and 2002 by gender. IQR, interquartile range.

U.S. universities are concerned about faculty retention in science and engineering (*1–4*). When a faculty member leaves prematurely, they suffer disruptions in teaching and mentoring as well as significant economic losses (*1*). Start-up costs in engineering and natural sciences can range from $110,000 to nearly $1.5 million (*3*), and it may take up to 10 years to recoup this investment (*4*).

50% of faculty leave by year 10

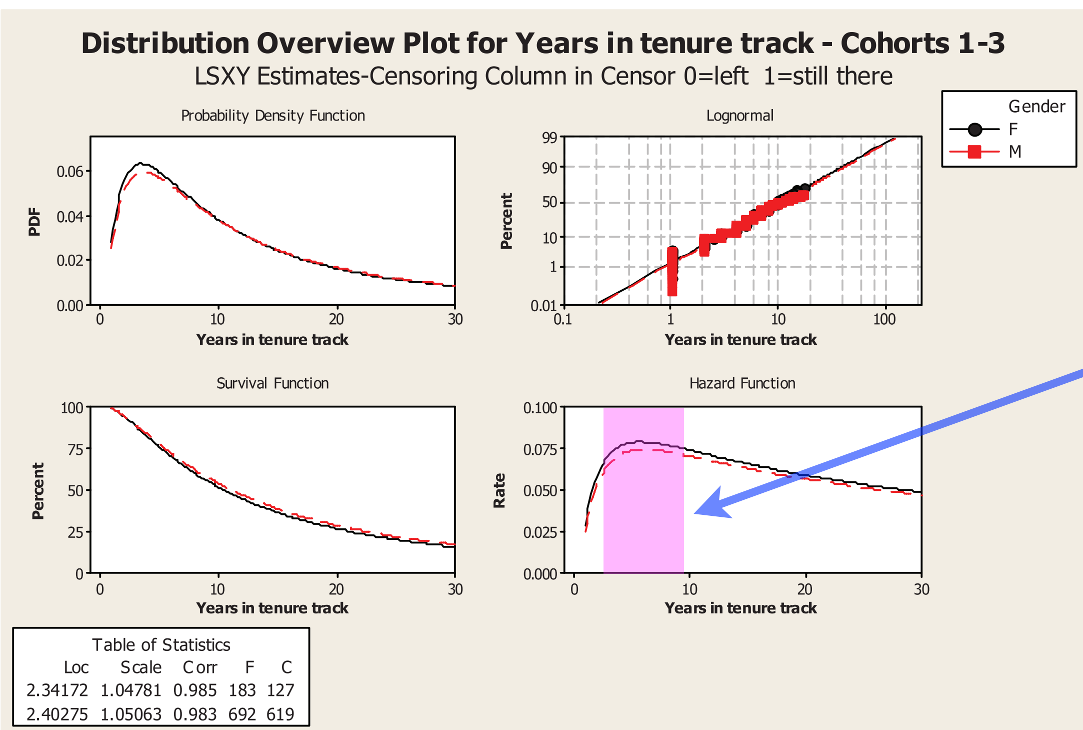


Fig. 2. Survival analysis of faculty who entered between 1990 and 2002 by gender. LSXY, least squares; F, number that left; C, number still remaining.

high-risk period peaks around pre-tenure year 6

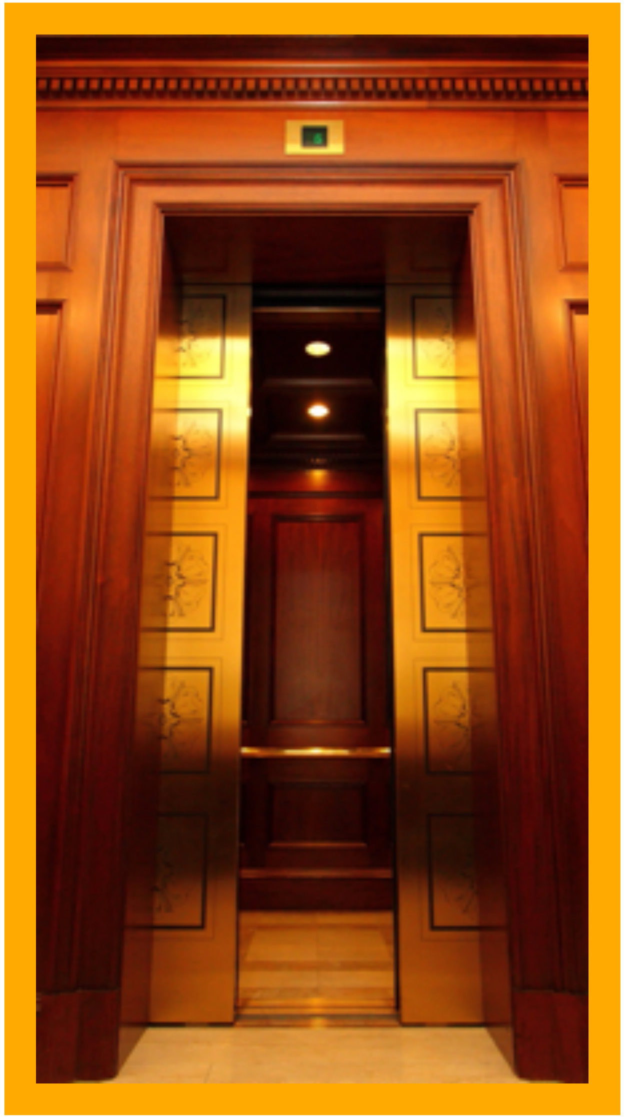
Survival Analysis of Faculty Retention in Science and Engineering by Gender

Deborah Kaminski[1]* and Cheryl Geisler[2]

17 FEBRUARY 2012 VOL 335 **SCIENCE** www.sciencemag.org

E

A "real-world" example of the tenure decision: Italian Abilitazione



courtesy of F. Radicchi

A practical example

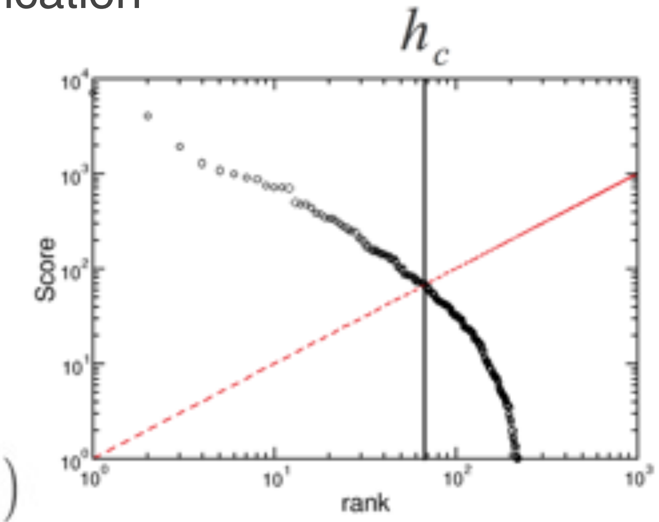
the Italian National Scientific Qualification

1) Number of papers: $I(N_p, A_A) = \frac{10 N_p}{A_A}$

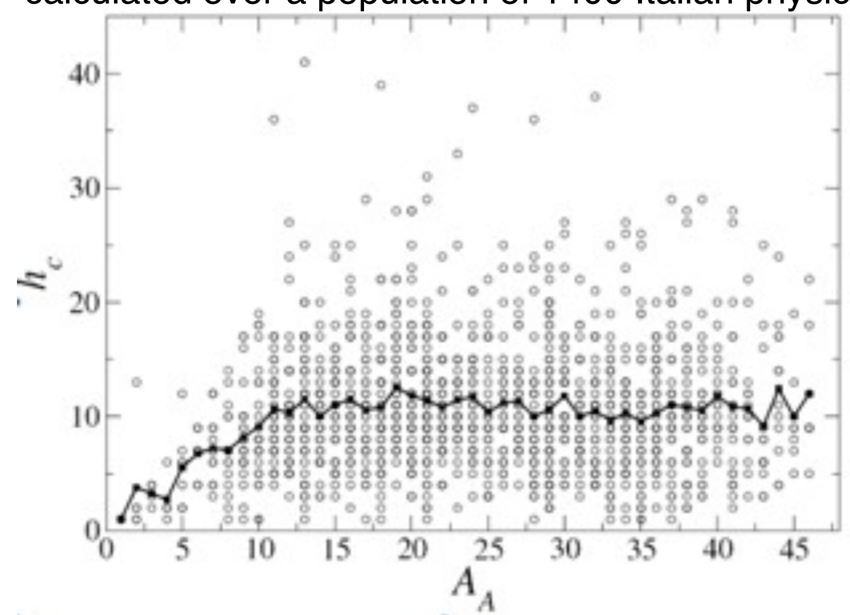
2) Number of citations: $I(N_C, A_A) = \frac{N_C}{A_A}$

3) Contemporary h-index: $S(i, t_i, t) = \frac{4}{(t - t_i + 1)} C(i, t_i, t)$

Sidiropoulos A et al. Scientometrics 72, 253 (2007)



calculated over a population of 1400 Italian physicists



R. Manella and P. Rossi, arXiv:1207.3499 (2012)

N_p total number of publications

A_A academic age

N_c total number of citations

$C(i, t_i, t)$ citations accumulated up to year t by paper i published in year t

source: "Abilitazione Scientifica Nazionale – La normalizzazione degli indicatori per l'eta' accademica", ANVUR, Jul. 2012 courtesy of F. Radicchi

A practical example

the Italian National Scientific Qualification

		associate professor			full professor		
		norm. pub.s	norm. citations	h-c index	norm. pub.s	norm. citations	h-c index
Mathematics	01/A1	5	1.74	2	4	1.37	2
	01/A2	8	1.65	2	9	3.23	3
	01/A3	10	4.34	4	14	8	5
Physics	02/A1	59.5	104.08	18	78	105.03	22
	02/B2	**37.5**	**40.08**	**11**	**47.5**	**75.94**	**14**
Biology	05/A2	14	24.45	8.5	20	37.47	10
	05/C1	21.5	15.77	8	26	18.63	9
Chemistry	03/A1	26	29.47	9	41	53.81	12
	03/B1	31	47.05	11	49.5	62.38	13

source: "Mediane per candidati all'abilitazione scientifica nazionale a professore associato/ordinario", ANVUR, Jul. 2012

General take-home messages

- Complex career dynamics: Knowledge, reputation, and collaboration spillovers are major factors leading to increasing returns along the scientific career trajectory

- Science as an evolving institution: An institutional setting that neglects specific features of academic career trajectories (increasing returns from knowledge spillovers and cumulative advantage, collaboration factors, career uncertainty) is likely inefficient and unfair.

- Nano-sociology: A data-centric ("big data") understanding of the production function of individual scientists can improve academic policies aimed at increasing career sustainability and decreasing career risk

- Competition and Reward: There are many analogies between the superstars in science and the superstars in professional sports, possibly arising from the generic aspects of competition. Currently, the contract length, compensation, and appraisal timescale in these two professions are VERY different. Is Science becoming more like professional sports?

I) "Methods for measuring the citations and productivity of scientists across time and discipline," A. M. Petersen, F. Wang, H. E. Stanley. Phys. Rev. E 81, 036114 (2010).

II) "Quantitative and empirical demonstration of the Matthew effect in a study of career longevity," A. M. Petersen, W.-S. Jung, J.-S. Yang, H. E. Stanley. Proc. Natl. Acad. Sci. USA 108, 18-23 (2011).

III) "On the distribution of career longevity and the evolution of home run prowess in professional baseball," A. M. Petersen, W.-S. Jung, H. E. Stanley. Europhysics Letters 83, 50010 (2008).

IV) "'Methods for detrending success metrics to account for inflationary and deflationary factors," A. M. Petersen, O. Penner, H. E. Stanley. Eur. Phys. J. B 79, 67-78 (2011).

V) "Statistical regularities in the rank-citation profile of scientists," A. M. Petersen, H. E. Stanley, S. Succi. Scientific Reports 1, 181 (2011).

VI) "Persistence and Uncertainty in the Academic Career," A. M. Petersen, M. Riccaboni, H. E. Stanley, F. Pammolli. Proc. Natl. Acad. Sci. USA 109, 5213-5218 (2012).

Thank You!

A special thanks to my collaborators: Santo Fortunato, Woo-Sung Jung, Fabio Pammolli, Raj Pan, Orion Penner, Massimo Riccaboni, Gene Stanley, Sauro Succi, Fengzhong Wang, and Jae-Sook Yang

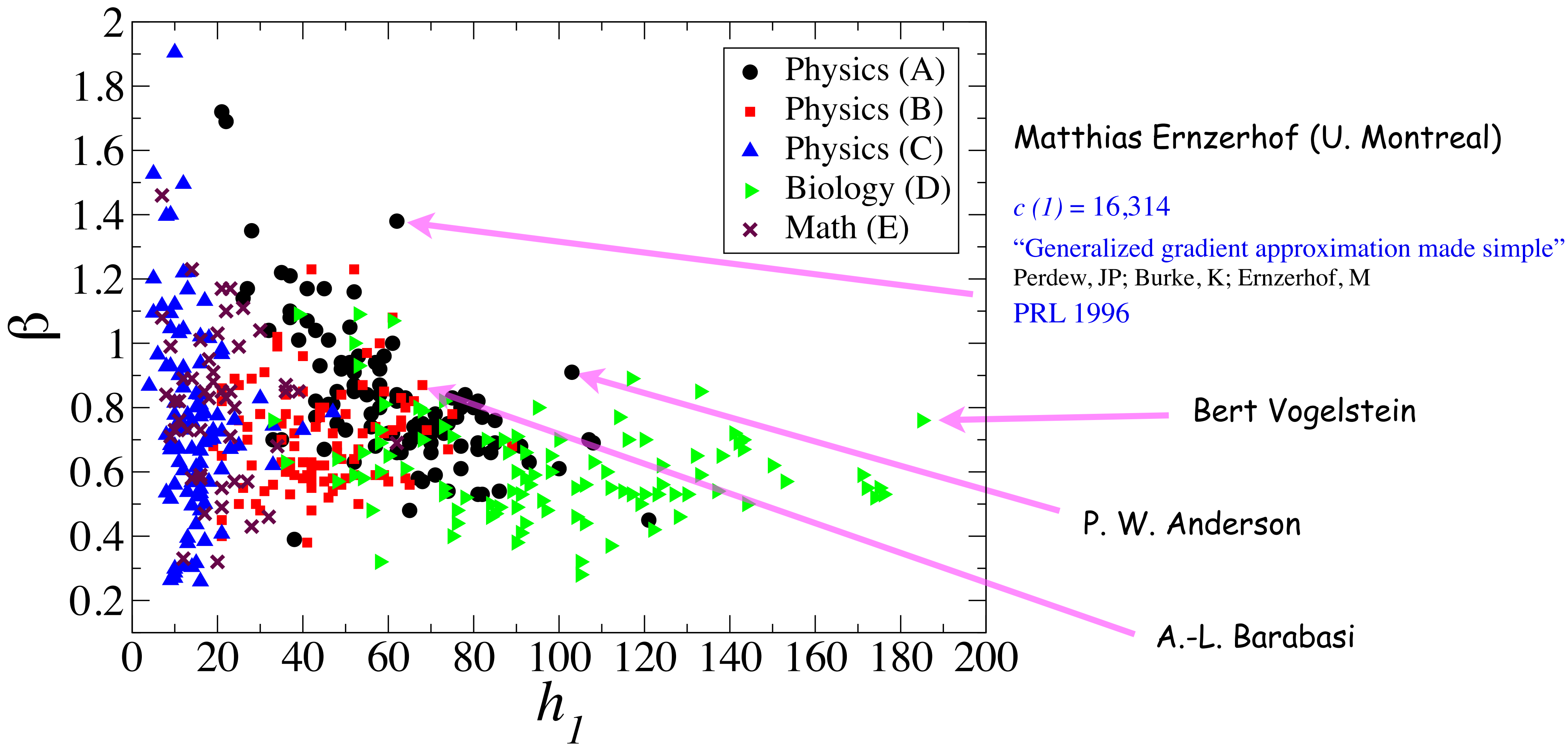
http://physics.bu.edu/~amp17/

Title: **Identifying potential pitfalls in the quantitative appraisal system for scientific careers**

Abstract:

Quantitative measures are becoming increasingly prevalent in the scientific appraisal of countries, universities, departments, and notably, individuals. In this talk I will discuss the potential pitfalls arising from the appraisal of individual careers based on citation metrics, a proceeding which is likely to occur at several stages of an academic career, from postdoctoral and faculty appointments to career achievement awards. Using longitudinal career data for 450 scientists, ranging from assistant professors to Nobel laureates and Fields medal winners, I will demonstrate a graphically intuitive method for visualizing an individual's publication profile. While much ado has been made about the h-index, a metric intended to measure simultaneously the productivity and impact of a scientist, I will argue for the careful use of this and related quantitative measures. With the remaining time, I will illustrate the complex dichotomy of competition and collaboration in science.

The β -vs- h parameter space

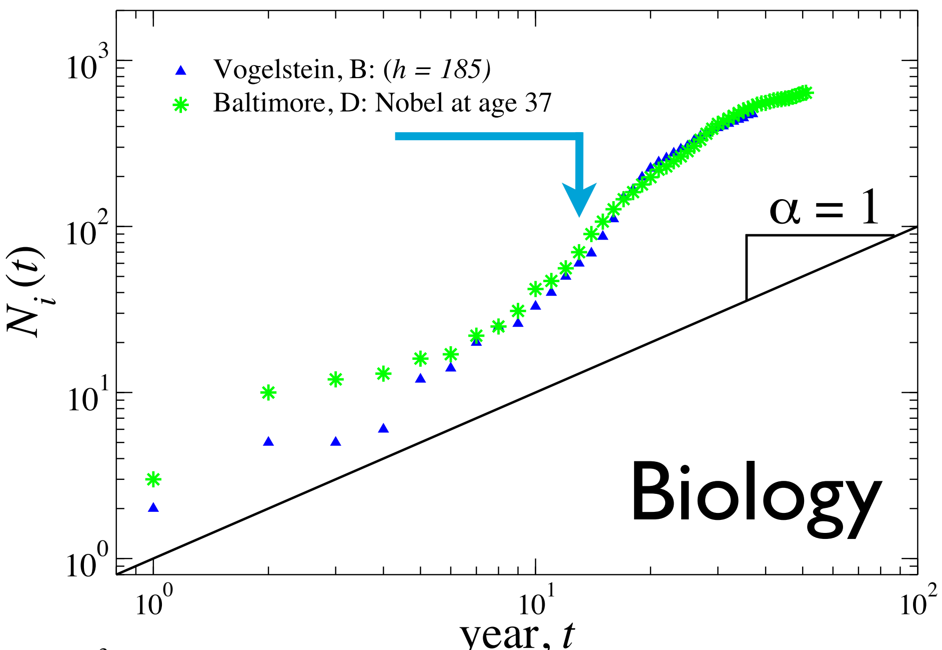


For a given h, a large β value corresponds to a larger total citations,

$$C_i \sim h^{1+\beta},$$

which is a proxy for career publication impact

The career trajectory in science: a tale of knowledge, collaboration, and reputation spillovers



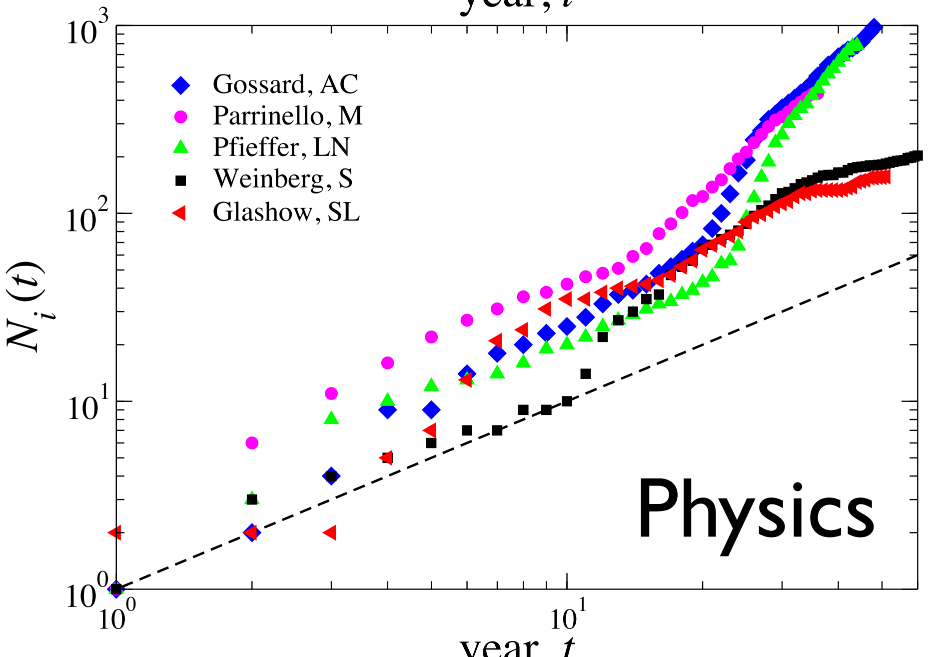
Annual production of individual i

$n_i(t)$ number of publications in year t

Cumulative production, a proxy for career reputation

$$N_i(t) \equiv \sum_{t'=1}^{t} n_i(t')$$

$$\approx A_i t^{\alpha_i}$$ ← for many prolific careers!

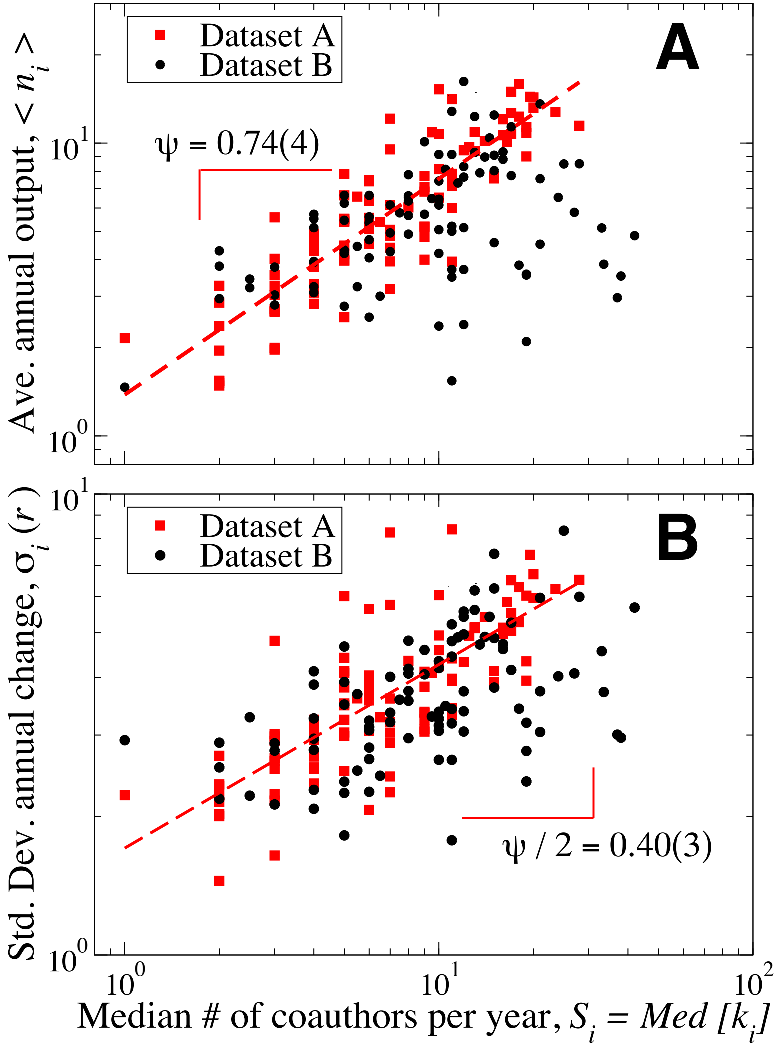


Careers are also subject to "shocks":

sometimes good: (Nobel Prize, promoted to lab director)

sometimes bad: accused of scientific fraud

Collaboration Radius and team efficiency



Towards a micro-level production function:

$$\langle n_i \rangle \sim S_i^{\psi}$$

average number of publications per year

S_i is median number of coauthors per year

Output change ("growth fluctuation"),

$$r_i(t) \equiv n_i(t) - n_i(t - \Delta t)$$

std. deviation of publication change

$$\sigma_i(r) \sim S_i^{\psi/2}$$

team efficiency parameter ψ

Not surprisingly, there is a decreasing marginal returns with increasing collaboration radius, likely attributable to team management inefficiencies, however inefficiencies aggregate sub-linearally, $\psi < 1$