Multilevel networks in science: from individual careers to Europe

Econophysics and Networks Across Scales

Lorentz Center
International Center for workshops in the Sciences

27 May 2013 through 31 May 2013

A. Chessa,
A. Morescalchi,
F. Pammolli,
O. Penner,
M. Riccaboni.

Alexander M. Petersen
IMT Institute for Advanced Studies, Lucca Italy

Institutions Markets Technologies
IMT
INSTITUTE FOR ADVANCED STUDIES LUCCA

Outline

- **Europe:** using "big data" to quantify the rate of cross-border R&D integration in Europe as compared to non-EU countries
- **Careers:** the intensity of collaboration within superstar careers
  - rank-coauthorship profile
  - collaboration life-cyce
  - "Batman & Robin(s)"
  - implications in the "big science" era

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

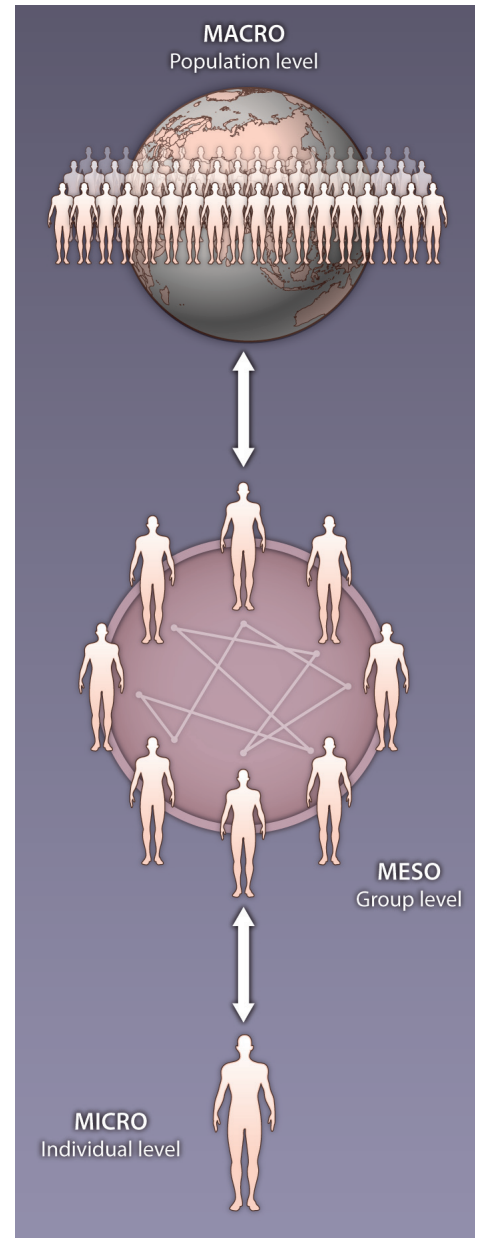
Institutional factors

- Country
- Institution
- Lab / Team

Behavioral factors

- Individual
- Paper

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



Institutions Markets Technologies
IMT
INSTITUTE FOR ADVANCED STUDIES LUCCA

Quantifying Institutional Impact on Science

Is Europe Evolving Toward an Integrated Research Area?

Geopolitical borders

EU borders

EU research initiatives are aimed at forming an integrated and competitive ERA

- 1998 : 5th Framework Programme
- 2000 Lisbon European Council
- ....ongoing

} - directed funding
- increased high-skill labor mobility
- streamlined trans-national innovation policies

cost
EUROPEAN COOPERATION
IN SCIENCE AND TECHNOLOGY

1971-present

European transnational program to build cooperation in science and technology. Funds and promotes integration via mobility and cross-border workshops

AN OPEN SPACE FOR KNOWLEDGE

The European Research Area is composed of all research and development activities, programmes and policies in Europe which involve a transnational perspective. Together, they enable researchers, research institutions and businesses to increasingly circulate, compete and co-operate across borders. The aim is to give them access to a Europe-wide open space for knowledge and technologies in which transnational synergies and complementarities are fully exploited.

EU initiatives towards cross-border mobility & collaboration

EUROPEAN COMMISSION

Community research

Framework programmes (FP):

STREP (specific targeted research project):
min 3 partners from three different member/ associated states

NoE (network of excellence) :
min 3 partners from 3 different countries

Mobility Actions

Ready to spend some time abroad ?

Introduction: In FP7, Marie Curie Actions have been regrouped in the PEOPLE programme. The "Marie Curie in a nutshell" document gives you an overall introduction to the Marie Curie Actions.

Objectives: With the Marie Curie Actions, PEOPLE aims to increase the trans-national mobility of researchers. It encourages European researchers to stay in Europe, and attract to Europe researchers from the entire world.

Thematic domains: Marie Curie Actions are cross cutting, e.g. they cover any scientific topic. They also follow a bottom-up approach: the research topic is chosen by researchers. Find more information in the Work Programme 2013.

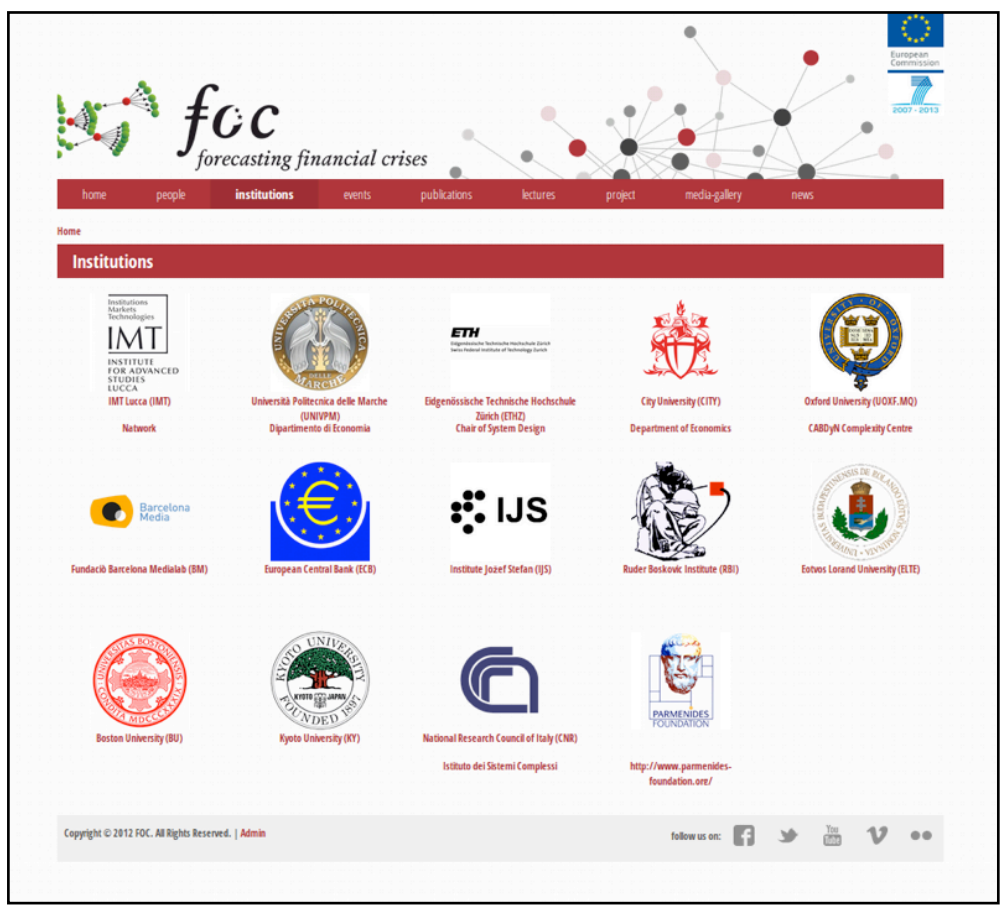
Target public: Individual researchers, Universities, Universities of Applied Sciences, Swiss Federal Institutes of Technology (EPFL and ETHZ), SMEs and large companies.

Calls: You'll find here all the open calls; expected calls; past calls.

Events: Events offered for preparation to the Marie Curie 2013 calls will be published on our events page.

Four categories of Marie Curie Actions:

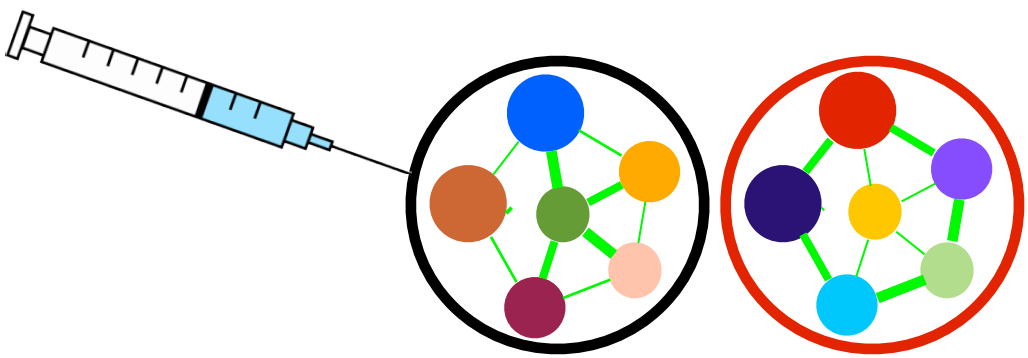
- Life-long training: Actions aiming at supporting *experienced researchers* in acquiring new skills thank to stays abroad: Intra-European Fellowship (IEF), Career Integration Grant (CIG) and COFUND
- International dimension: Actions aiming at increasing the co-operation with third countries: International Outgoing Fellowship (IOF), International Incoming Fellowships (IIF) and International Research Staff Exchange Scheme (IRSES)
- Initial training: The Initial Training Network (ITN) aims to improve *early-stage researchers'* career perspectives in both public and private sectors
- Industry-Academia: The Industry-Academia Partnerships & Pathways scheme (IAPP) aims to increase the co-operation between private & private sectors



Quantifying the impact of EU policies on cross-border R&D integration

The EU spends ~ 10% of government level R&D budget on programs with explicit cross-border criteria, compared to < 1% for non-EU countries

EU initiatives aimed at an integrated ERA serve as a "treatment" (think a vaccine)



EU countries

non-*EU* countries

POLICYFORUM

EUROPEAN POLICY

Is Europe Evolving Toward an Integrated Research Area?

A. Chessa, A. Morescalchi, F. Pammolli,* O. Penner, A. M. Petersen, M. Riccaboni*

Despite efforts to integrate across borders, Europe remains a collection of national innovation systems.

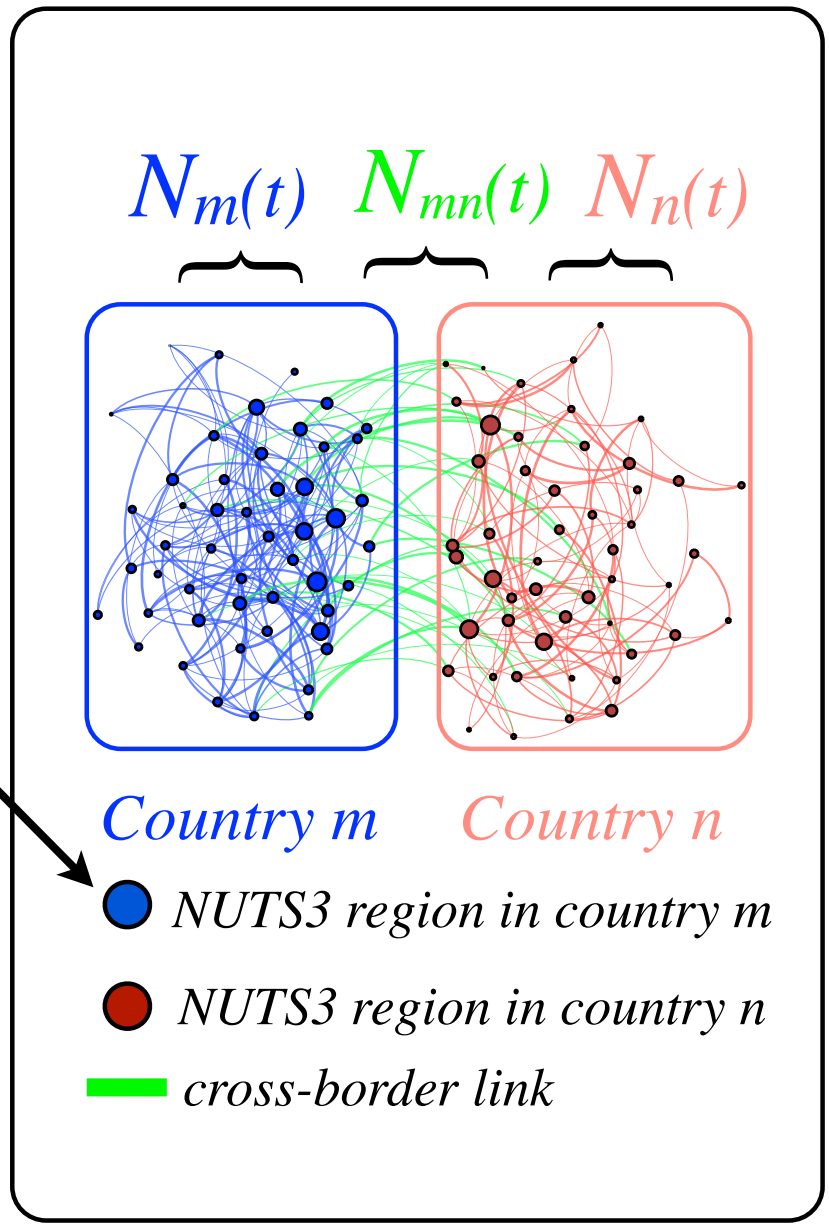
SCIENCE VOL 339 8 FEBRUARY 2013

DIMENSIONS OF SCIENCE COOPERATION

Methods and Data

NUTS3 regions ≈ province/ district/ county size

Complex networks approach



Geocoded data

(a) 2.4 million patents filed in the EPO and
(b) 0.26 million scientific publications

4 patent networks

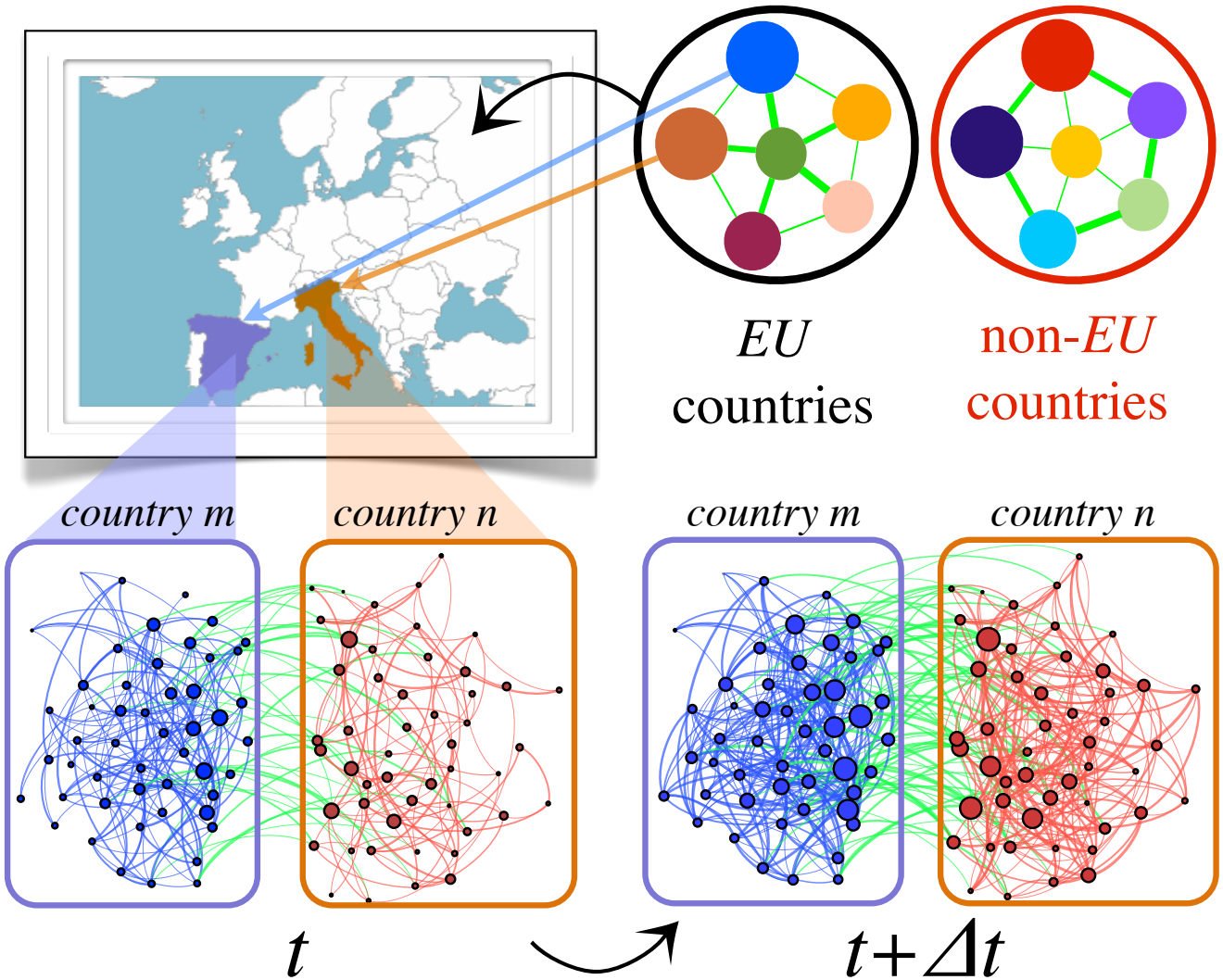
(i) co-inventor
(ii) co-applicant
(iii) citations
(iv) mobility

1 publication network

(v) co-author

Methods and Data

Complex networks approach



Geocoded data

(a) 2.4 million patents filed in the EPO and
(b) 0.26 million scientific publications

4 patent networks

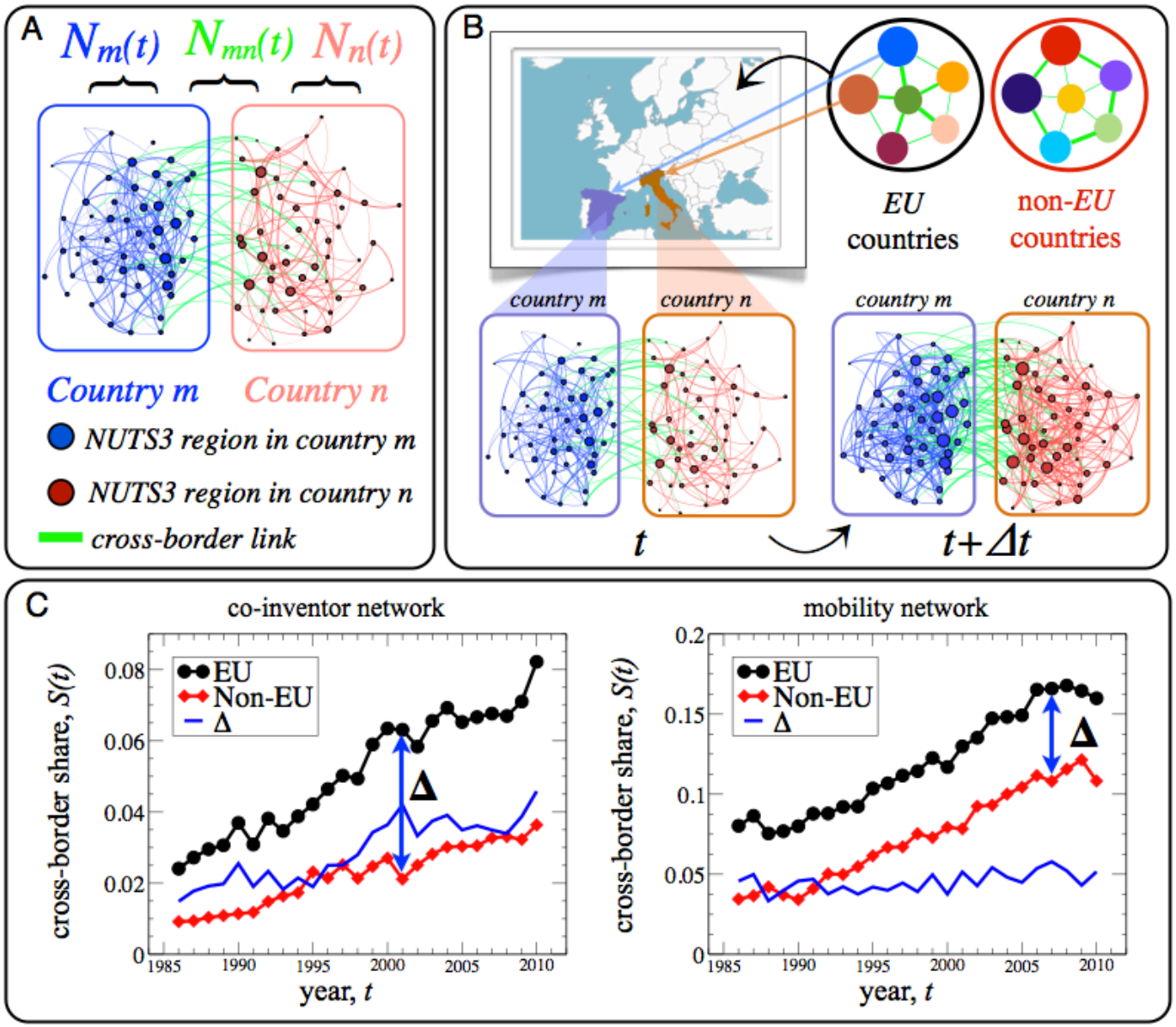
(i) co-inventor
(ii) co-applicant
(iii) citations
(iv) mobility

1 publication network

(v) co-author

Methods and Data

Complex networks approach



Geocoded data

(a) 2.4 million patents filed in the EPO and

(b) 0.26 million scientific publications

4 patent networks

(i) co-inventor

(ii) co-applicant

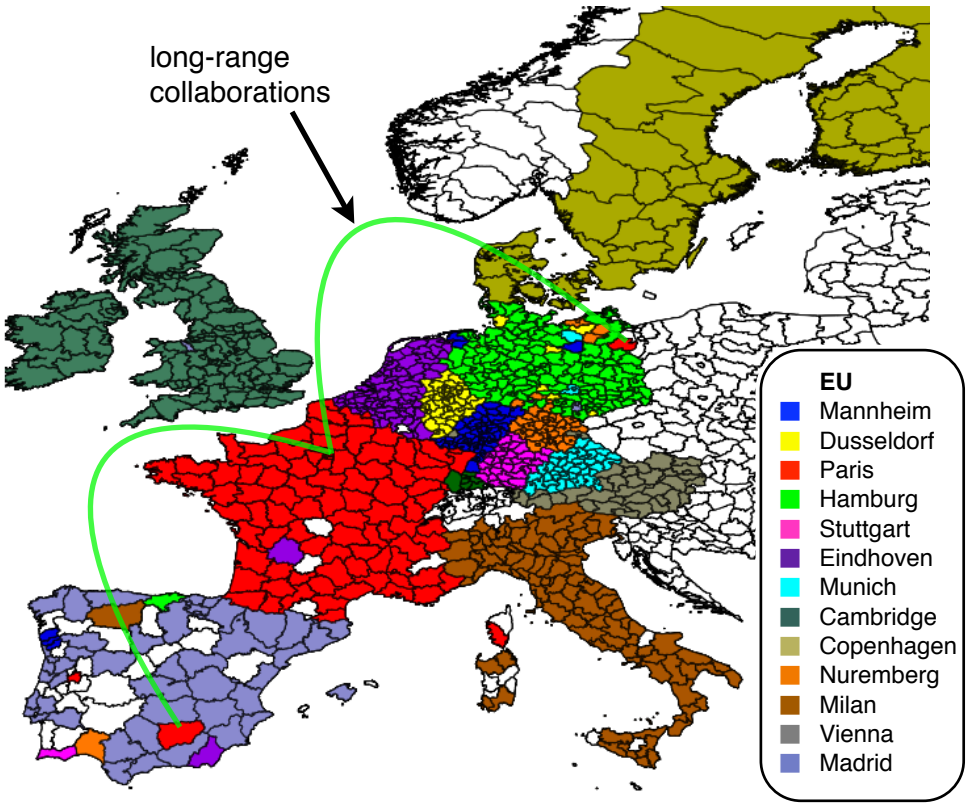
(iii) citations

(iv) mobility

1 publication network

(v) co-author

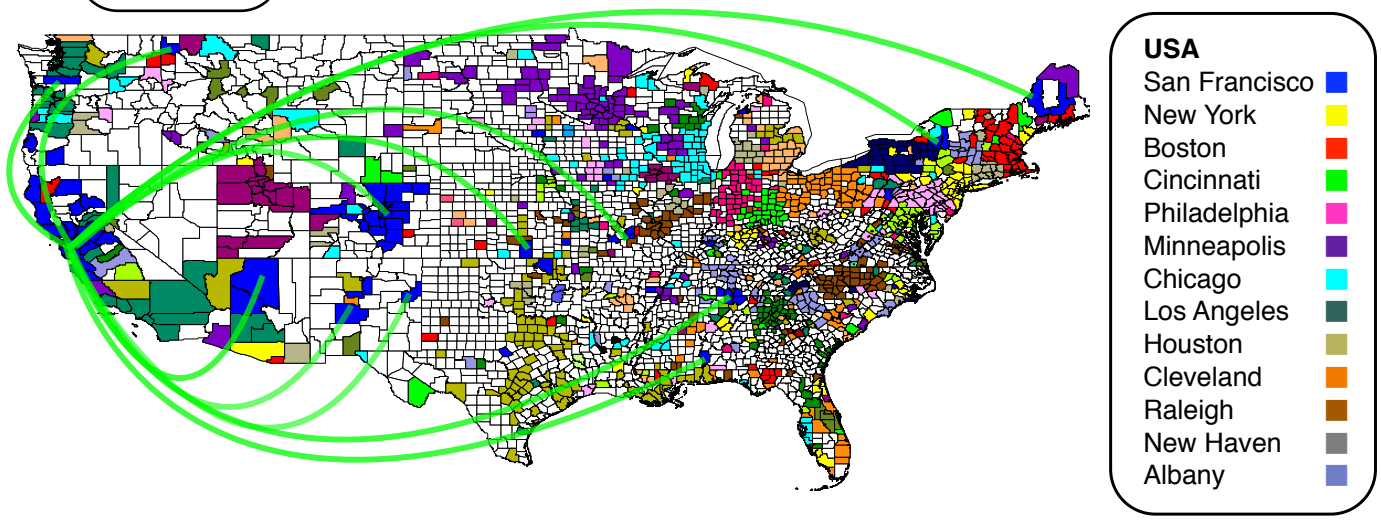
Comparing the community structure of the 2009 EU-15 and US coinventor networks



Pasteur's Quadrant: policy-oriented network science

Q1: Are the scientific borders in the EU any different than the geo-political borders?

Q2: has there been an intensification in cross-border R&D activity in Europe vis-a-vis other OECD countries (control group used for counterfactual data to quantify the "treatment effect")

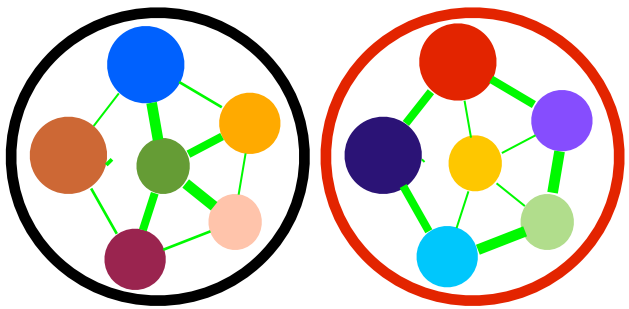


Community structure of the 2009 EU-15 and USA coinventor network.

Communities (color blocs) are labeled by their most-central region and were generated by iteratively aggregating NUTS3 regions into clusters of increasing size.

Quantifying the impact of EU policies on cross-border R&D integration

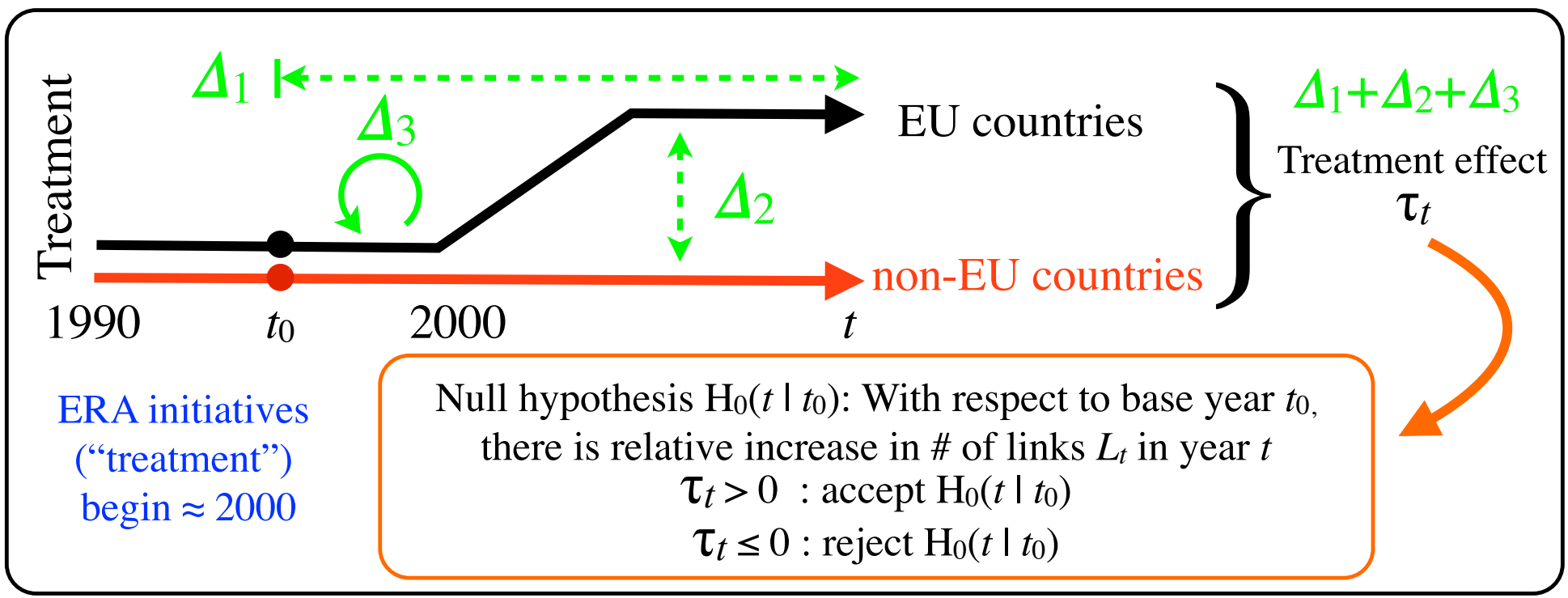
Q: do "treated" EU countries have different cross-border collaboration patterns than "untreated" non-EU countries



EU countries

non-*EU* countries

Econometric model controls for: borders, distance, technological distance, neighbors, EU vs non-EU

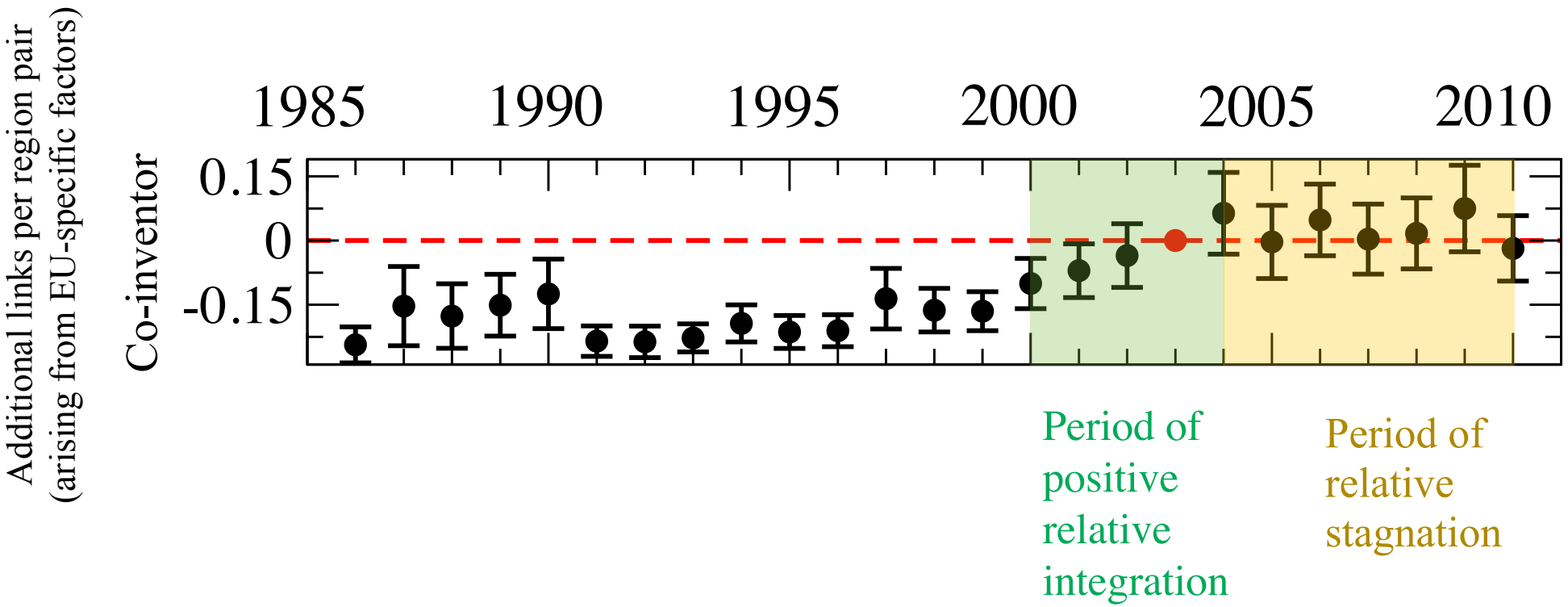


Econometric "treatment effect" model

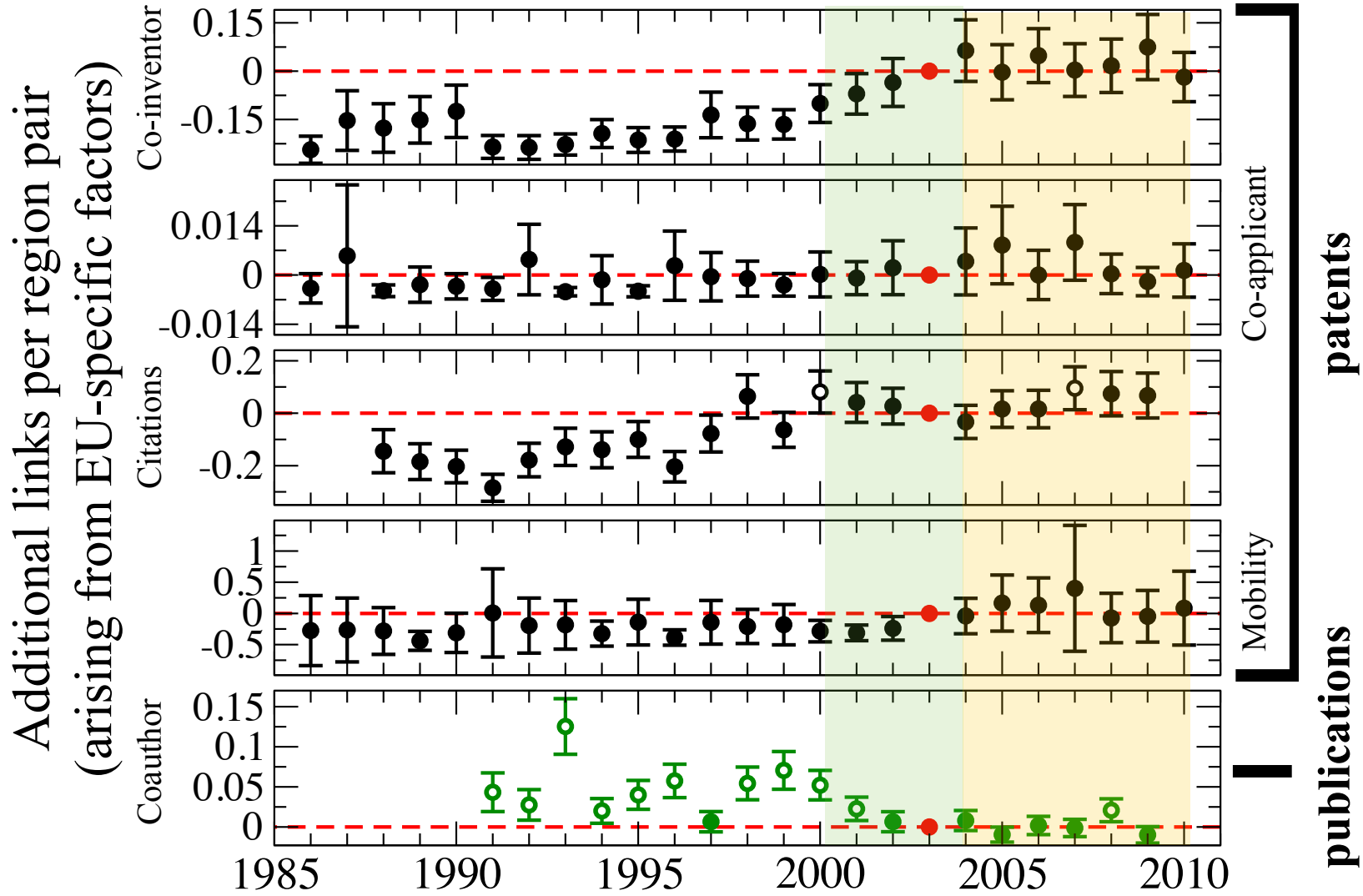
The "treatment effect" on an outcome variable is defined as the difference between

(i) the outcome actually observed under the treatment, and
(ii) the counterfactual, the outcome that would have been observed without treatment.

Under this treatment-effect framework, our analysis seeks to quantify the effect of EU institutional integration factors within the EU, by measuring the relative rate of cross-border links within a given network. Moreover, to isolate the signal arising only from EU factors, we must control for the global rate of cross-border integration.



Cross-validation using 5 different R&D networks



Evolution of European integration in five R&D networks. The *y* axis reports the additional number of cross-border links for an average pair of regions (i) relative to within-border links, (ii) due to EU-specific factors as compared with non-EU OECD countries, and (iii) relative to 2003 baseline year. Error bars indicate 95% confidence intervals for four different patent networks (black circles) and a scientific publication network (green circles). Open circles indicate statistically significant (0.05 level) positive deviations from the baseline year.

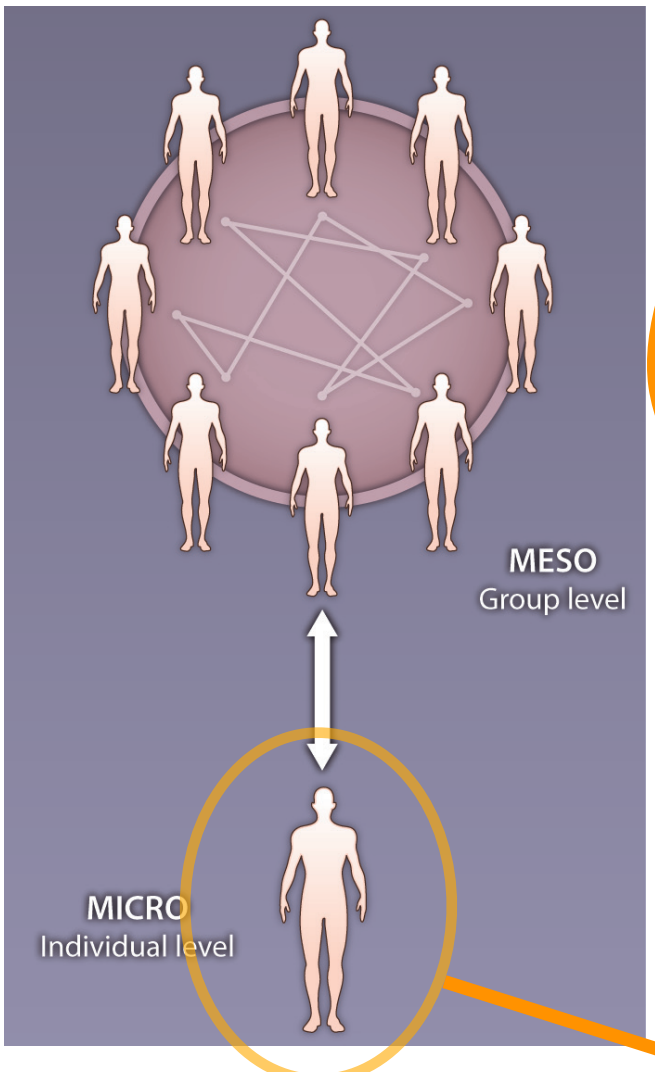
Dynamic duos in Science

=

Fun facts:

- Batman had multiple robins (non-overlapping, except for the "Earth-2" Robin who lives on a parallel universe earth)
- Batman also tag-teamed with Batgirl and several other "side-kicks"
- There is even an episode where Batman pretends to be Robin!

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:

451 publications

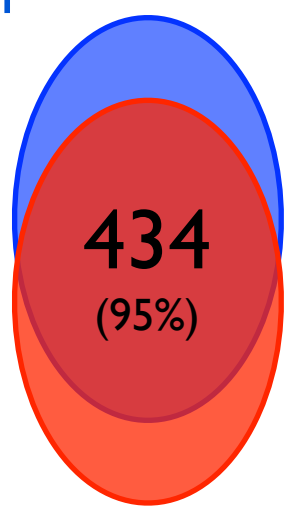
* **Michael Stuart Brown**
* **Joseph L. Goldstein**

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

Solo-artist strategy:

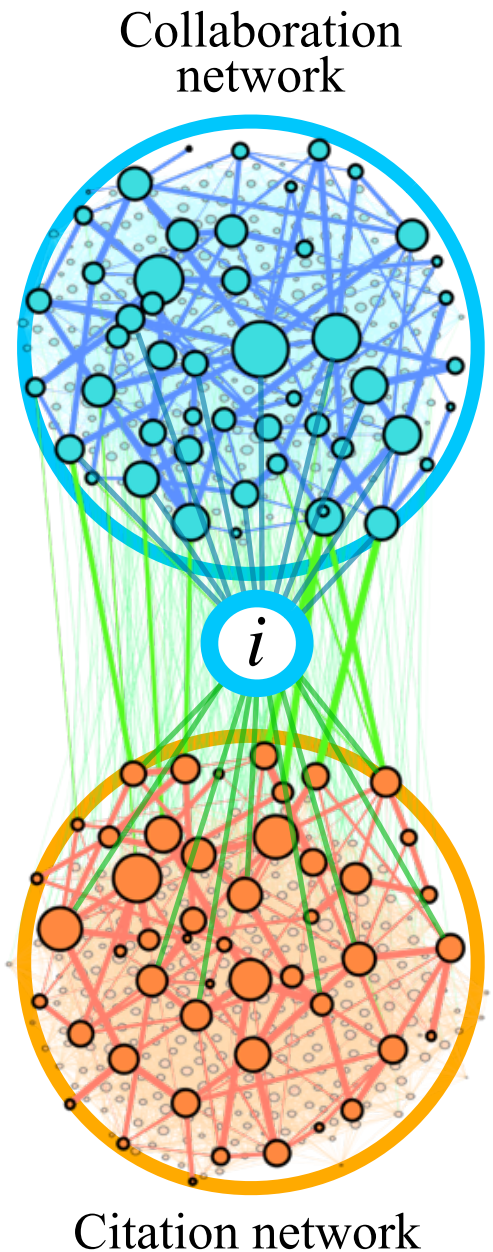
* **Marilyn Kozak** (also cell biologist)

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



458 publications

Scientific networks, spillovers, and career growth

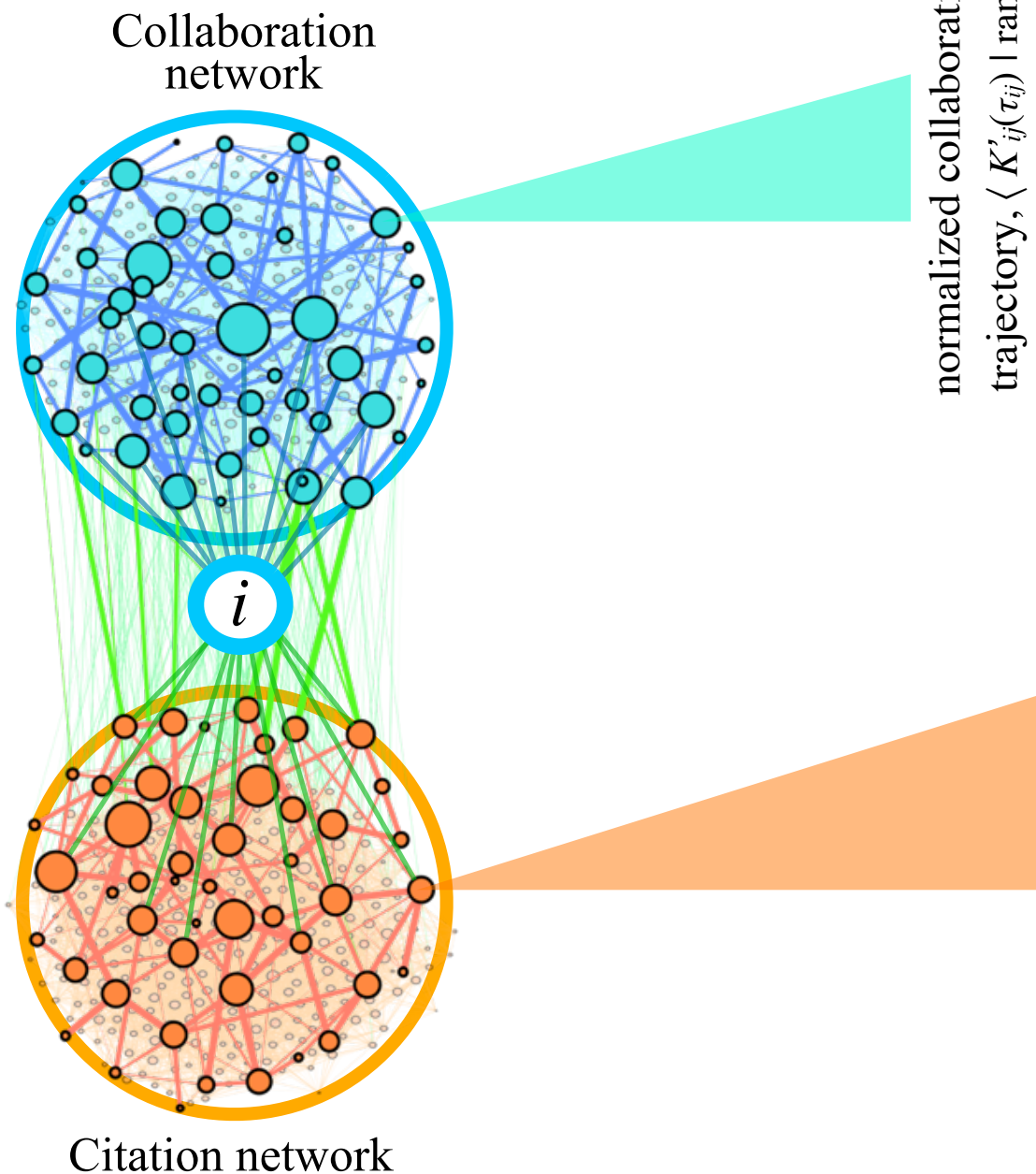


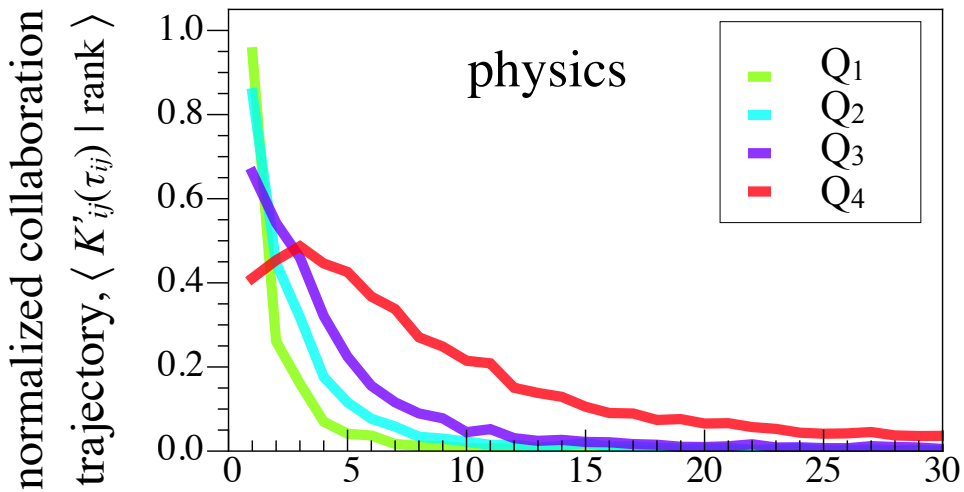
Complexity

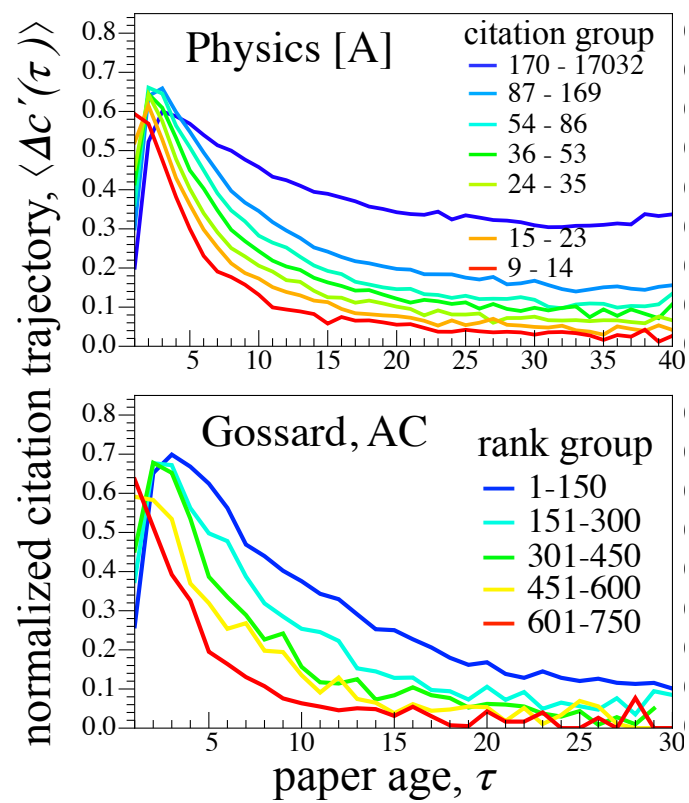
- coevolutionary system
- behavioral components
- embedded social processes
  - reputation
  - economic incentives (e.g. to collaborate)

Reputation and Impact in Academic Careers, ArXiv:1303:7274
A. M. Petersen, S. Fortunato, R. K. Pan, K. Kaski, O. Penner, M. Riccaboni, H. E. Stanley, F. Pammolli

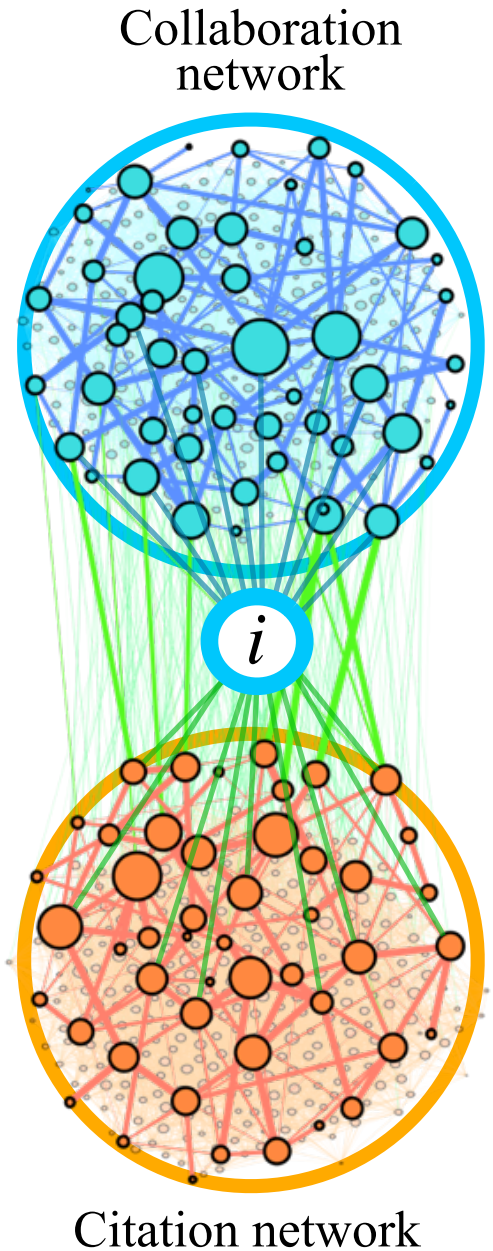
Dynamic network characterized by life-cycles







Even stars die!



Reputation and productivity "spillovers" are mediated by the network

SUPERSTAR EXTINCTION

PIERRE AZOULAY
JOSHUA S. GRAFF ZIVIN
JIALAN WANG

We estimate the magnitude of spillovers generated by 112 academic "superstars" who died prematurely and unexpectedly, thus providing an exogenous source of variation in the structure of their collaborators' coauthorship networks. Following the death of a superstar, we find that collaborators experience, on average, a lasting 5% to 8% decline in their quality-adjusted publication rates. By exploring interactions of the treatment effect with a variety of star, coauthor, and star/coauthor dyad characteristics, we seek to adjudicate between plausible mechanisms that might explain this finding. Taken together, our results suggest that spillovers are circumscribed in idea space, but less so in physical or social space. In particular, superstar extinction reveals the boundaries of the scientific field to which the star contributes—the "invisible college."

"Greater is the merit of the person who facilitates the accomplishments of others than of the person who accomplishes himself."
Rabbi Eliezer, *Babylonian Talmud, Tractate Baba Bathra 9a*

The Quarterly Journal of Economics, May 2010

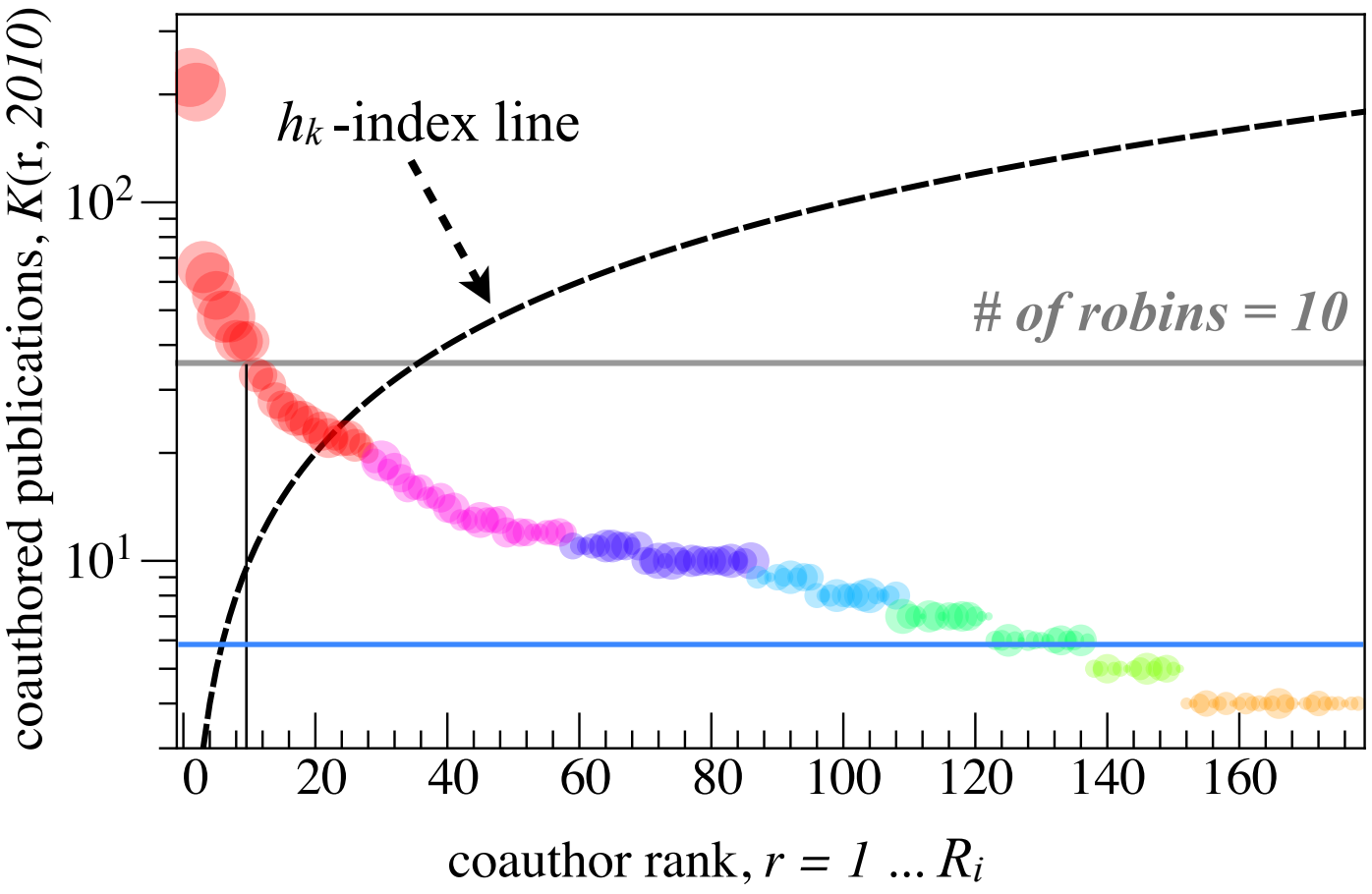
What is the intensity of collaboration around a "superhero"?

Stanley, HE

 # publications, N_i (2010) = 909

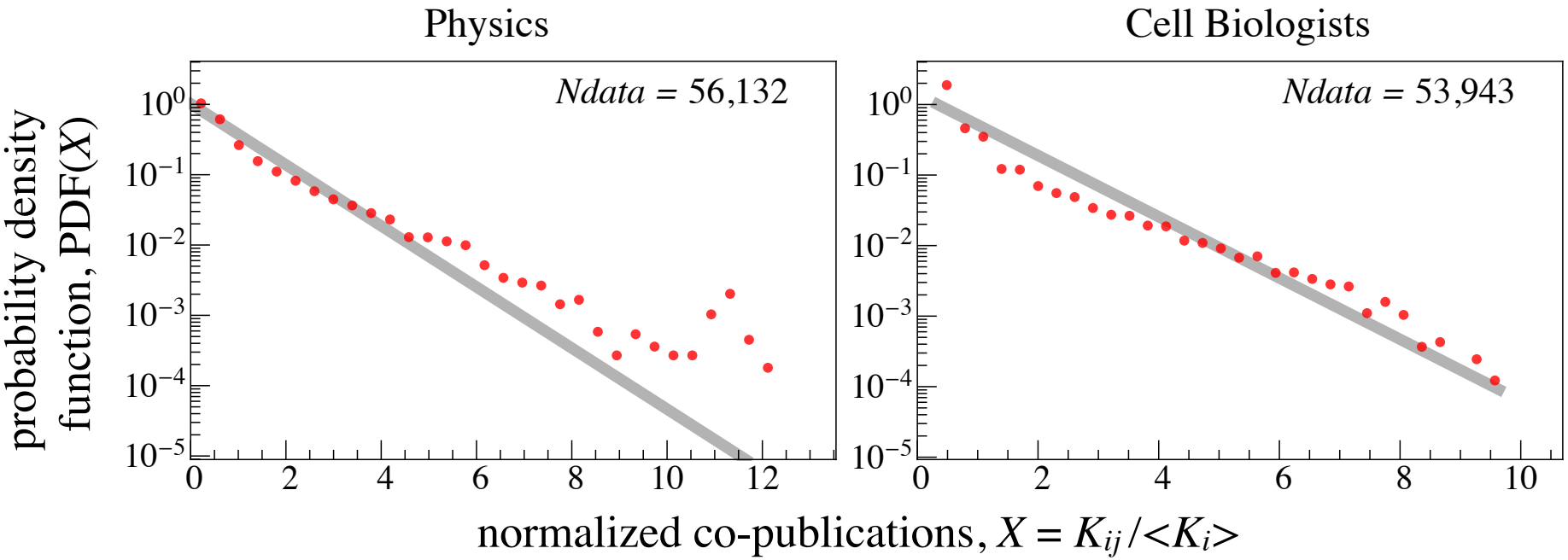
R_i = 541 coauthors

The average #, $\langle K_i \rangle$ = 5.7 papers



rank		K_{ij}	*%*
1	HAVLIN, S	223	*.25*
2	BULDYREV, SV	203	*.22*
3	AMARAL, LAN	66	*.07*
4	SCIORTINO, F	62	*.06*
5	IVANOV, PC	55	*.05*
6	GOLDBERGER, AL	48	
7	PENG, CK	48	
8	GOPIKRISHNAN, P	41	
9	PLEROU, V	41	
10	STARR, FW	41	
11	DOKHOLYAN, NV	33	*.04*
12	PAUL, G	33	
13	BUNDE, A	31	*.03*
14	GIOVAMBATTISTA, N	28	
15	MAKSE, HA	27	
16	CONIGLIO, A	26	
17	URBANC, B	25	
18	CRUZ, L	25	
19	SCALA, A	24	
20	LARRALDE, H	23	
21	MANTEGNA, RN	23	
22	POOLE, PH	22	*.02*
23	ANDRADE, JS	22	
24	VISWANATHAN, GM	22	
25	KUMAR, P	22	
26	CHANG, TS	21	
27	SUKI, B	21	
28	PODOBNIK, B	20	
29	MILOSEVIC, S	19	
30	SIMONS, M	19	

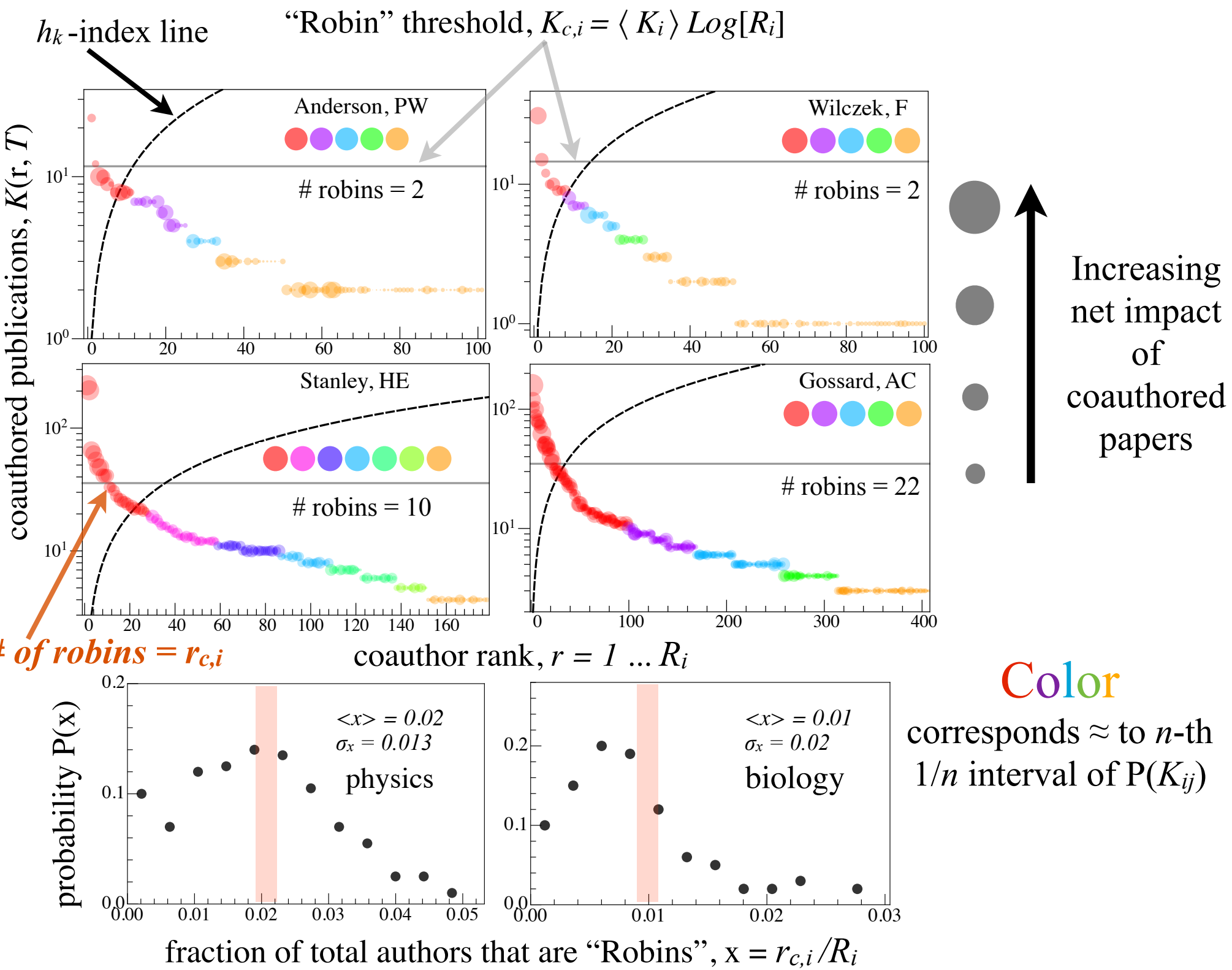
Is there a characteristic collaboration intensity scale?



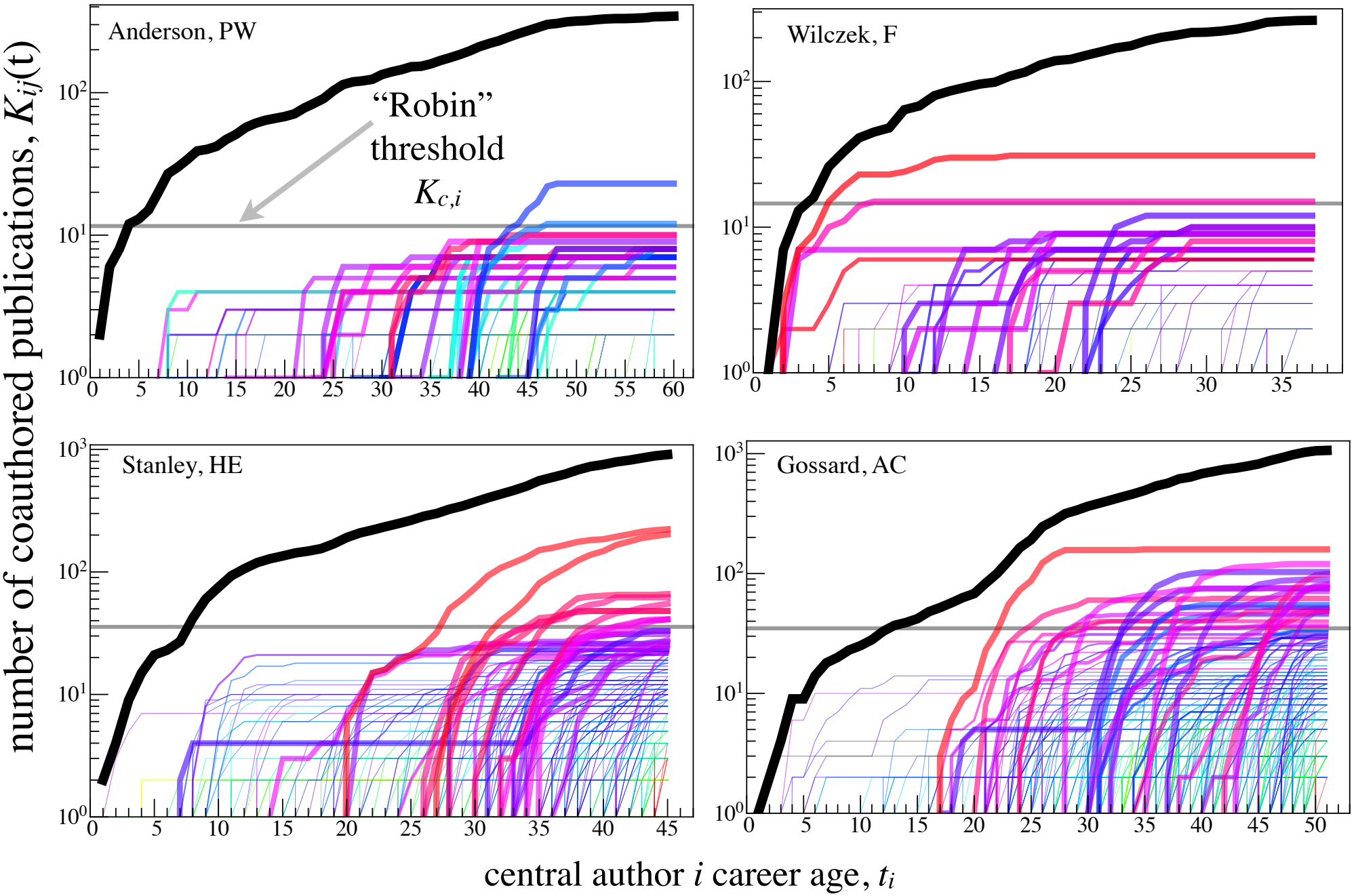
In order to aggregate across careers with varying coauthorship patterns, we use the normalized variable $X = K_{ij} / \langle K_i \rangle$

P(X) is well-described by an exponential distribution

Hence, we can define the extreme value
"Robin" threshold, $K_{i,c} = \langle K_i \rangle \, Log[R_i]$

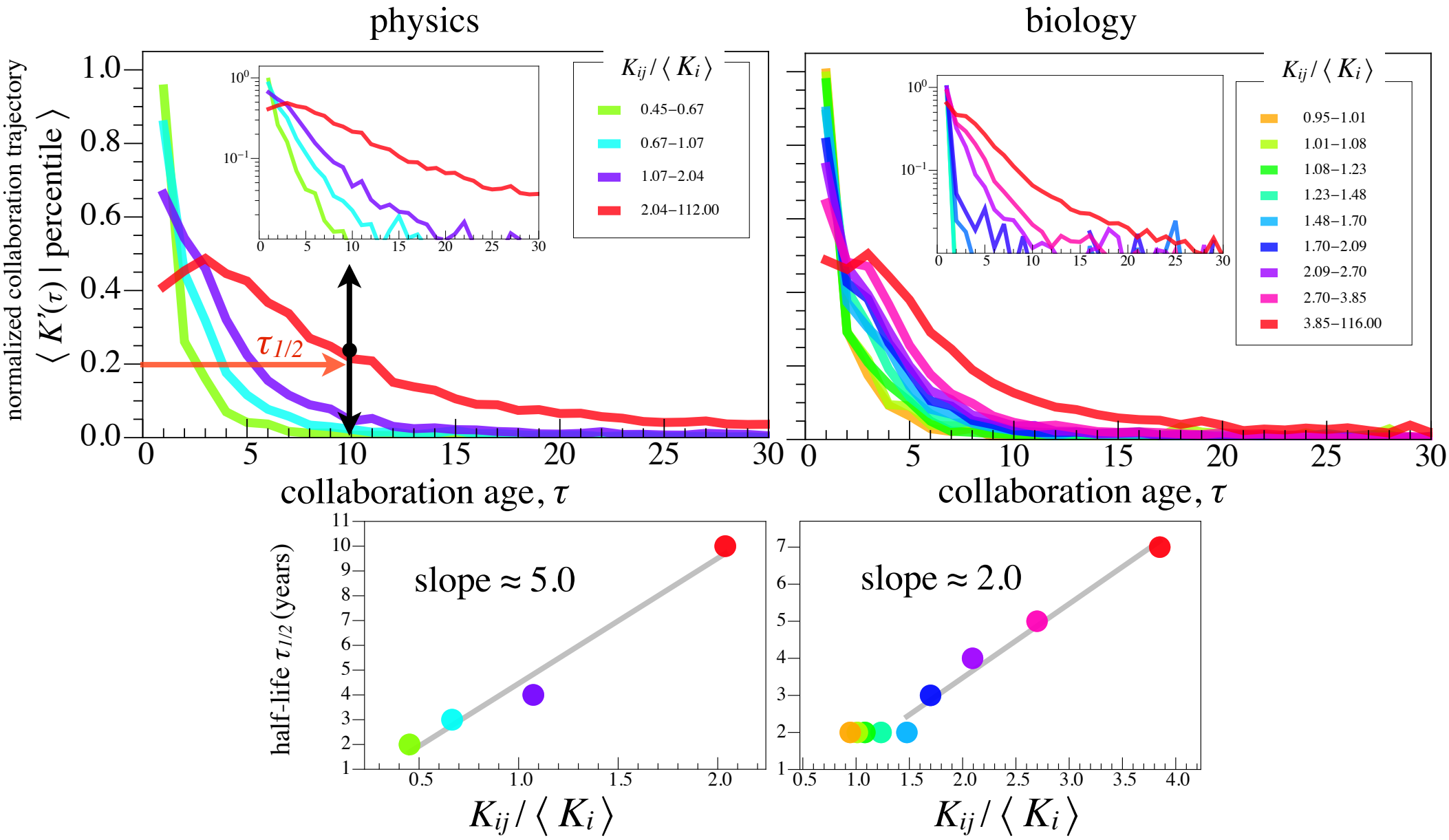

h_k-index line
"Robin" threshold, $K_{c,i} = \langle K_i \rangle Log[R_i]$
Anderson, PW
robins = 2
Wilczek, F
robins = 2
Stanley, HE
robins = 10
Gossard, AC
robins = 22
coauthored publications, $K(r, T)$
# of robins = $r_{c,i}$
coauthor rank, $r = 1 \ldots R_i$
Increasing net impact of coauthored papers
Color
corresponds ≈ to n-th $1/n$ interval of P(K_{ij})
probability P(x)
<x> = 0.02
σ_x = 0.013
physics
<x> = 0.01
σ_x = 0.02
biology
fraction of total authors that are "Robins", x = $r_{c,i}/R_i$

Evolution of the collaboration profile over time



color represents net impact of papers published together

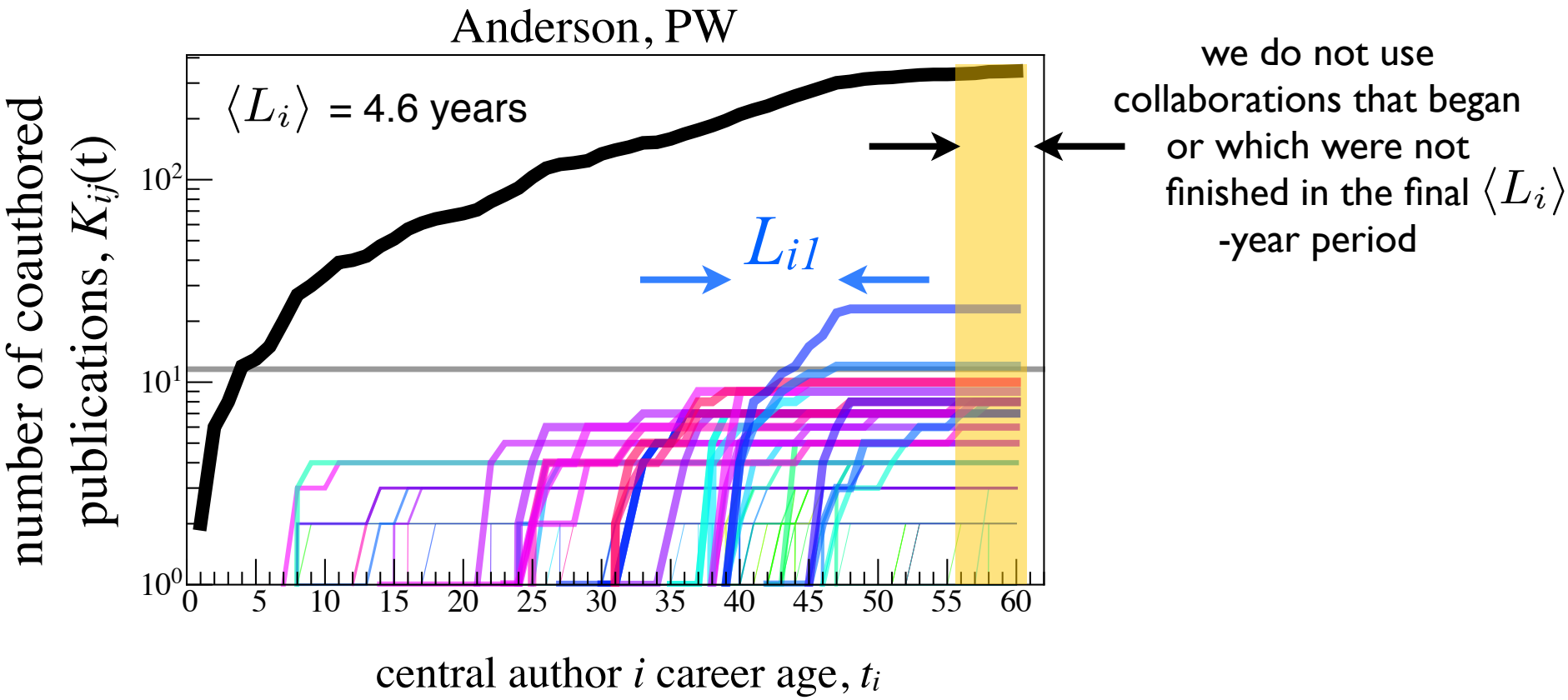
Average properties of the collaboration life-cycle $K_{ij}(\tau_{ij})$



$\tau_{1/2}$ increases faster per percentile in physics than biology

Biology: the state of the labor market in biomedical fields is one of a pyramid scheme, characterized by an indentured "postdoc" class [P. Stephan, How Economics Shapes Science, Harvard Univ. Press, 2012]

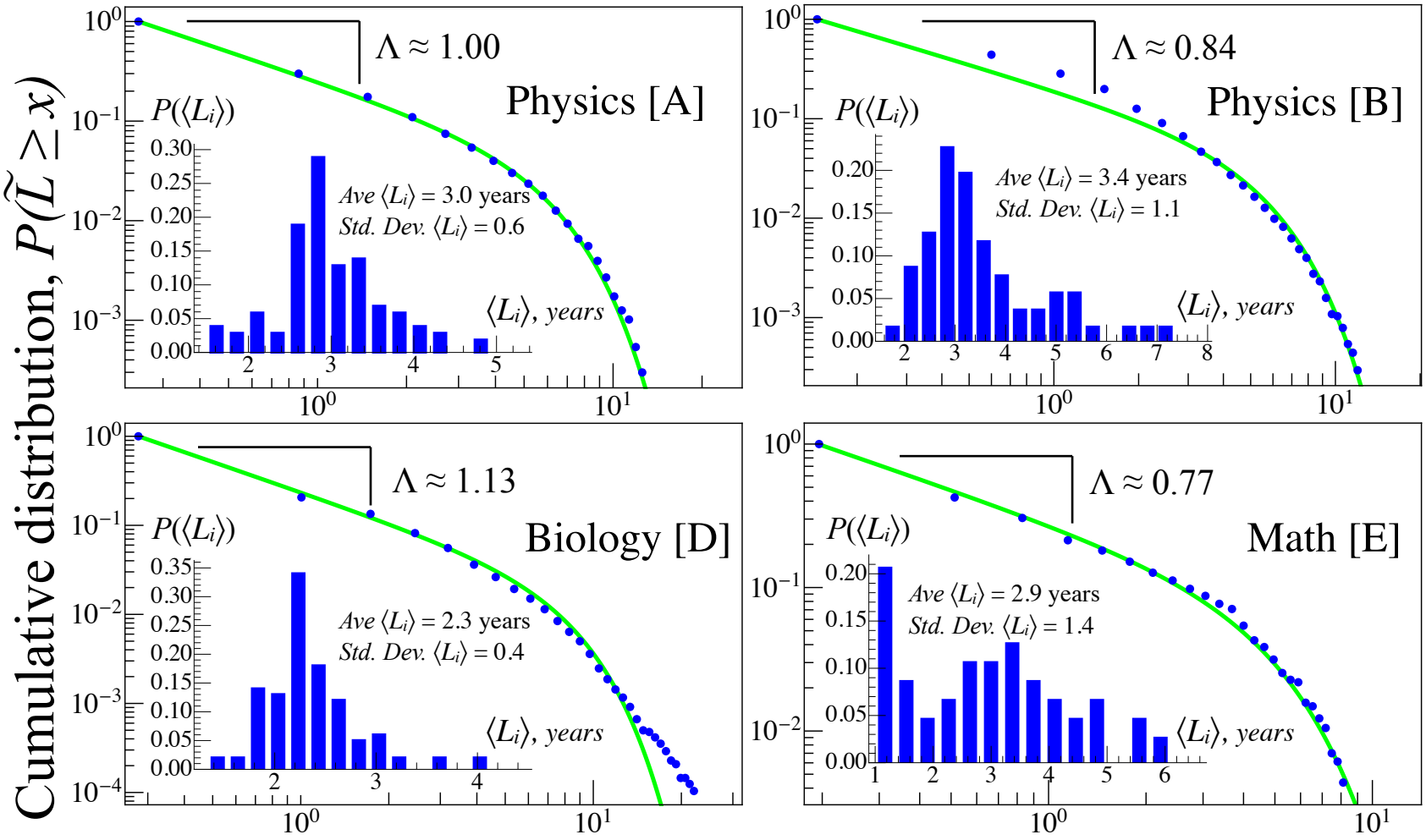
Measuring the characteristic collaboration longevity



$L_{ij} \equiv t_{ij}^f - t_{ij}^0 + 1$ between author i and one of his/her coauthors j, using their first joint publication appearing in year t_{ij}^0 and their last joint publication in year t_{ij}^f.

normalized longevity $\quad L \equiv L_{ij}/\langle L_i \rangle$

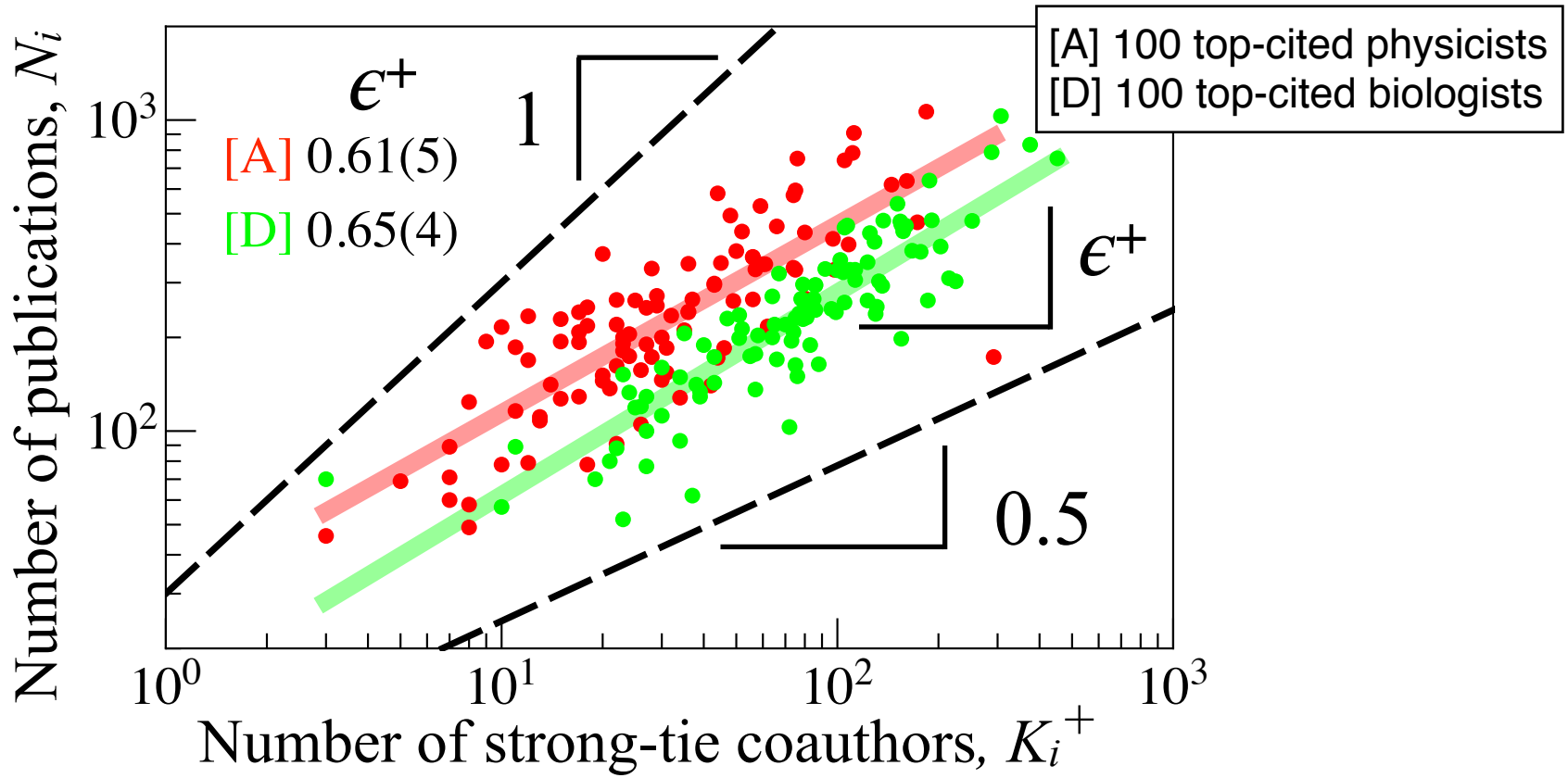
Spurious collaborations & common longevity patterns across disciplines



$$P(\tilde{L} \geq x) \propto x^{-\Lambda} \exp\left[-(x/x_c)^{1+\Lambda}\right]$$

70% of coauthorships last less than <*L*>
and only ~1% last longer than 7 <*L*>

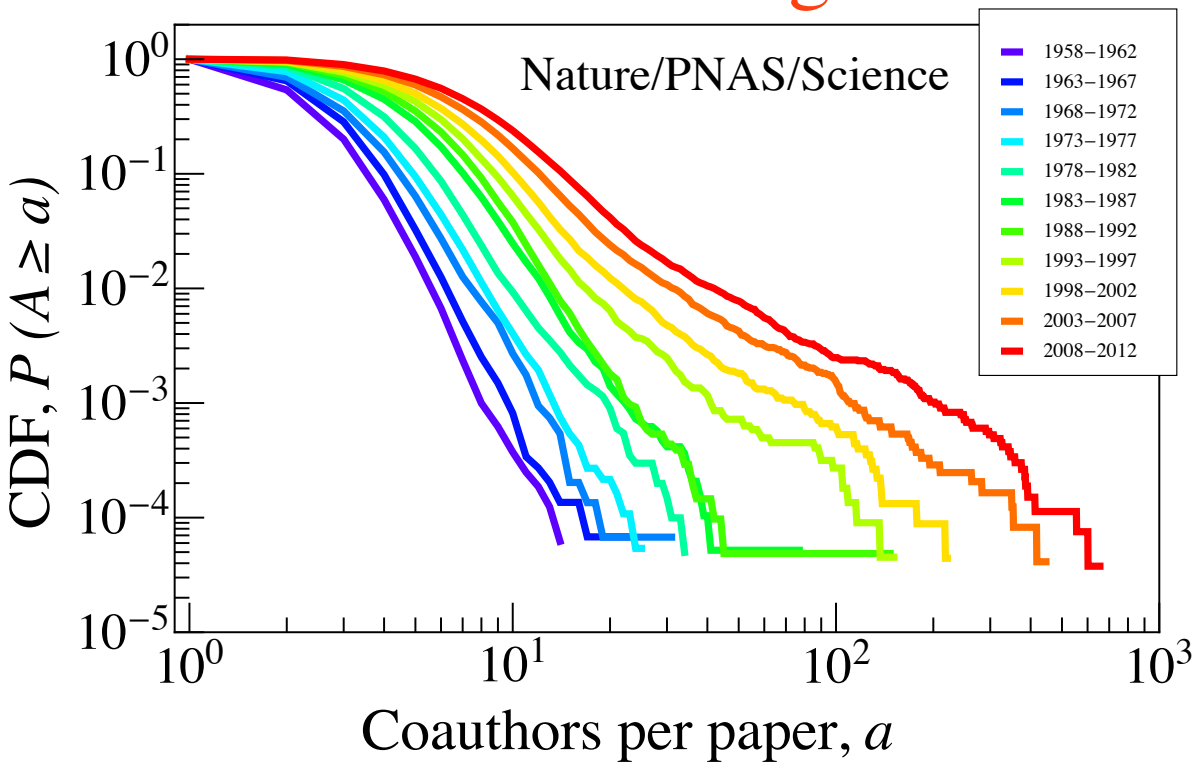
Collaboration in-efficiencies

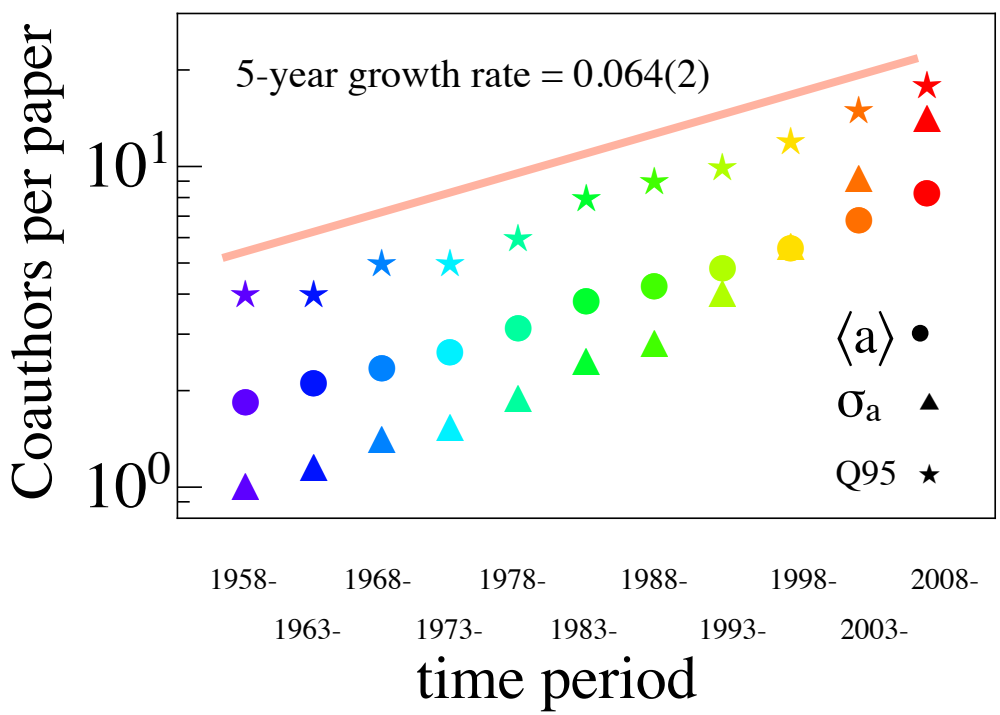


Productivity relation between total number of publications N_i and the number of strong ties K_i^+. The strong positive relation highlights the fundamental role of social processes underlying production in science. A sub-linear efficiency value $\varepsilon+ < 1$ is indicative of team inefficiencies which are here shown to arise from factors above and beyond spurious collaborations with $L_{ij} < \langle L_i \rangle$ for which there are costly training inefficiencies.

How to quantify the marginal value of strong-tie collaborations versus weak-tie collaborations ?

Emergence of "big science"



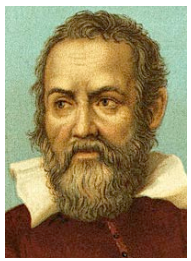


Q: how to "fairly" distribute credit in a system dominated by teams?

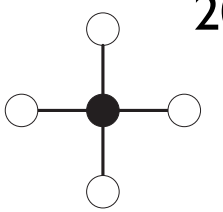
0.013 annual growth rate of the mean collaboration size ⟨a⟩ is consistent with the growth rate in the grad/postdoc populations

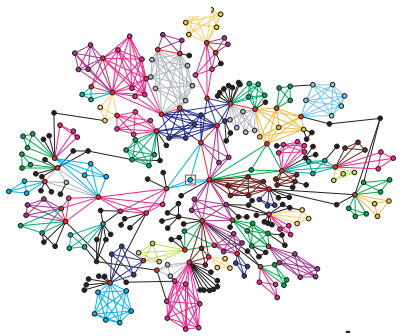
Institutional trends in Science

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



200+ years





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).

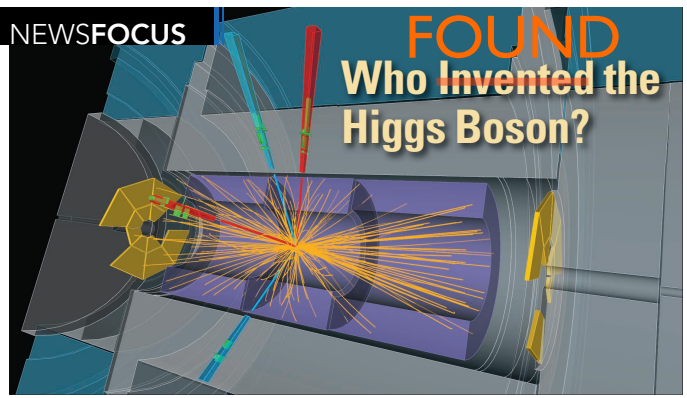
G. Palla, A.-L. Barabasi, T. Vicsek. Quantifying social group evolution. Nature 446, 664-667 (2007)

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
- redefining the role of teaching -vs- research faculty
- shifts in the competitive aspects of science, universities, and scientists: reputation tournaments in omnipresent competition arenas

Institutions Markets Technologies
IMT
INSTITUTE FOR ADVANCED STUDIES LUCCA

... increasing team size & changing incentive system



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

CITATION IMPACT 9 DECEMBER 2011 VOL 334 **SCIENCE**

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

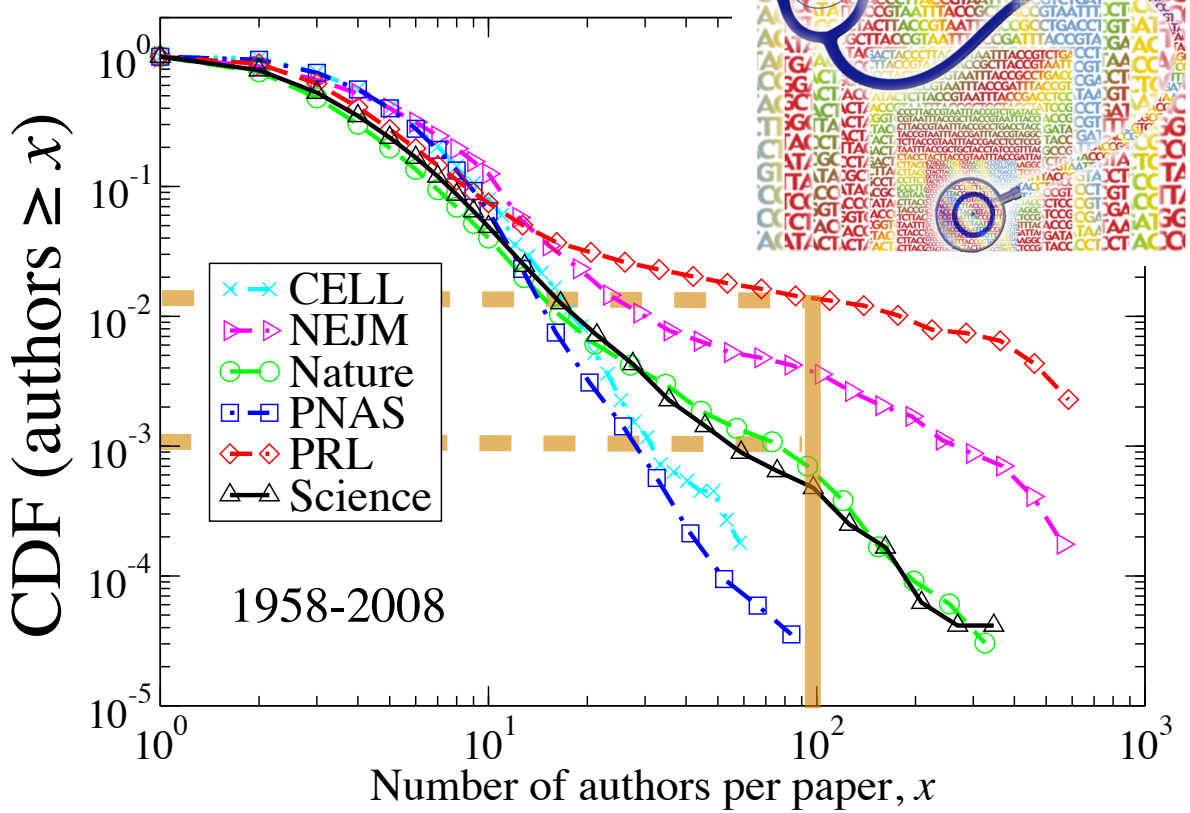
SCIENCE POLICY 5 AUGUST 2011 VOL 333 **SCIENCE**

Changing Incentives to Publish

Chiara Franzoni,[1] Giuseppe Scellato,[2,3] Paula Stephan[4,5,6]*

genome.gov
National Human Genome Research Institute
National Institutes of Health

ENCODE



Institutions Markets Technologies
IMT
INSTITUTE FOR ADVANCED STUDIES LUCCA

Food for thought

- Globalization of Science: What are the roles of institutions on the evolution of science? Are government policies aimed at increasing productivity having an impact, are the "treatments" efficiently allocated? Can we better understand the pipeline from Academia to Industry, and the impact of basic science on economic growth, using R&D productivity measures?

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

- Complex career dynamics: Knowledge, reputation, and collaboration spillovers are major factors leading to increasing returns along the scientific career trajectory. Finite collaboration life-cycles and extremely large team sizes have implications in the allocation of credit in science.

- 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

✳ 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).

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

✳ 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).

✳ The case for caution in predicting scientists' future impact, O. Penner, R. K. Pan, A. M. Petersen, S. Fortunato. Physics Today 66, 8-9 (2013)

✳ Is Europe Evolving Toward an Integrated Research Area, A. Chessa, A. Morescalchi, F. Pammolli, O. Penner, A. M. Petersen, M. Riccaboni. Science 339, 650-651 (2013)

✳ Reputation and Impact in Academic Careers, A. M. Petersen, S. Fortunato, R. K. Pan, K. Kaski, O. Penner, M. Riccaboni, H. E Stanley, F. Pammolli. ArXiv:1303:7274 (2013)

Thank You!

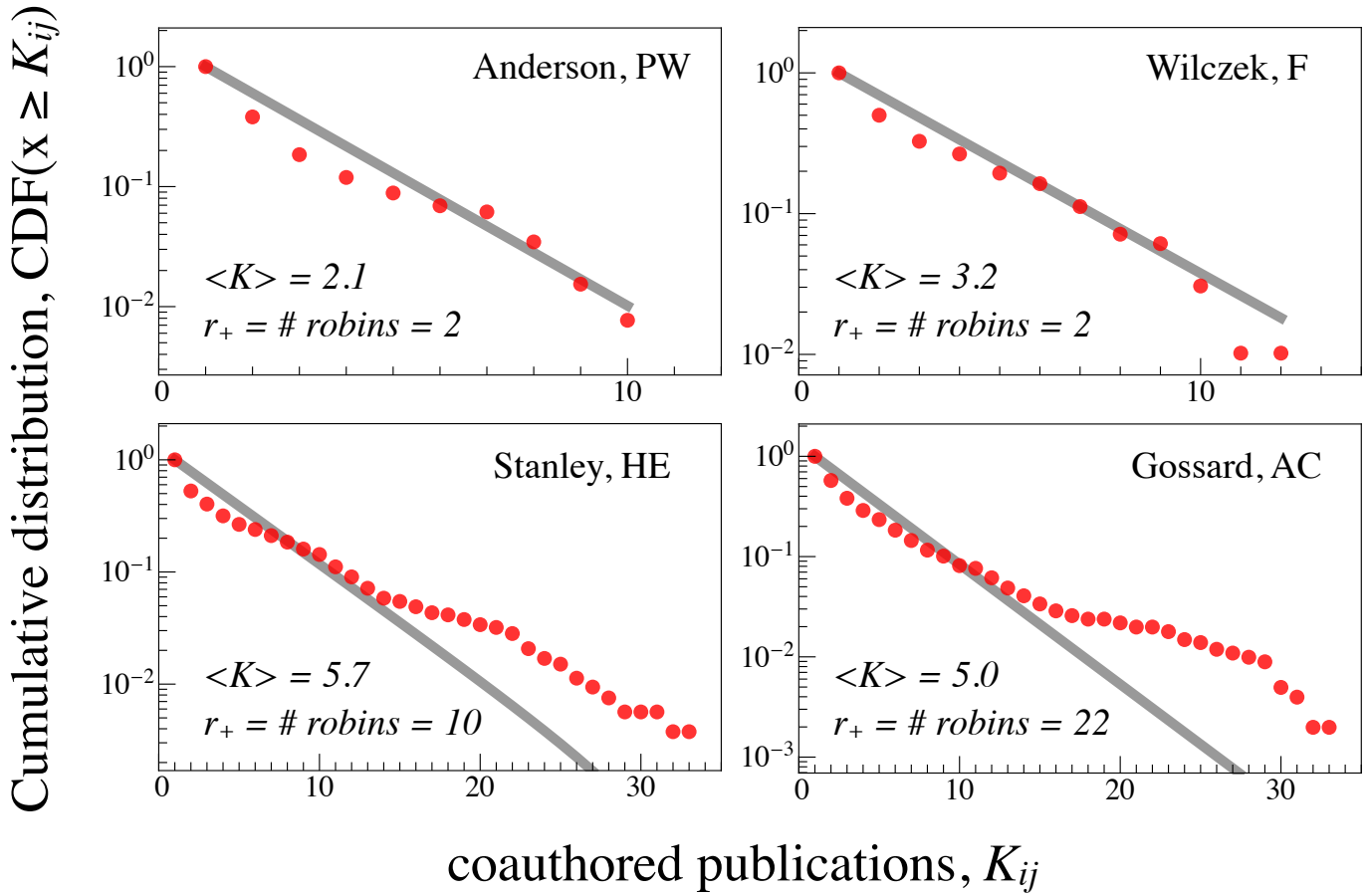
A special thanks to my collaborators: Santo Fortunato, Woo-Sung Jung, Fabio Pammolli, Raj Pan, Orion Penner, Massimo Riccaboni, Gene Stanley, Sauro Succi, Fengzhong Wang, and Jae-Sook Yang

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

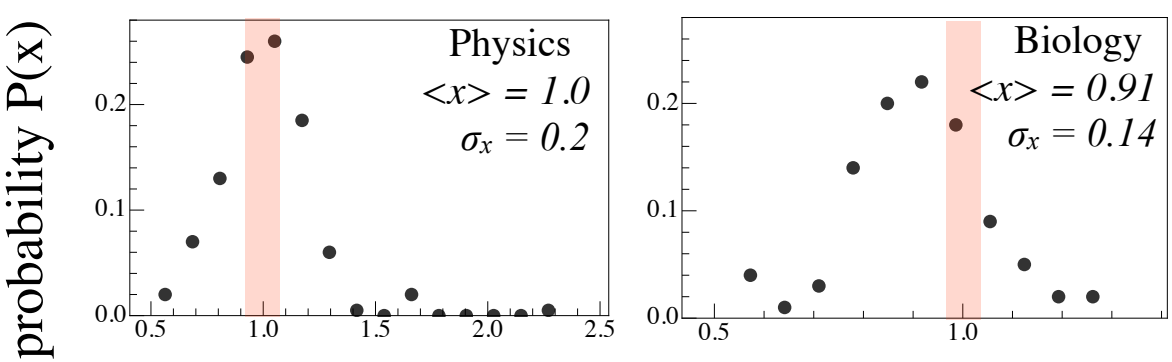
Title: Multilevel networks in science: from individual careers to Europe

Abstract:
Quantitative measures are becoming increasingly prevalent at all scales of scientific evaluation, from countries, to universities, departments, laboratories, and individuals. In this talk I will discuss the multi-level scientific networks that can be constructed from these output measures and the growth factors associated with the knowledge, human, and public capital spillovers which are facilitated by the network structure. Indeed, there is mounting evidence that both career growth and economic growth are intrinsically related to underlying features of co-evolving scientific networks. At the level of careers, I will discuss the role of strong ties in superstar careers, and the evolution of these ties longitudinally across the career. At the level of countries, I will discuss recent results obtained by analyzing 4 networks constructed from 2.4 million patent applications filed with the European Patent Office (EPO) over the 25-year period 1986-2010 [Science 339, 650-651 (2013)]. Combining econometric methods with network science we perform a comparative network analysis across time and between EU and non-EU countries to determine the “treatment effect” resulting from EU integration policies. Using non-EU countries as a control set, we provide quantitative evidence that, despite decades of efforts to build a European Research Area, there has been little integration above global trends in patenting and publication. This analysis provides concrete evidence that Europe remains a collection of national innovation systems.

Neglecting the "Robins" using the threshold, $K_{c,i} = \langle K_{ij} \rangle Log[R_i]$



How good is the exponential distribution approximation across all scientists?



Biologists have underestimated coeff. of determination due to excessive K_{ij} = 1

Visualizing the collaboration life-cycle $K_{ij}(\tau_{ij})$ of individual scientists