Economics 2880- 1: Introduction

Study of economics of science has developed rapidly in past decade or so bcs:

FACTOR 1)Sci-tech Illuminate major economic problems and Behavior

a) *Macro/economic growth accounting* leaves residual labeled TC that we need to shrink to do more than "observe"/measure growth. First efforts to fill it in were with Human Capital … changing the quality of labot but this analysis treats, say, an engineering degree was the same over time. Then an effort using R&D and patents to get some non-residual measure of technology – disruptive or otherwise – into aggregate production function.

STOCK OF USEFUL KNOWLEDGE as key input in long term.

b) *Micro-analysis labor and industrial organization/valuation of firms* also forced to examine role of science and engineering. Labor has problem of explaining why despite large increase in education/skills the wage distribution widened from 1970s-80s to present. First effort was "residual" skill-biased TC. But then wage distribution changed in a different way – polarization, Don't want too many epicycles/unobserved TC. Industrial org/valuation of firms has problem explaining firm value, which is nowadays 80% intangible capital – ideas, ability to monetize ideas, etc.

c)*The sluggish productivity problem* Despite the huge increase in number of Scientists &Engineers and in RD spending and in output– papers, patents, productivity growth had STAGNATED in advanced countries and declined in China, Korea, etc. Are there diminishing returns to S&E? is it ideas? Are Ideas Getting Harder to Find? Nicholas Bloom, Charles I. Jones, John Van Reenen, Michael Webb NBER Working Paper No. 23782

Is economy/organizations structured so cannot "absorb" new scientific-engineering ideas?
Is GDP and standard economic calculations failing to measure what matters?

STOCK OF USEFUL KNOWLEDGE

The magic of SOUK is that it i**s public good** – available freely – you don't deplete it when used; marginal cost (after someone has the idea and makes it public) = 0. If you make advance, why tell people and make it freely available? Three answers:

1)Publishing your great idea on how to turn greenhouse gases into Fountain of Youth pills gains you incredible prestige and fame. Robert Merton, founder of sociology of science, stressed that there is property right – the reward is **prestige related to priority** in discovery. In the day of the wealthy scientist – think Newton or Darwin – maybe that was enough to motivate people but today money incentives also enter.

Competition in science is that many people often seeking answer to same question. Someone takes credit for theorem/discovery – only way to get your priority is by making the discovery public. You establish *priority of discovery* by being **first** to communicate an advance in knowledge with reward to priority in *the recognition the scientific community gives you for being first*. Since multiple people have access to the public knowledge on which most discoveries rest, the system may produce ARMS RACE//PATENT WAR – with risk of multiple discovery but with benefit of fast disclosure.

Arrow: "The incentive compatibility literature needs to learn the lesson of the priority system; rewards to overcome shirking and free-rider problems need not be monetary in nature; society is more ingenious than the market." But many theorems/discoveries are named after the wrong guy. Maybe better to publicize idea; to establish yourself as part of scientific group and be the one who grabs the idea. Was Darwin really first in getting the key idea in evolution? Or was it the young guy?

Issue of who gets credit for scientific advance? More difficult today because papers have more authors than ever before and science more dependent on equipment. An funding goes largely to older PIs. If they make decisions, does this lower new ideas? If they protect younger scientists from bureaucracy, does it help produce new ideas? Must consider life-cycle of ideas and would need data on whether work was done in X's lab, funded by Z, protocol of senior person's name on paper.

I don't do research.
I don't know what the paper is about.
But I RAISED the money.
It's my paper

2)Turn science into patent and either company or leasing/selling patent If the real payoff is in innovative products or services that earn money, you *form the company* – Edison – and use patent or other legal way to limit what others can do with your knowledge. If the advance is based on new science/tech, to benefit from knowledge user must often make substantive investment in knowledge as well. Resurgence of Schumpeter type thinking – entrepreneurship – as key. Universities under Bayh-Dole 1980 Act allows federal government, which is the largest supporter of university research nationwide, to assign its patent rights to universities, which then have the obligation to pursue intellectual property protection and commercialization. Prior to Bayh-Dole, only a handful of universities had technology commercialization offices. Now there is a big association – Association of University Technical Managers AUTM is the non-profit leader in efforts to educate, promote and inspire professionals to support the commercialization of academic research that changes the world and drives innovation forward. "Our community is comprised of more than 3,000 members who work in more than 800 universities, research centers, hospitals, businesses and government organizations around the globe" Because most patents are useless and a few make lots of money, very uneven set of rewards.

3)Grants/Prize/Competitions: a) For work that will be done in future, important because it pays salary and equipment; b) for past work Nobel, Fields, for past work, numerous medals etc. Prizes for specific things as well.

https://www.xprize.org/ to attract new people/approaches to difficult problems; Create breakthrough results that are real and meaningful; compel teams around the world to invest the intellectual and financial capital required to solve seemingly intractable challenges.

The $5 million IBM Watson AI XPRIZE is a global competition challenging teams to develop and demonstrate how humans can collaborate with powerful AI technologies to tackle the world's grand challenges

The $10M ANA Avatar XPRIZE is a four-year global competition focused on the development of an Avatar System that will transport a human's sense, actions, and presence to a remote location in real time, leading to a more connected world.

The $20 million NRG COSIA Carbon XPRIZE is a global competition to develop breakthrough technologies that will convert CO_2 emissions from power plants and industrial facilities into valuable products like building materials, alternative fuels and other items that we use every day.

The $10 million Rainforest XPRIZE is a global 4-year competition to incentivize the development of integrated and advanced biodiversity assessment technologies that will yield previously undiscovered and meaningful insights into these magnificent environments. The competition will be followed by an additional two years of impact activities geared towards establishing a new bioeconomy for the next century. The prize respects the people of the forests and their culture, as well as the Convention on Biological Diversity and other national and international policies.

CANCELLATION OF The Archon Genomics XPRIZE, the second XPRIZE to be offered by the foundation, was announced October 4, 2006. The goal of the Archon Genomics XPRIZE was to greatly reduce the cost and increase the speed of human genome sequencing to create a new era of personalized, predictive and preventive medicine, eventually transforming medical care from reactive to proactive. The $10 million prize purse was promised to the first team that can build a device and use it to sequence 100 human genomes within 10 days or less, with an accuracy of no more than one error in every 100,000 bases sequenced, with sequences accurately covering at least 98% of the genome, and at a recurring cost of no more than $1,000 per genome.

Actual competition events were originally scheduled to take place twice a year with all eligible teams given the opportunity to make an attempt, starting at precisely the same time as the other teams. This was changed to a single competition scheduled for September 5, 2013 to October 1, 2013, which was **canceled on August 22, 2013.** The rationale for the change was articulated by the CEO: "Today, companies can do this for less than $5,000 per genome, in a few days or less – and are moving quickly towards the goals we set for the prize. For this reason, we have decided to cancel an XPRIZE for the first time ever."[21] A public debate concerning the validity and potential implications of the cancellation was published March 27, 2014.[22]

The **first Kavli prizes (http://kavliprize.org/)**were handed out in 2008 and, every two years since then, three Kavli Prizes have been awarded in the fields of astrophysics, nanoscience, and neuroscience; each winner, or team of winners, receives a million dollars. The Kavli is a partnership between The Norwegian Academy of Science and Letters, The Norwegian Ministry of Education and Research and The Kavli Foundation.

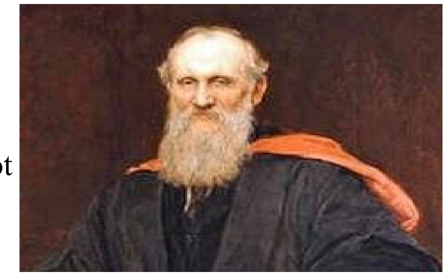
The Kavli Prize recognizes scientists for pioneering advances that shape our world and our understanding of existence at the very large, very small and very complex scales. From unexpected scientific breakthroughs to the creation of entirely new fields of research, the laureates selected are forward-looking and their work alters how we think about and interact with science today.
Presented every two years in the fields of astrophysics, nanoscience and neuroscience, e**ach of three international prizes consists of US$ 1 million**. Laureates are chosen by three committees whose members are recommended by six of the world's most renowned science societies and academies. Laureates are celebrated in Oslo, Norway in a ceremony presided over by His Majesty King Harald V, where they receive gold medals for their achievements.

The Breakthrough Prize, the largest monetary science prize in the world awarded in three categories: life sciences, fundamental physics, and math. Each prize is worth $3 million. In November, at the Breakthrough Prize ceremony, in San Francisco This is three times the amount of money awarded through the Nobel Prizes. EIGHTH ANNUAL BREAKTHROUGH PRIZE – **"THE OSCARS OF SCIENCE"** – CELEBRATES TOP ACHIEVEMENTS IN LIFE SCIENCES, PHYSICS & MATHEMATICS; AWARDS OVER $21 MILLION IN PRIZES AT GALA TELEVISED CEREMONY IN SILICON VALLEY
Evening's Theme: "Seeing the Invisible" Salutes Discoveries at Cosmic, Cellular and Quantum Levels.
2019 Prize in Fundamental Physics Awarded to Sergio Ferrara, Daniel Z. Freedman and Peter van Nieuwenhuizen.
2020 Prize in Fundamental Physics **Awarded to 347 members of the Event Horizon Telescope Collaboration.**
2020 Prizes in Life Sciences Awarded to David Julius, Virginia Man-Yee Lee, Jeffrey M. Friedman, F. Ulrich Hartl and Arthur L. Horwich.
2020 Prize in Mathematics Awarded to Alex Eskin.
New Horizons in Physics Prizes Awarded to Simon Caron-Huot, Xie Chen, Jo Dunkley, Lukasz Fidkowski, Michael Levin, Max A. Metlitski, Samaya Nissanke, Kendrick Smith and Pedro Vieira.
New Horizons in Mathematics Prizes Awarded to Tim Austin, Emmy Murphy and Xinwen Zhu.
T he awards brought together luminaries in the science and tech communities alongside celebrities, athletes, musicians and dozens of current and prior Breakthrough Prize laureates for a festive celebration of science.
A combined total of over $21 million was awarded this evening, in recognition of groundbreaking research achievements in Life Sciences, Fundamental Physics and Mathematics. Each Breakthrough Prize is $3 million, the largest monetary prize in science.

Smaller prizes … all kinds. https://en.wikipedia.org/wiki/List_of_general_science_and_technology_awards
Economist: Does linking money with fame have bigger effect than money or fame ?

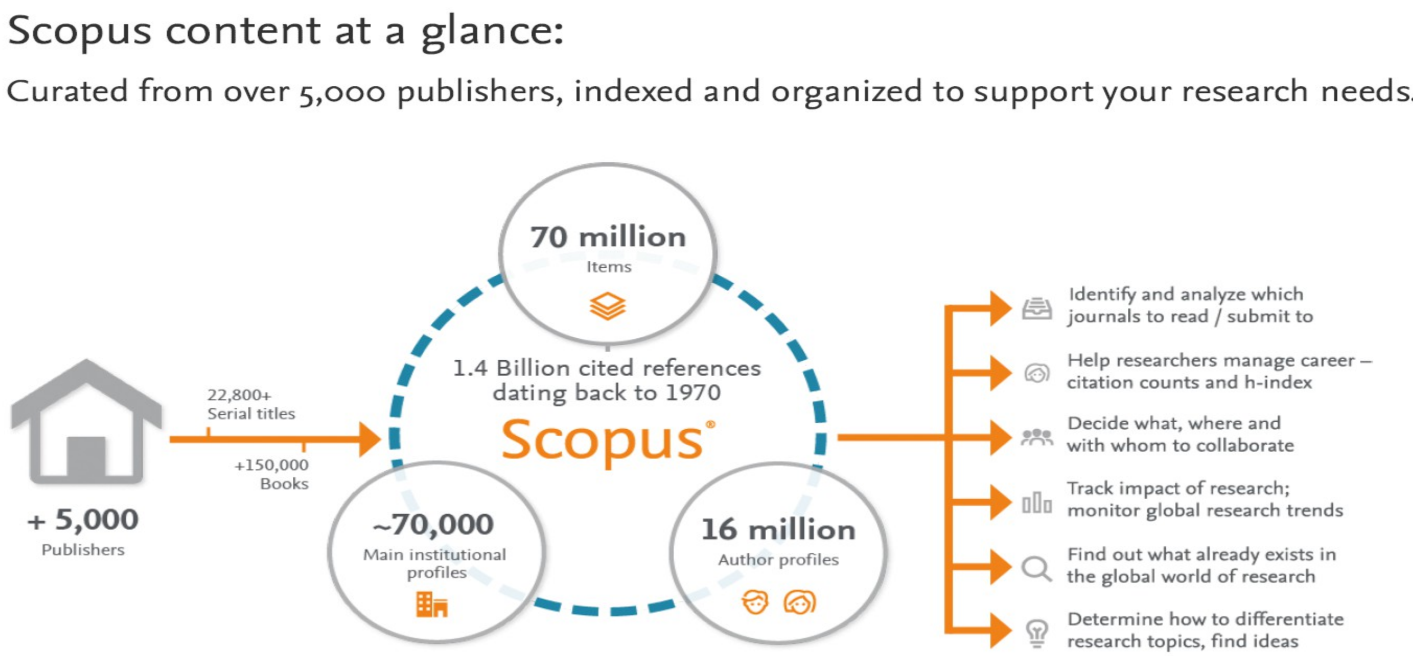
Krauss, L. DO THE NEW, BIG-MONEY SCIENCE PRIZES WORK? New Yorker February 3, 2016
https://www.newyorker.com/tech/annals-of-technology/do-the-new-big-money-science-prizes-work

FACTOR 2:Data/empirical/models tools Lord Kelvin: "In physical science the first essential step in the directionof learning any subject is to find principles of numerical reckoning and practicable methods for measuring some quality connected with it. I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind." **…**



(Kelvin also said,"There is nothing new to be discovered in physics now. All that remains is more and more precise measurement," and "All science is either physics or stamp collecting.") . This was before string theory.

Since most science appears in papers, we have better measures of **individual output** than in almost any other area of work, exclusive of sports. SCOPUS and WEB OF SCIENCE contain papers that allow us to document papers, citations, references, addresses, names, key words … BIG DATA. USPTO, SIPO and WIPO have data on patents in US, China, World.



MAIN RESULT: POWER LAW/ZIPF PLOT/PARETO DISTRIBUTION: A few researchers/papers/whatever garner a disproportionate number of papers/citations/collaborators etc. UNEQUAL DISTRIBUTION WITH BIG TAIL

"**A Zipf plot** of the number of citations of a paper vs its citation rate appears to be be consistent with a power-law dependence for leading rank papers, with exponent close to -1/2. This in turn suggests that the number of papers with x citations N(x) has a large x power law decay $N(x) \sim x^{-a}$ with a~3" People have estimated power law of papers for many fields. Get different coefficients but same basic pattern – few scientists contribute most papers; few papers contribute most citations. Log normal fits; exponential or "stretched exponential" sometimes does better. Economics coefficient is 1.84 in power law of papers. But **Fitting power laws raises statistical questions since few observations at extreme.** Statistics of extreme events needed to estimate the shape of the distribution.

Power laws; giving prizes reflect view of science as depending on **great individuals** but if teams are important, we have issue of how to apportion credit to a team. How much should being 3rd, 5th, whatever author, help my career? Depends on significance of spot on list. Many fields have first author most responsible; last author most senior. But econ has alphabetic. Movie productions list lots of people who came together to collaborate and they get "credits"

Lotka's law for paper productivity of scientists[1] – # of scientists who publish= a (# papers)$^{-2}$. This is a power law y = x raised to a negative power. Taking logs we have ln # authors = ln a -2 ln # papers. Let a=100 then we have:

1Lotka, Alfred 1., "The Frequency Distribution of Scientific Productivity, " Journal of the Washington Academy of Science , 16 (No. 12): 317—323 (1926).

papers written	number of scientists
1 paper ...	100
Two papers	25
Three ...	11
four	6
five	4
ten	1

Rank of scientist	# in group	Papers / scientist	Total for group
1	1	10	10
2-5	4	5	20
6-11	6	4	24
12-22	11	3	33
23-47	25	2	50
48-147	100	1	100
Total	147	1.6 mean; 1.0 med	237

Proportion of papers by top 5 30/237 13% – bcs so many produce 1-2 papers, they contribute "most" of science.

Power laws; giving prizes reflect view of science as depending on **great individuals** but if majority of science is done by normal folk, that is "shoulders" science stands on. And if teams are important, we have issue of how to apportion credit to a team. How much should being 3rd, 5th, whatever author, help my career? Depends on significance of spot on list. Many fields have first author most responsible; last author most senior. But econ has alphabetic. Movie productions list lots of people who came together to collaborate and they get "credits"

Danger of PI exploiting post-docs and grad students.
:what determines the decision to co-author? Will you contribute enough? Will you gain your just credit?
: what determines optimal team size? Are scientific teams too big or too little?
: does complementarity of skills made older scientists as productive as last names on papers may indicate?

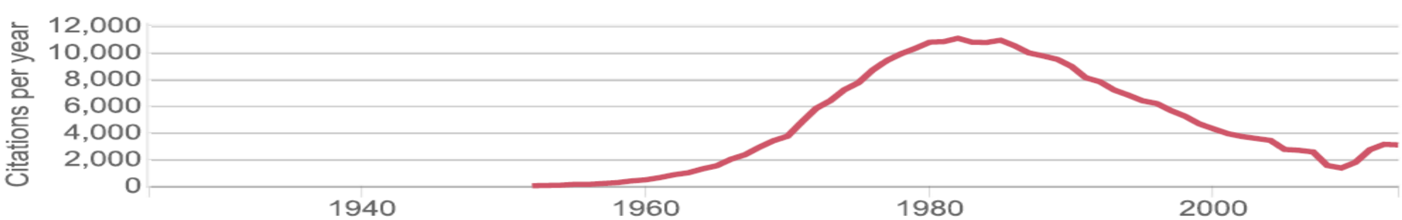
Science editorial: "a team is a team, and the members should share the credit or the blame":"A faculty member is only as good as his or her best postdoc"

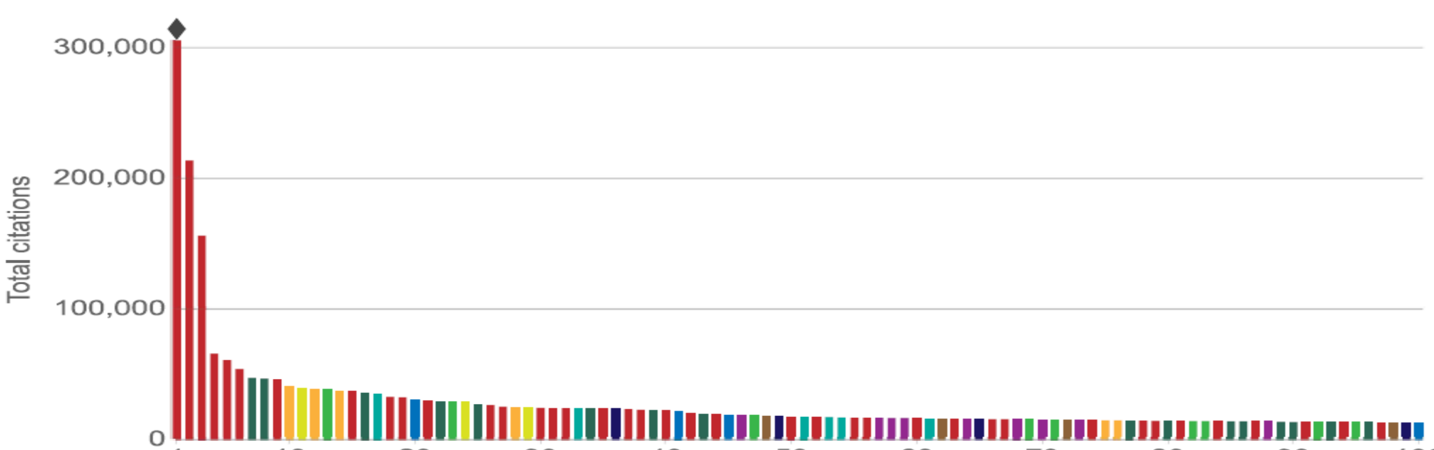
But PI suggests to A and B experiments with equal chance of success. A and B are equally adept. A's experiment succeeds; B's fails. Should B be listed on team for A? Candace Pert and Sol Snyder "lost" a Nobel for discovering opiate receptor for endorphins in brain because they got into a fight over her discovery in his lab after he shifted project to another post-doc.

Industry gives credit to firm: you don't know the name of the person who did some important applied research or development job inside Pfizer or the full team that worked on patent. But for years R&D magazine competition companies listed full teams in contest for being one of top 100 technology projects of year.

SIMILAR GET POWER LAWS IN CITATIONS, in papers, etc. The top 100 papers **_Nature_ explores the most-cited research of all time** , 29 October 2014

Rank: **1** Citations: **305,148**
Protein measurement with the folin phenol reagent.
Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J.
J. Biol. Chem. **193**, 265–275 (1951).





But many papers are "never" cited: 3 YEAR FORWARD CITATIONS IN SCOPUS, China Address, English

3 year citation_CO_ENG &CHI_2000	Freq.	Percent
0	154	84.15%
1	20	10.93%
2~5	9	4.92%
6~10	0	0.00%
11~20	0	0.00%
>20	0	0.00%
Total	183	100.00%

3 year citation_CO_ENG &CHI_2014	Freq.	Percent
0	63	28.64%
1	46	20.91%
2~5	79	35.91%
6~10	20	9.09%
11~20	10	4.55%
>20	2	0.91%
Total	220	100.00%

Distributions with long tails: economics – Pareto distribution

Team production; The number of co-authored papers has risen

1981-1999 from 2.8 to 4.2 in "all science"; 1975 to 1995, 3.3 to 6.8 in life sciences; 1980 to 2009 in labor NBER – initially most papers are single-authored; now almost no one writes alone. Wuchty, Jones and Uzzi (2007 Science) https://www.kellogg.northwestern.edu/faculty/jones-ben/htm/Teams.ScienceExpress.pdf report that co-authorship size has grown in all but one of the 171 S&E fields studied during the past 45 years.

Here is a tabulation of papers that shows ..

Tabulation of 22 068 239 papers in WOS, 1985-2008

# AUTHORS	# Papers	Proportion
1	4 452 443	0.20
2	4 443 492	0.20
3	4 026 248	0.18
4	3 109159	0.14
5	2 163 670	0.10
6	1 469 647	0.07
7	886 525	0.04
8	554 946	0.03
9	333 554	0.02
10	220 963	0.01
11+	407592	0.02

Paper topic: Has increase in # authors leveled off? Is difference among fields due to machines/ tools used?

Bibliometric data from WOS/Scopus on published papers-->**networks** of author and citations, combined with other data such as PROQUEST ondegree/thesis advisor, maybe living in same country, etc create huge data for analysis. Using natural language processing-->understanding links of papers.

Most university laboratories are small businesses run by self-employed scientists, who raise money. Example: Cori Bargman, Torsten N. Wiesel Professor Lulu and Anthony Wang Laboratory of Neural Circuits and Behavior Rockefeller: 19 people in addition to Bargman: lab manager, assistant professor, 4 RAs, technicians, 14 graduate fellows or post-docs. **Neuroscientist Cori Bargmann Named 2017 Scientist of the Year** works on biology of smell.

But there are also large centers: LAWRENCE BERKELY NATIONAL LABORATORY

At a Glance

A U.S. Department of Energy National Laboratory
Managed by: University of California
Headquarters: Berkeley, California
Director: Dr. Michael Witherell

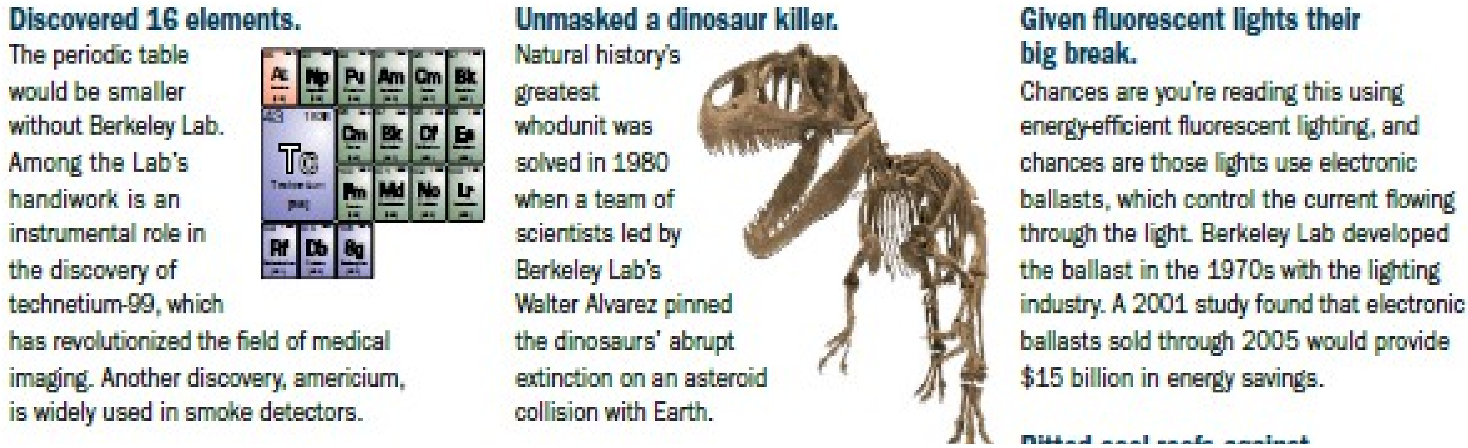
Budget / FY2017

$858 million total
$364 million total procurement
$133 million procurement to California businesses

By the Numbers

3,816 employees
13 Nobel Prizes
15 National Medal of Science recipients
16 new elements discovered
48 startups launched using Berkeley Lab technology
207 disclosed inventions and software in FY 2017
678 U.S. patents issued in last 10 years
913 IP licenses for software and inventions in last 10 years

18-LD-5345 0918


Discovered 16 elements.
The periodic table would be smaller without Berkeley Lab. Among the Lab's handiwork is an instrumental role in the discovery of technetium-99, which has revolutionized the field of medical imaging. Another discovery, americium, is widely used in smoke detectors.

Unmasked a dinosaur killer.
Natural history's greatest whodunit was solved in 1980 when a team of scientists led by Berkeley Lab's Walter Alvarez pinned the dinosaurs' abrupt extinction on an asteroid collision with Earth.

Given fluorescent lights their big break.
Chances are you're reading this using energy-efficient fluorescent lighting, and chances are those lights use electronic ballasts, which control the current flowing through the light. Berkeley Lab developed the ballast in the 1970s with the lighting industry. A 2001 study found that electronic ballasts sold through 2005 would provide $15 billion in energy savings.

FACTOR 3:NIH/NSF/firm policy. Area where ideology/political power has to "bow" to what really works. Huge sums of money where bipartisan support overwhelms even the "cut science spending" Trump. And where agencies see need for social science guidance in economic decisions since economic incentives drive much of science.
. Micro –Effect of economics on science --Paula Stephan "“How Economics Shapes Science" (Harvard University Press) gives best summary of what we know about economic considerations that drive science and first part of course

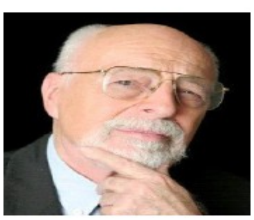
Does economics of science know enough to help government, firms, and researchers make better decisions about the allocation of resources to science? NSF developed a Science of Science Policy program in the belief that we do. NIH also funds work in this area while some of its institutes provide program officers with data on publications, citations, patents, etc in making research funding decisions.

While there is more rational decision-making in science than in many other areas, lots of "out of the blue" decisions that provide experiments, but not ideal allocation of resources: the preferences of Arlen Spector produced two "experiments" with how bio-medical sciences adjust to massive changes in funding—the 1998-2003 NIH doubling and ARRA burst of $10 Billion for NIH, of which 6.5B was added in conference reconciling House and Senate versions. Here are some areas where economics should help:

swings in government R&D funding; internationalization of research; attracting and retaining S&E workers, particularly from underrepresented minorities and women with family responsibilities

speed with which basic life science research produces drugs and better medical treatment; way physical science-based innovations affect jobs and earnings.

Policies regarding patents which balance monopoly incentives against desire to spread of ideas freely/products at low cost. Balancing risks in portfolio of science projects and funding novel "transformative" research



Hmm. You characters help us without a Large Hadron Collider … or a $30B NIH budget?

Oh Yeah!
Social science is cool and cheap.

Contrary to Merton and Arrow downplaying role of money in science, it is an economic activity. scientists are economic actors; firms do research to make profits; and even universities and research centers follow the money, up to a point, vide MIT taking Epstein money and covering it up as much as they can. Endogenizing" inputs on the basis of monetary incentives CAN lead to better policies. . Firms do research to make profits. Funding agencies provide money to advance goals. Governments are main funder of basic R&D, conducted largely in universities or non-profit research centers. Private sector funds applied research and development and establishes property rights through patents or non-disclosure. Firms have been reducing their investment in research, scientists at firms write fewer papers than in past but more patents.

FIELD: Articles on how science operates published in wide number of journals and in specialized bibliometric journals that publish on "science of science" : Journal of Informatics/Quantitative Science Studies see below, Research Policy, and Scientometrics. Since co-authors and citations are major sources of node-link-graph theory stuff many articles in Science, Nature, PNAS, elsewhere. Increased number of papers in NBER WP/

Topics

1 – China's rise – magnitude – papers, citations, ; CNKI literature; "Is S&E the "one ring that rules them all?"

TABLE 5A-1

S&E articles in all fields, for 15 largest producing regions, countries, or economies: 2008 and 2018

(Number)

Rank	Region, country, or economy	2008	2018	Average annual growth rate 2008–18 (%)	2018 world total (%)	2018 cumulative total (%)
-	World	1,755,850	2,555,959	3.83	-	-
1	China	249,049	528,263	7.81	20.67	20.67
2	United States	393,979	422,808	0.71	16.54	37.21
3	India	48,998	135,788	10.73	5.31	42.52
4	Germany	91,904	104,396	1.28	4.08	46.61
5	Japan	108,241	98,793	-0.91	3.87	50.47
6	United Kingdom	91,358	97,681	0.67	3.82	54.29
7	Russia	31,798	81,579	9.88	3.19	57.49
8	Italy	56,157	71,240	2.41	2.79	60.27
9	South Korea	44,094	66,376	4.17	2.60	62.87
10	France	66,460	66,352	-0.02	2.60	65.47
11	Brazil	35,490	60,148	5.42	2.35	67.82
12	Canada	53,296	59,968	1.19	2.35	70.17
13	Spain	44,191	54,537	2.13	2.13	72.30
14	Australia	37,174	53,610	3.73	2.10	74.40
15	Iran	17,034	48,306	10.99	1.89	76.29
-	EU	528,938	622,125	1.64	24.34	-

International coauthorship of S&E articles, for the 15 largest producing countries of S&E articles by country: 2018

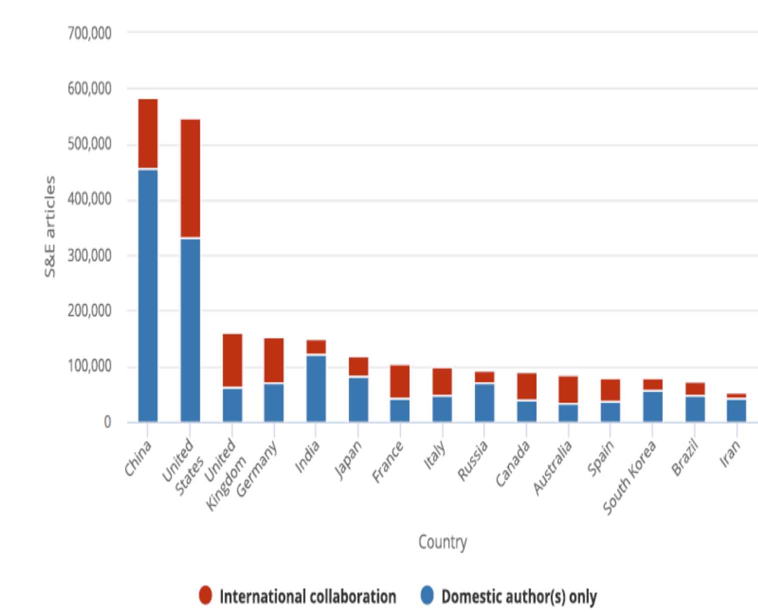
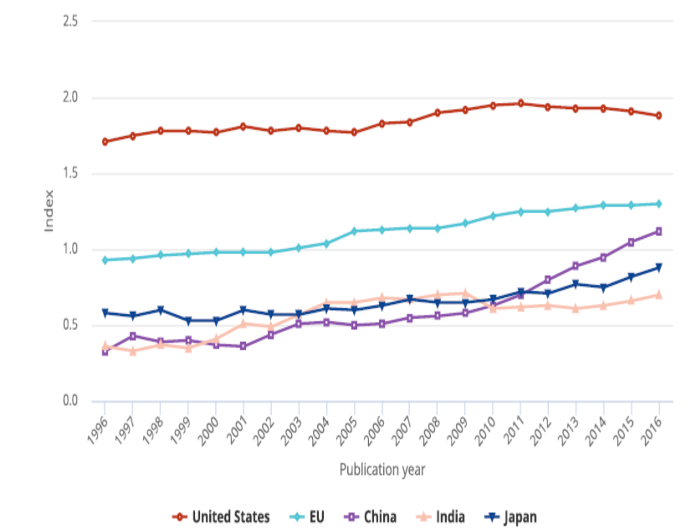
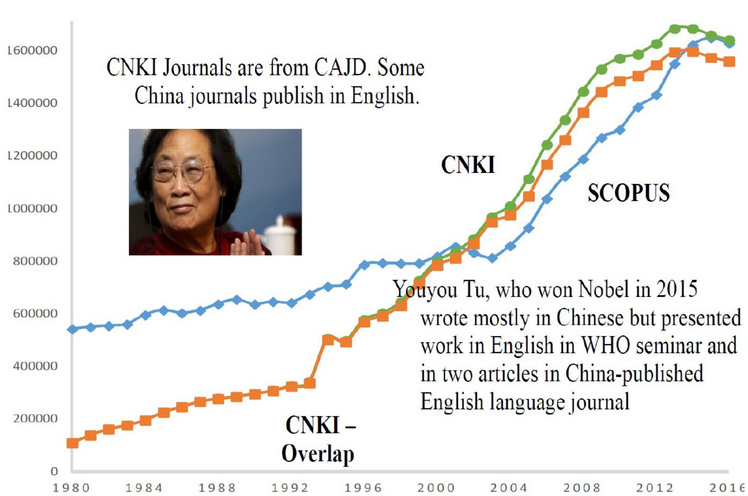
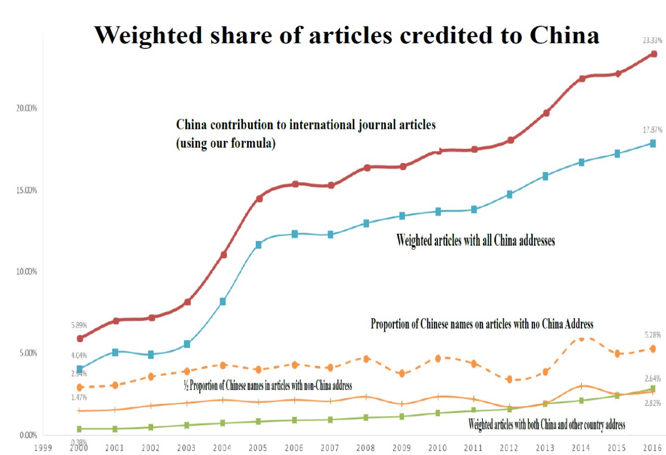


FIGURE 5A-9

S&E publication output in the top 1% of cited publications, by selected country or economy: 1996–2016







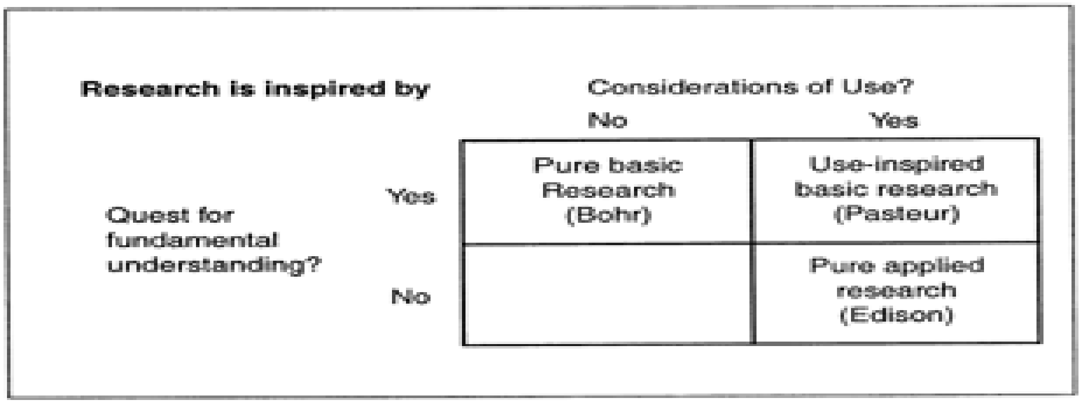
2- Innovation

If we do Sci-Eng → Increased SOUK then ec value is in taking knowledge and using it in market setting. Distinction between **Basic Knowledge vs Applied Knowledge and Innovation**

1)**Public knowledge does not mean all people can use the knowledge.** You have to know something or have access to capital and technology to use it. This creates a policy issue about publicly funded research. On the one side, incentive is to let others do it and then use public access to exploit it for economic purposes, but you need knowledge to use knowledge and much knowledge is tacit. How much spillover of knowledge is there? How sticky is knowledge creation? Enough for taxpayer to support basic science?

2) To turn basic knowledge into market goods and services, need some property right. Patents would seem ideal – every time you use my idea you pay for it, but this means charging for good with low/zero marginal cost. Patents last for a given period of time, and require administrative work to determine whether or not they should be granted. Who was the most famous patent officer in history? Question of what should be covered by patents? Discoveries about nature? New chemicals or medicines? Ways of producing?

Division between Pure and Applied : Pasteur's Quadrant



Most research funding is in the consideration of use column with industry in the Edison area and most NSF/NIH government funding in the Pasteur area. Pure basic still gets some support as always possible some pure basic may "break the bank" in an area, either directly or spurring new thoughts in the boxes that pay off

\

The standard way to link S&T to the economy is to relate measures of outputs to S&T inputs in a production function framework. Analysts regress market value of goods and services (GDP, sales or value added) on labor and capital input and R&D/S&T inputs, often in outputs/labor form.

Since S&T changes quality of goods and services, correct measurement of output depends on price deflators: if you build a better computer and charge the same price as previous computer and people buy the same number, PQ will be the same but the machine is more productive. Hedonic prices estimate the value consumers place on the attributes of items and derives the social value of the innovation in terms of a fixed bundle of attributes. When intermediate inputs improve, important to deflate purchased materials/energy/etc in gross output model; and to include measure of materials in VA regression but VA lacks price deflator so people use gross output deflator.

But this is not a good patch for the **absence of data on the innovations** that S&T produces and that underlie growth of productivity or declines in prices of actual innovations, quality adjusted. Standard production function creates a "black box" of what the actual innovations were that led to the improved productivity or higher value.

OECD "Oslo"Innovation Manual Guidelines for Collecting and Interpreting Innovation Data 163 pages in 2005 edition – defines an innovation as ".. **the implementation of a new or significantly improved** : (p 46)

product innovations – new or significantly improved goods or services
process innovations – new or significantly improved methods for production or delivery (operational processes)
organisational innovations – new or significantly improved methods in a firm's business practices, workplace organisation or external relations (organisational or managerial processes)
marketing innovations – new or significantly improved marketing methods.

US Industrial Research Institute surveyed R&D at member firms from 1991 to 1999. Data ollected at firm and line-of-business level, including patents, new sales ratio (revenues realized this year from new products introduced in the last 5 years divided by total revenues realized this year), and **cost savings realized** (cost savings from process improvements made in the last 5 years/gross profits this year). Results were reported annually in Research-Technology Management between 1993 and 1999. The data file is maintained and available through the Center for Innovation Management Studies at North Carolina State University.

Then NSF stopped funding. Low overall response rates--45% -53% Far below other Census Bureau surveys of firms. Why? Voluntary survey –response rates tend to be lower than for mandatory surveys. Not linked to other surveys so can't follow through R&D / innovation / diffusion cycle. Lower response rates for larger firms. NBER researchers recommended no further surveys. Instead put resources into 1999 Computer Network Use supplement to the Annual Survey of Manufactures; 55,000 plants: 83% response rate; mandatory survey. Found strong links between productivity and Computer network use. Questions on intensity of network use, and how used. Supply chain activities important, production not. But other countries followed innovation survey route.

More recently NSF decided they were wrong and added innov questions to BRDIS and American Business Survey

Product (good or service) innovation

A product innovation is the market introduction of a **new** or **significantly** improved good or service with respect to its capabilities, user friendliness, components, or sub-systems.

- Product innovations (new or improved) must be new to your company, but they do not need to be new to your market.
- Product innovations could have been originally developed by your company or by other companies.

1-11 During the three years 2014 to 2016, did your company introduce:

a. New or significantly improved goods (Exclude the simple resale of new goods purchased from other companies and changes of a solely aesthetic nature)? . . . . . ☐ Yes ☐ No

b. New or significantly improved services? . . . . . ☐ Yes ☐ No

1-12 If you answered "yes" to either 1-11, line a, or 1-11, line b, were any of your product innovations during the three years 2014 to 2016:

a. New to your market?. . . . . ☐ Yes ☐ No

Your company introduced a new or significantly improved good or service to your market before your competitors. (It may have been available in other markets.)

b. New only to your company?. . . . . ☐ Yes ☐ No

Your company introduced a new or significantly improved good or service that was already available from your competitors in your market.

1-13 Using the definitions above, please give the percentage of your total sales in 2016 from:

a. New or significantly improved goods and services introduced during 2014 to 2016 that were **new to your market** . . . . . ___ %

b. New or significantly improved goods and services introduced during 2014 to 2016 that were **new only to your company**. . . . . ___ %

c. Goods and services that were **unchanged or only marginally modified** during 2014 to 2016 (include the resale of new goods or services purchased from other companies) . . . . . ___ %

d. **Total sales in 2016** . . . . . 1 0 0 %

Process innovation

A process innovation is the implementation of a **new** or **significantly** improved production process, distribution method, or support activity for your goods or services.

- Process innovations must be new to your company, but they do not need to be new to your market.
- The innovation could have been originally developed by your company or by other companies.
- Exclude purely organizational innovations.

1-14 During the three years 2014 to 2016, did your company introduce:

a. New or significantly improved methods of manufacturing or producing goods or services? . . . . . ☐ Yes ☐ No

b. New or significantly improved logistics, delivery or distribution methods for your inputs, goods, or services? . . . . . ☐ Yes ☐ No

c. New or significantly improved supporting activities for your processes, such as maintenance systems or operations for purchasing, accounting, or computing? . . . . . ☐ Yes ☐ No

And

PRODUCT OR PROCESS INNOVATION ACTIVITIES

Innovation activities include the acquisition of machinery, equipment, buildings, software, and licenses; engineering and development work, feasibility studies, design, training, R&D and marketing when they are specifically undertaken to develop and/or implement a product or process innovation. This includes also all types of research and development activities to create new knowledge or solve scientific or technical problems.

During the three years 2015 to 2017, did this business engage in the following product or process innovation activities? ***Select one for each row.***

	Product or process innovation activities only: Yes	No
a. **In-house R&D** Research and development activities undertaken by this business to create new knowledge, solve scientific or technical problems, or devise new applications of available knowledge (include software development that meets this requirement) If yes, did this business perform R&D during the three years 2015 to 2017: ☐ Continuously (business had permanent R&D staff in-house) ☐ Occasionally (as needed only)	☐	☐
b. **External R&D** This business contracted-out R&D to other companies or to public or private research organizations	☐	☐
c. **Acquisition of machinery, equipment, software & buildings** New machinery, equipment software and buildings that were acquired for the purpose of developing new or significantly improved goods, services, manufacturing or logistics	☐	☐
d. **Acquisition of existing knowledge from other companies or organizations** Acquisition of existing know-how, copyrighted works, patented and non-patented inventions, etc. from other companies or organizations for the development of new or significantly improved products and processes	☐	☐
e. **Training for innovative activities** In-house or contracted out training for your personnel specifically for the development and/or introduction of new or significantly improved products and processes	☐	☐
f. **Market introduction of innovations** In-house or contracted out activities for the market introduction of your new or significantly improved goods or services, including market research, launch advertising, and social media announcements	☐	☐
g. **Brand Building** In-house or contracted out activities such as advertising or promotion to build this business's brand identity or brand name	☐	☐
h. **Design** In-house or contracted out activities to alter the shape, appearance or usability of goods or services	☐	☐
i. **Other** Other in-house or contracted out activities to develop or implement new or significantly improved products or processes such as feasibility studies, testing, industrial engineering, etc.	☐	☐

RESULTS OF INNOVATION ACTIVITIES

During the three years 2015 to 2017, did this business have any innovation activities that did not result in a product or process innovation because the activities were: ***Select one for each row.***

	Yes	No
a. Abandoned or suspended before completion	☐	☐
b. Still ongoing at the end of 2017	☐	☐

REGULATIONS AND INNOVATION

What is the effect of the following types of legislation or regulations on this business's innovation activities during the three years 2015 to 2017? ***Select all that apply.***

Legislation or regulation	Stimulated Innovation	Created no major problems	Created uncertainty	Generated an excessive burden	Not applicable
Product safety /consumer protection	☐	☐	☐	☐	☐
Operational and worker safety	☐	☐	☐	☐	☐
Environmental	☐	☐	☐	☐	☐
Intellectual property	☐	☐	☐	☐	☐
Tax	☐	☐	☐	☐	☐
Employment or social affairs	☐	☐	☐	☐	☐
Other, specify ↙	☐	☐	☐	☐	☐

Skip to SECTION D: RESEARCH AND DEVELOPMENT

ORGANIZATIONAL AND MARKETING INNOVATION

During the three years 2015 to 2017, did this business introduce new: ***Select one for each row.***

		Yes	No
a.	**Business practices** for organizing procedures *(for example, first time use of supply chain management, business re-engineering, knowledge management, lean production, quality management, etc.)*	☐	☐
b.	Methods of **organizing work responsibilities and decision making** *(for example, first time use of a new system of employee responsibilities, team work, decentralization, integration or de-integration of departments, education/training systems, etc.)*	☐	☐
c.	Methods of **organizing external relations** with other companies or public organizations *(for example, first time use of alliances, partnerships, outsourcing or sub-contracting, etc.)*	☐	☐
d.	Aesthetic **design or packaging** of a good or service *(exclude changes that alter the product's functional or user characteristics – these are product innovations)*	☐	☐
e.	Media or techniques for **product promotion** (*for example, first time use of a new advertising media, a new brand image, introduction of loyalty cards, etc.)*	☐	☐
f.	Methods for **product placement** or sales channels *(for example, first time use of franchising or distribution licenses, direct selling, exclusive retailing, new concepts for product presentation, etc.)*	☐	☐
g.	Methods of **pricing** goods or services *(for example, first time use of variable pricing by demand, discount systems, etc.)*	☐	☐

If "No" is selected for a. *and* b. from the PRODUCT INNOVATION question **AND** "No" is selected for a. – c. from the PROCESS INNOVATION question **AND** "No" is selected for a. – g. from the ORGANIZATIONAL AND MARKETING INNOVATION question, skip to BUSINESS REASONS FOR NOT INNOVATING.

FACTORS INTERFERING WITH BUSINESS INNOVATION

During the three years 2015 to 2017, how important were the following factors in interfering with this business's ability to innovate? ***Select one for each row.***

		Very Important	Somewhat Important	Not at all Important
a.	Lack of internal finance	☐	☐	☐
b.	Lack of credit or private equity	☐	☐	☐
c.	Innovation costs too high	☐	☐	☐
d.	Lack of skilled employees within the business	☐	☐	☐
e.	Lack of collaboration partners	☐	☐	☐
f.	Difficulties in obtaining government grants or subsidies	☐	☐	☐
g.	Uncertain market demand for your ideas	☐	☐	☐
h.	Too much competition in your market	☐	☐	☐

Some Patterns:

Percentage of self-reported innovative firms varies across countries in odd ways. In the 2009 BRDIS, only 6.4 percent of U.S. firms reported introducing new or significantly improved goods in the previous three years, while only 10.3 percent reported new or significantly improved services. By contrast, the percentage of German firms that reported themselves as innovative from 2005 through 2007 was far higher, at 79.9 percent

Arundel et al., 2012 examines how people interpret questions concerning innovation. The study asked respondents to describe their most important innovation, and then had experts classify whether they met the requirements of an innovation. It indicates which types of innovations are deemed most important for business performance – often marketing and organizational innovations were cited.

R&D Firms Are More Likely to Innovate but Most Innovating Firms Do Not Do R&D

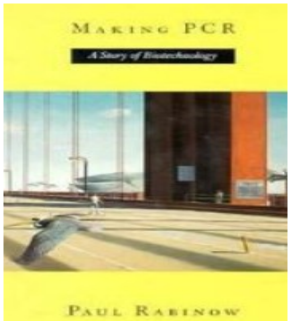
	%of all firms in scope	%firms in row with new/ significantly improved product or process	%firms with new or significantly improved product or process
Firms doing R&D	4%	65%	16%
Firms not doing R&D	96%	14%	84%

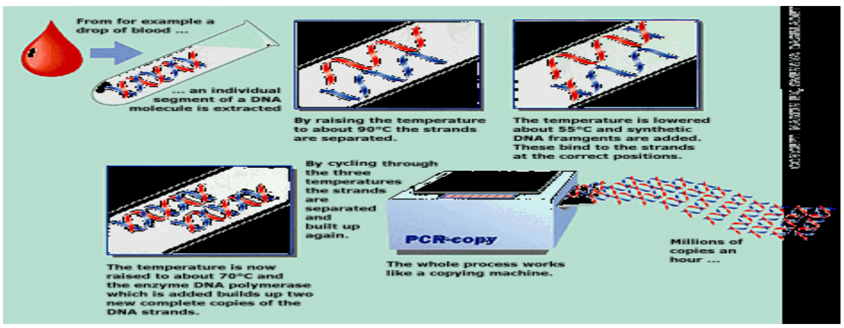
Source: Tabulations from the 2009 BRDIS.

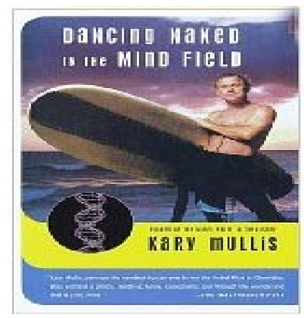
In some cases, a potential innovation may not make it to market because it needs complementary innovations to succeed. Chemcor, an ultra-hard glass, was invented in the 1960s by Corning. Because of the cost of production, Chemcor did not find a place in the market until 2006. Renamed as Gorilla Glass, it became the product of choice for cell phone screens. The concept of unmarketed innovations is implicit in an existing question from the Community Innovation Survey.

3. Me or my machine/AI?

Technology of instruments – tools affect what science actually accomplishes;: In many areas scientists must have lab or equipment to proceed. Expensive startup packages for young scientists (~1M$). Examples of important equipment: Scanning tunneling microscopy (STM) for viewing surfaces at the atomic level. Its development in 1981 won its inventors, Binnig and Rohrer (at IBM Zürich), the Nobel Prize in Physics in 1986; created huge advances in nano-technology. PCR in biology – tool that opens door for amplifying DNA. Paul Rabinow's book Making PCR questions whether Mullis invented PCR or “merely” came up with the concept and argues that scii discovery is often product of group work: ”Committees and science journalists like the idea of associating a unique idea with a unique purpose, the lone genius. PCR is, in fact, one of the classic examples of teamwork”

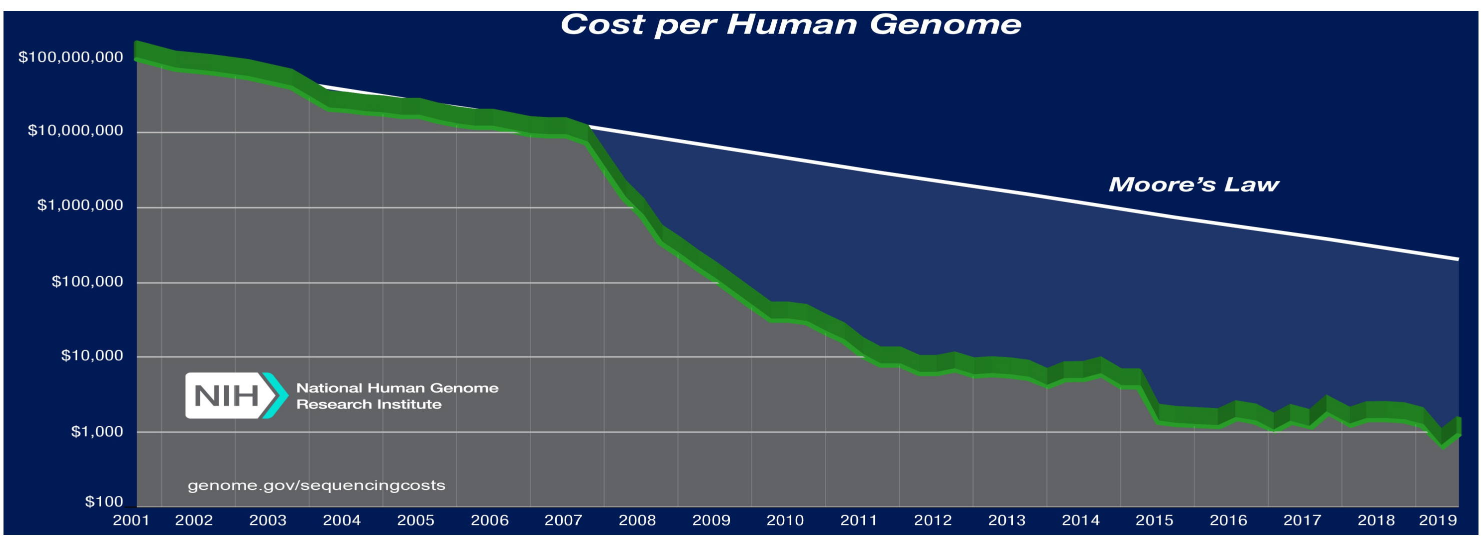






Human genome project and improved technology for reading gene sequences. As technology improved Ventor Institute/Celerus and Broad Institute retrained/fired workers since the “next generation” approach to sequencing DNA was hundreds of times more efficient than existing technique but required entirely different technologies.

Paula Stephan: "... in 1990 the best equipped lab could sequence 1000 base pairs a day, by January 2000 the 20 labs involved in mapping the human genome were collectively sequencing 1000 base pairs a second. The cost per finished base pair fell from $10.00 in 1990 to under $.05 in 2003 (Francis S. Collins, Michael Morgan, Aristides Patrinos 2003) and is roughly $.01 in 2007 (www.biodesign.asu./edu/news/232/). Measured in base pairs sequenced per person per day, productivity increased more than 20,000 fold from the early 1990s to 2007, doubling approximately every 12 months (http://www.bio-era.net/news/add_news_18.html

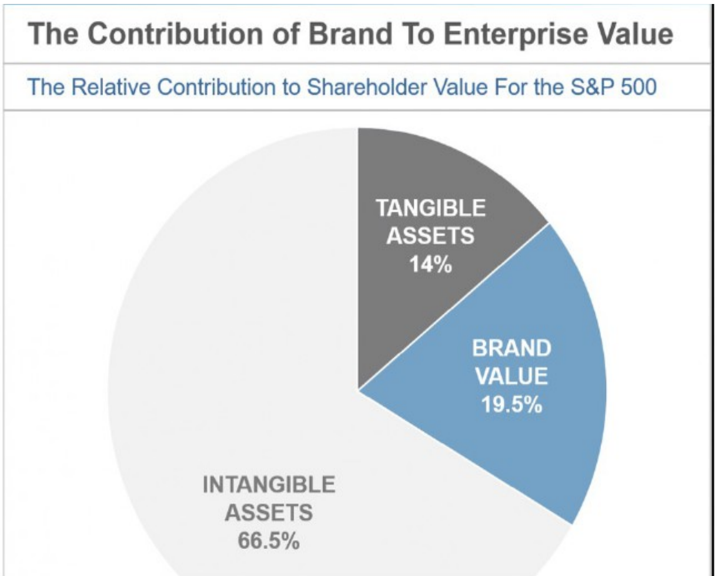


There are AI machines that form hypotheses etc.

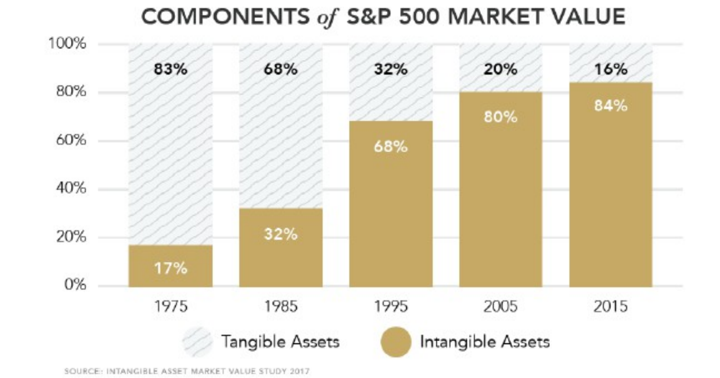
https://www.quantamagazine.org/how-artificial-intelligence-is-changing-science-20190311/

https://www.graphcore.ai/posts/why-artificial-intelligence-will-allow-us-to-make-new-scientific-discoveries

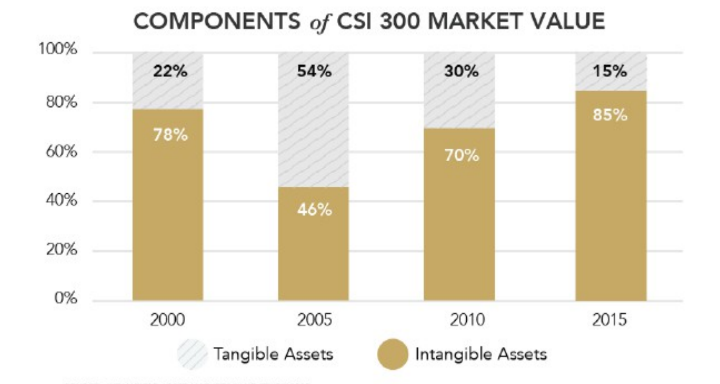
3 Intangible capital and firms: costing environment.



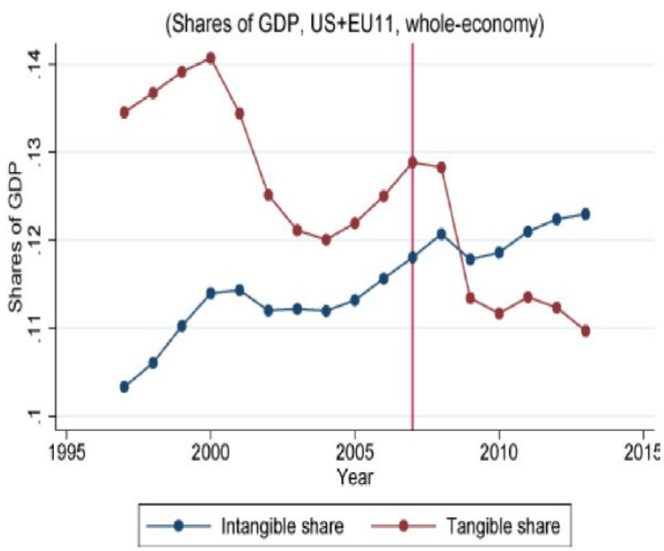
Previously reported in the 2017 study update, the Components of the S&P 500 Market Value continues to support the role of intellectual capital as the leading asset class. The term "intellectual capital" refers generally to traditional Intellectual Property assets - patents, trademarks and copyrights. The growth in the value of Intellectual Capital Equity® can be seen when evaluating the market capitalization of the S&P 500 as shown in the chart below.



While emphasis is often on technology-driven intangible assets such as patents and trade secrets, brand value is also an important component of IAMV. For the first time in 2017 we compared the US IAMV calculations to Interbrand's Best Global Brands 2016 calculation of brand value. Comparing the 39 companies appearing on both the S&P 500 and the Interbrand list suggests brand value may represent roughly one quarter or more of average IAMV.



The CSI 300 is a capitalization-weighted stock market index designed to replicate the performance of top 300 stocks traded in the Shanghai and Shenzhen stock exchanges. As the Index aims to reflect the overall performance of the China A-share market, it serves as a useful comparative barometer.



: GDP adjusted to include intangibles.

Established in 2003, Ocean Tomo, LLC, is the Intellectual Capital Merchant Banc™ firm providing Opinion, Management and Advisory services centered on intellectual property assets. Practice offerings address financial expert testimony, valuation, strategy, risk management, venture development, investments and transaction brokerage. Ocean Tomo assists clients – corporations, law firms, governments and institutional investors – in realizing Intellectual Capital Equity® value broadly defined. Chicago, Houston, Greenich Conn, SF, and CHINA.

Haskel, J, and S Westlake (2017), Capitalism Without Capital The Rise of the Intangible Economy, Princeton, NJ: Princeton University Press. examines if intangible can explain stagnation in GDP in any of four specific ways: 1)The first, and most straightforward link between intangible investment and secular stagnation is mismeasurement. Spillovers; 2) Because intangibles tend to generate more spillovers, a slowdown in intangible capital services growth would manifest itself in the data as a slowdown in TFP growth; 3) widening gap between leading and lagging firms, with rents rising for top firms 4) intangibles are somehow generating fewer spillovers than they used to.

Ellen McGratten (NBER WP 23233, March 2017) Because firms invest heavily in R&D, software, brands, and other intangible assets—at a rate close to that of tangible assets—changes in measured GDP understate the actual changes in total output.... This mis-measurement leaves business cycle modelers with large and unexplained labor wedges accounting for most of the fluctuations in aggregate data. I incorporate intangible investments into a multi-sector general equilibrium model, with intangible investments reassigned from intermediate to final uses. I find that the model's common component of TFP is not correlated at business cycle frequencies with the standard measures of aggregate TFP used in the macroeconomic literature.

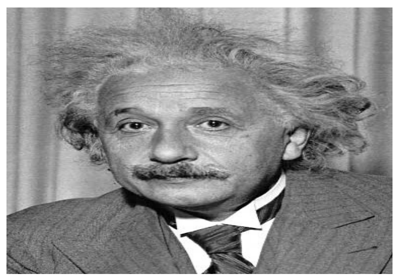
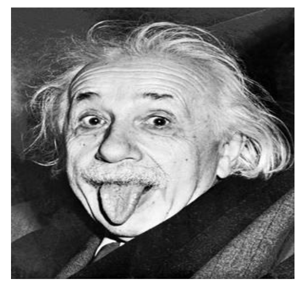
4 **Tournaments and Slippery Science: Is there Better Way to Operate Science?**

Science is a tournament where people compete at various stages … for getting into education programs; for getting fellowships; for getting into particular labs and projects; for grant funding; for publishing in top journals; for … It is the ultimate competitive market, but this can create problems? Tournament models show that with proper design and participants behaving honorably, tournaments can produce best possible solutions. But designs have problems, and any reward system can generate "gamesmanship" that produces undesireable outcomes.

Who Gets Funded? NIH funding goes largely to older PIs. They have track record and so look like safer bets than younger scientists, and review panels are almost entirely older experts, some of whom will prefer protecting their intellectual product than supporting work that could undo what they have done. If older PIs make decisions, does this lower new ideas? If they protect younger scientists from bureaucracy, does it help produce new ideas? Must consider life-cycle of ideas and comparison with some other ways to fund research.

Would it be better to give more money for graduate studies than big prizes? To award grants that would free post-docs from working in labs of senior professors? When is a prize tournament best? Borjas and Doran 2015. "Prizes and Productivity: How Winning the Fields Medal Affects Scientific Output," Journal of Human Resources, University of Wisconsin Press, vol. 50(3), pages 728-758. examined the production of mathematicians who win Fields Medal. Did winning medal increase or decrease their production?

Who is the GENIUS changing science? Young vs old – math Fields given at 40 vs Nobel given at 65, but when is the Nobel work done? Age of Nobels has gotten older: (Jones & Weinberg, PNAS, 2011 Nov 22)

Problems in tournament:

Peer review? Why America's Best Scientists Don't Get Funded headline based on **Nicholson & Ioannidis. "Research grants: Conform and be funded." Nature 492, 34-36 (06 December 2012)** claim that NIH gives most money to mediocre scientists, especially when they are in study section .. most authors of papers that gain hit status with 1000 citations did not gain NIH funding while scientists on NIH study sections with few than 1000 citation hits were funded. Fits with the vision of government and peer review as conservative in choice of what to fund. NIH responded **(Correspondence, Nature Jan 3, 2013) vs**

Li, Danielle, and Leila Agha. "Big Names or Big Ideas: Do Peer Review Panels Select the Best Science Proposals?" Science 348, no. 6233 (April 24, 2015): 434–438 track publication, citation, and patenting outcomes associated with more than 130,000 research project (R01) NIH grants from 1980 to 2008. Finds **better peer review scores consistently associated with better research outcomes** even with detailed controls for an investigator's publication history, grant history, institutional affiliations, career stage, and degree types. One-sd worse peer review associated with 15% less citations, 7% less publications, 19% less high impact papers, and 14% fewer patents.

Rush to Publication produces questionable results

Pharma Firms complain that many results in bio literature not reproducible:
Studies show only 10% of published articles are reproducible. What is happening? May 3, 2012 M. Pritsker :

- 1) biotech company Amgen had a team of about 100 scientists trying to reproduce the findings of 53 "landmark" articles in cancer research published by reputable labs in top journals. Only 6 were reproduced.
- 2) Scientists at Bayer, examined 67 target-validation projects in oncology, women's health, and cardiovascular medicine. Published results were reproduced in only 14 out of 67 projects (about 21%).
- 3) Why? Is it incentives to get results out? Is it problem of following recipe exactly? The Amgen and Bayer complaints are from people who want to develop drugs on academic results.

Slowdown in drug development and increased cost –>

NIH establishes National Center for Advancing Translational Sciences

New center to speed movement of discoveries from lab to patients

In a move to re-engineer the process of translating scientific discoveries into new drugs, diagnostics, and devices, the National Institutes of Health has established the National Center for Advancing Translational Sciences (NCATS). The action was made possible by Congress' approval of a fiscal year 2012 spending bill and the president's signing of the bill, which includes the establishment of NCATS with a budget of $575 million.

NCATS will serve as the nation's hub for catalyzing innovations in translational science. Working closely with partners in the regulatory, academic, nonprofit, and private sectors, NCATS will strive to identify and overcome hurdles that slow the development of effective treatments and cures.

Question: What other alternatives might NIH have chosen? How will we measure if this succeeds?
Paper: What is evidence on success/failure of NCATS? Is it possible to conduct a scientific assessment ?

Problems in Publishing World

Nature NEWS ·14 JANUARY 2019 Open-access row prompts editorial board of Elsevier journal to resign. *The board of the Journal of Informetrics (JOI)has launched a new open-access publication.* The board told *Nature* that given the journal's subject matter — the assessment and dissemination of science — it felt it needed to be at the forefront of open publishing practices, which it says includes making bibliographic references freely available for analysisvand reuse, and being open access and owned by the community. Board members also wanted Elsevier to lower the journal's article-publishing charges for authors and participate in the Initiative for Open Citations — a project aiming to free up citation data for study. The *JOI* typically charges researchers and institutions to access its content, but authors … can choose to make their paper freely available by paying a fee of US$1,800, plus tax (a model known as hybrid open access).The *JOI* editorial board, comprising 27 members and 2 associate editors, sent a joint resignation letter on 10 January. The response: "Elsevier needs to be able to continue investing in ways that add value to the research process, which it cannot do if it gives this value away for free."

Fresh start On 14 January, the same researchers launched a free-to-read journal called *Quantitative Science Studies* (*QSS*), which has the same editorial board and is published by MIT Press of Cambridge, Massachusetts, under the banner of the International Society for Scientometrics and Informetrics (ISSI)…. The German National Library of Science and Technology in Hanover invested more than €100,000 (US$115,000) to waive article-processing charges for authors submitting to *QSS* in the next three years, and to assist with other journal costs. The journal's publication fee is $600 per paper for ISSI members and $800 for non-members.

Space for a good paper on open access publishing, the journal fees to authors or readers or universities.

Slippery and Fraudulent Work and …Acclaimed Harvard Scientist Is Arrested,Accused Of Lying About Ties To China January 28, 2020 · 2:31 PM ET

Charles Lieber, the chair of Harvard University's Department of Chemistry and Chemical Biology, has been arrested and criminally charged with making "false, fictitious and fraudulent statements" to the U.S. Defense Department about his ties to a Chinese government program to recruit foreign scientists and researchers.
The Justice Department says Lieber, 60, lied about his contact with the Chinese program known as the Thousand Talents Plan, which the U.S. has previously flagged as a serious intelligence concern. He also is accused of lying about about a lucrative contract he signed with China's Wuhan University of Technology.
In an affidavit unsealed Tuesday, FBI Special Agent Robert Plumb said Lieber, who led a Harvard research group focusing on nanoscience, had established a research lab at the Wuhan university — apparently unbeknownst to Harvard. In response to the charges against Lieber, Harvard said in a statement to NPR: "The charges brought by the U.S. government against Professor Lieber are extremely serious. Harvard is cooperating with federal authorities, including the National Institutes of Health, and is initiating its own review of the alleged misconduct. Professor Lieber has been placed on indefinite administrative leave.". It says the deal called for Lieber to be paid up to $50,000 a month, in addition to $150,000 per year "for living and personal expenses."

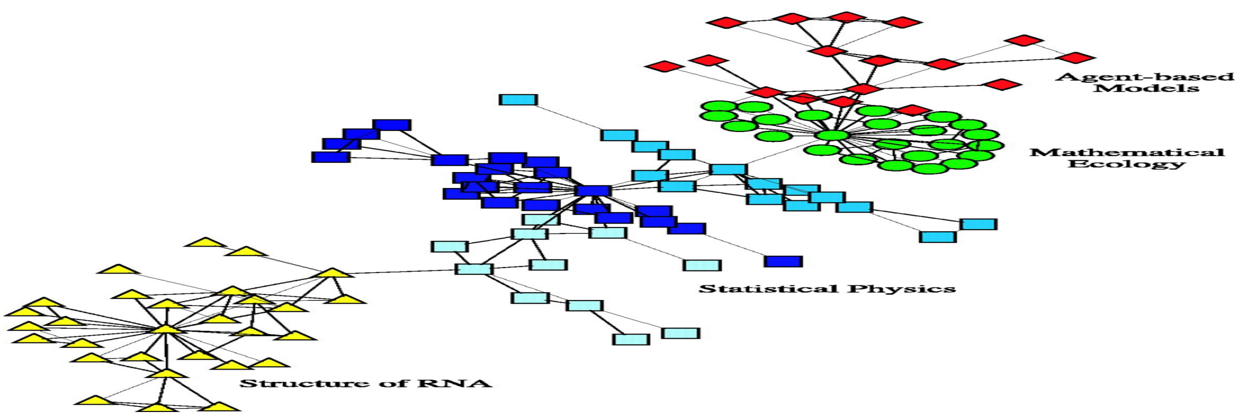
"Lieber was also awarded more than $1.5 million by WUT and the Chinese government to establish research lab and conduct research at WUT," the document states.

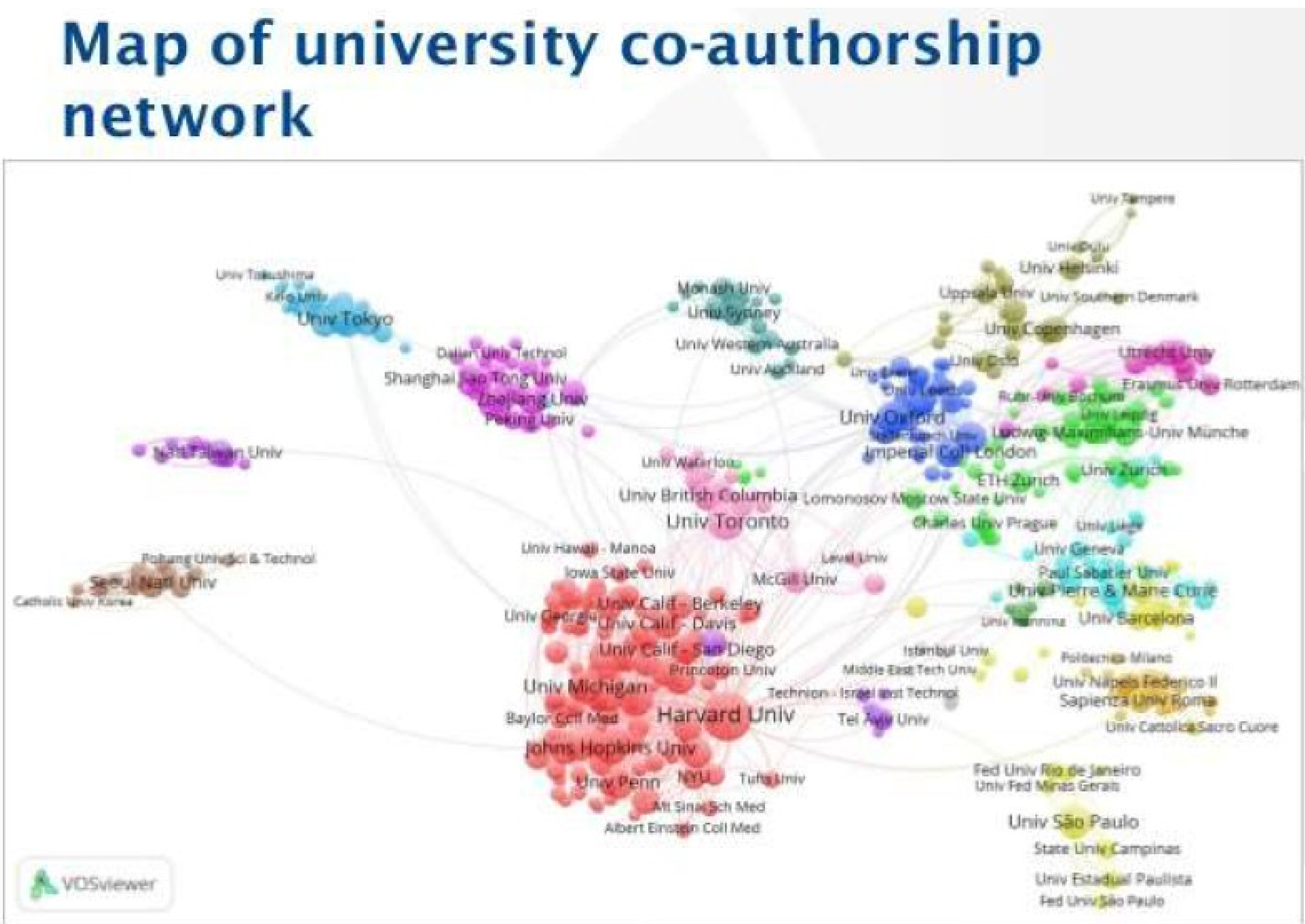
5. NETWORK ANALYSES: Co-authorships and relations of fields allow us to identify critical people and papers by how they are connected to others. This treats science as social activity. Great man as shoulder of giants–Riemann conceived the hypothesis but based on complex variables from French mathematicians, Gauss etc

But what about scientist who fails to communicate idea or communicates it poorly? Social interactions help produce ideas: Good to have outsiders and cross-disciplinary. Physicists have played major role in biology. But no empirical study nails down the value of cross-disciplinary work.

Machine learning algorithms allow us to follow the words/ideas so when I cite your paper, you can see I read it and used it rather than gave a perfunctory cite. **Tracing key words/ideas to delineate fields of study**

Connection among papers by -authorship network among universities/labs





Coauthorship Networks and Institutional Collaboration in Revista Española de Cardiología Publications
Redes de coautorías y colaboración institucional en Revista Española de Cardiología 2006 5 authors

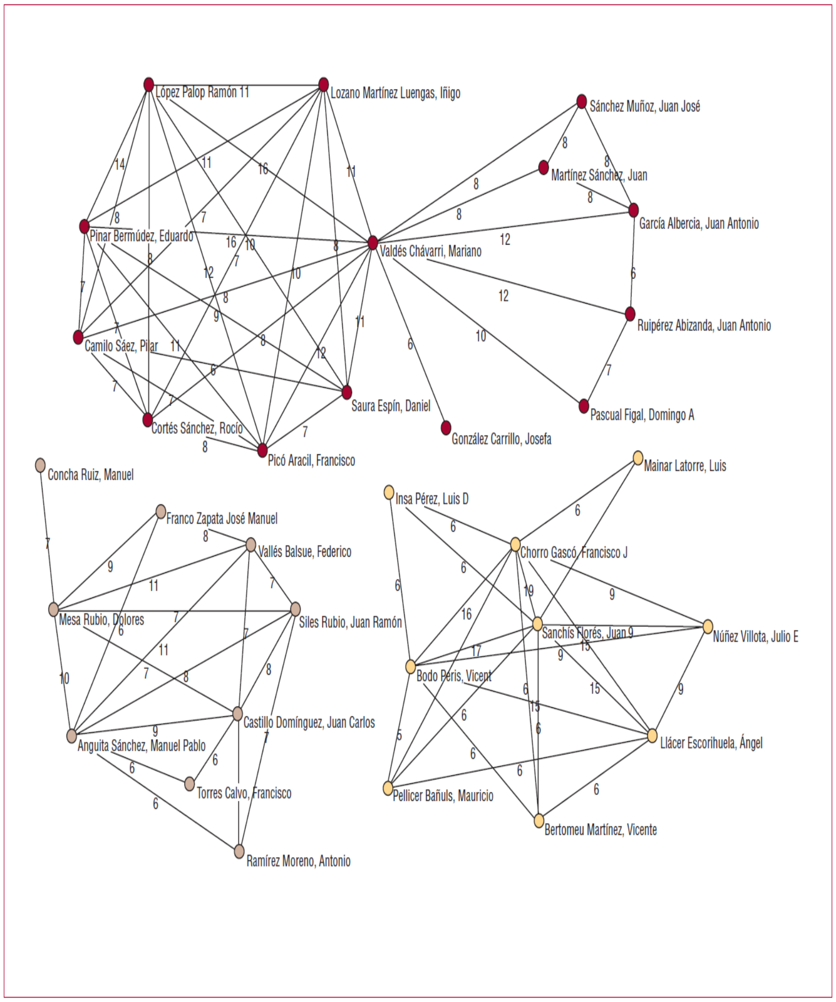


Figure 1. Clusters 1 to 3 (9 or more members) with an intensity of collaboration of 6 or more coauthors.

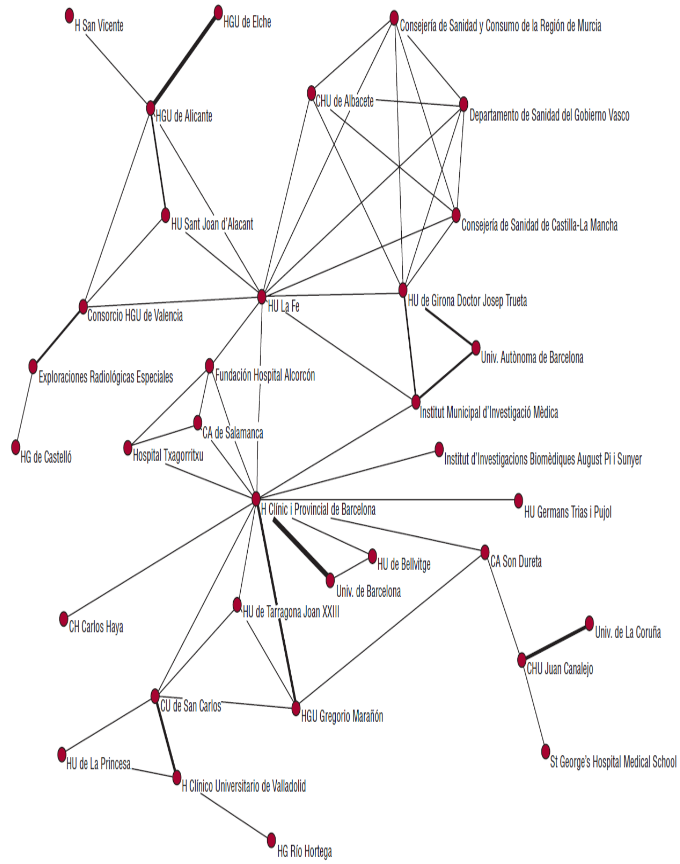


Figure 2. Principal nucleus of network of institutional collaborations with intensity of collaboration of 3 or more collaboration relationships.

TABLE 6. Number of Collaborations Between Institutions Grouped by Type of Collaboration and Total Number of Papers Published in Collaboration in *Revista Española de Cardiología* (2000-2005)

	Number of Collaborations						Total
Types of Collaboration	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2000-2005, %**
Intrainstitutional collaboration (type 1)	43	37	39	39	41	45	244 (40.87%)
Interinstitutional collaboration							
Same autonomous region (type 2a)	34	42	36	37	42	46	237 (39.7%)
Different autonomous regions (type 2b)	13	20	13	14	9	23	92 (15.41%)
International collaboration (type 3)	3	4	5	6	4	2	24 (4.02%)
Total, %*	93 (15.58%)	103 (17.25%)	93 (15.58%)	96 (16.08%)	96 (16.08%)	116 (19.43%)	597 (100%)
Number of papers in collaboration (%)†	77 (43.02%)	83 (46.63%)	73 (43.45%)	74 (49.01%)	77 (52.38%)	89 (56.69%)	473 (100%)

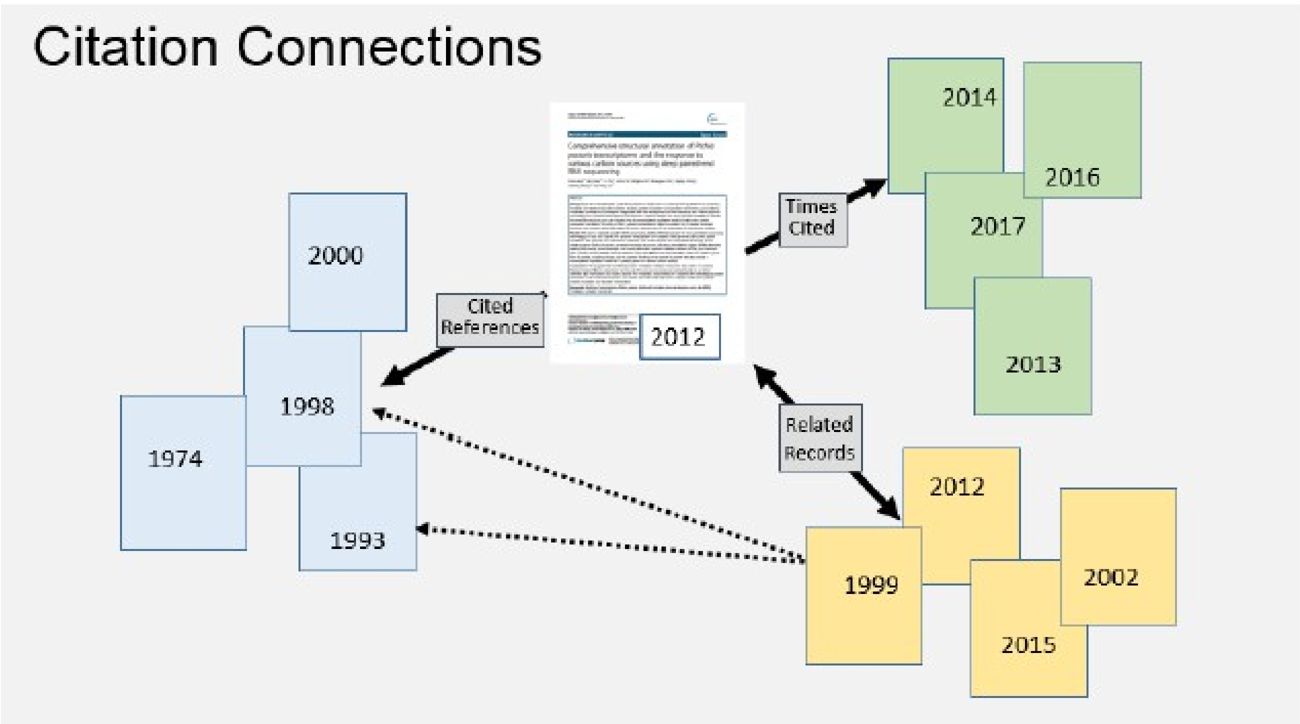
*Note that in number of collaborations, some papers include no collaboration and others involve more than one type of collaboration. To calculate percentages we did not include 31 papers authored in collaboration but not signed by Spanish institutions.
†We took into account when calculating the 473 papers signed by at least one Spanish institution. Percentages are calculated with respect to total number of papers per year, reported in Table 1.

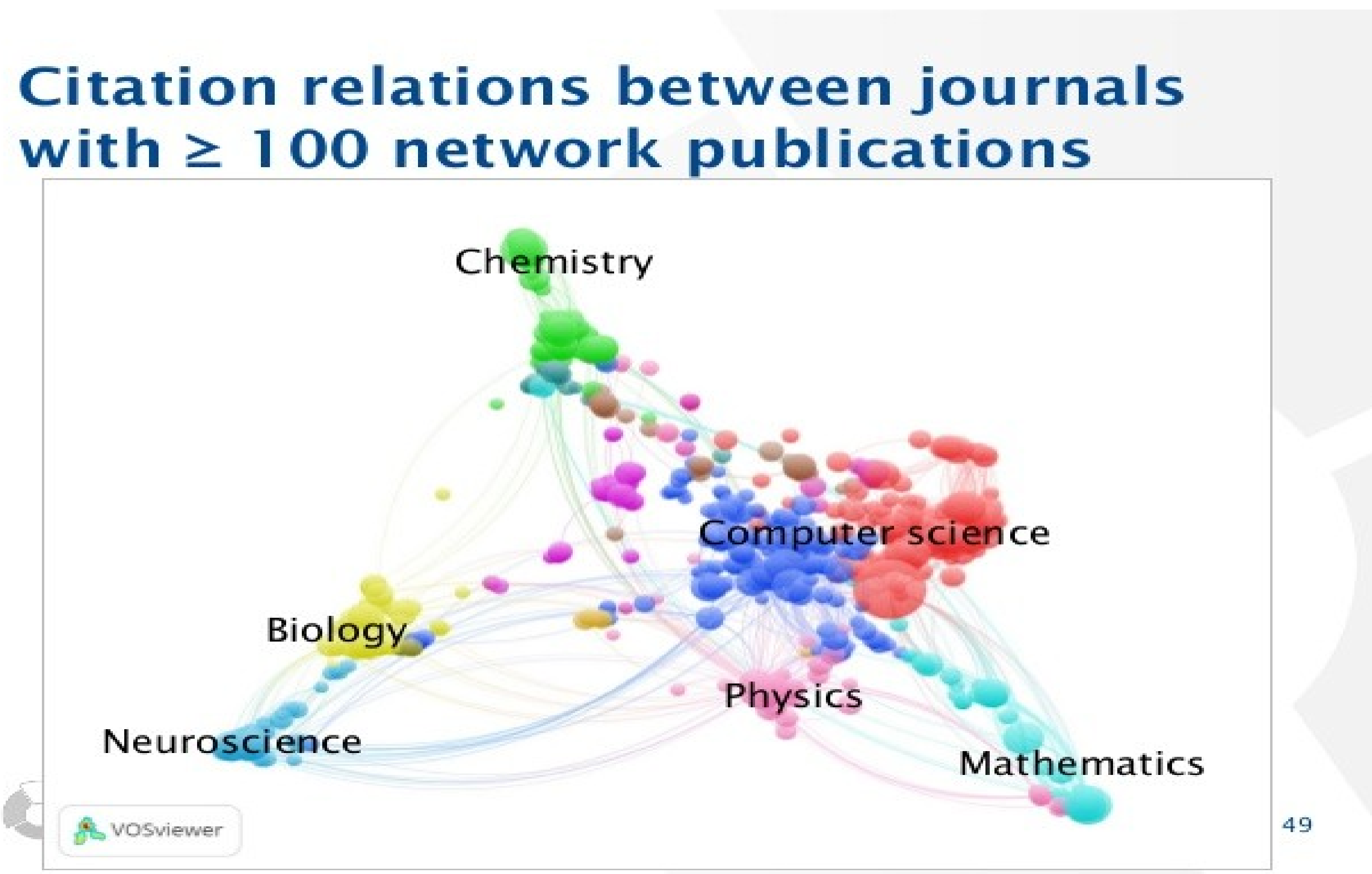
The Citation Network is represented for each paper through:

Cites References – the backward references - the research that this paper cites,

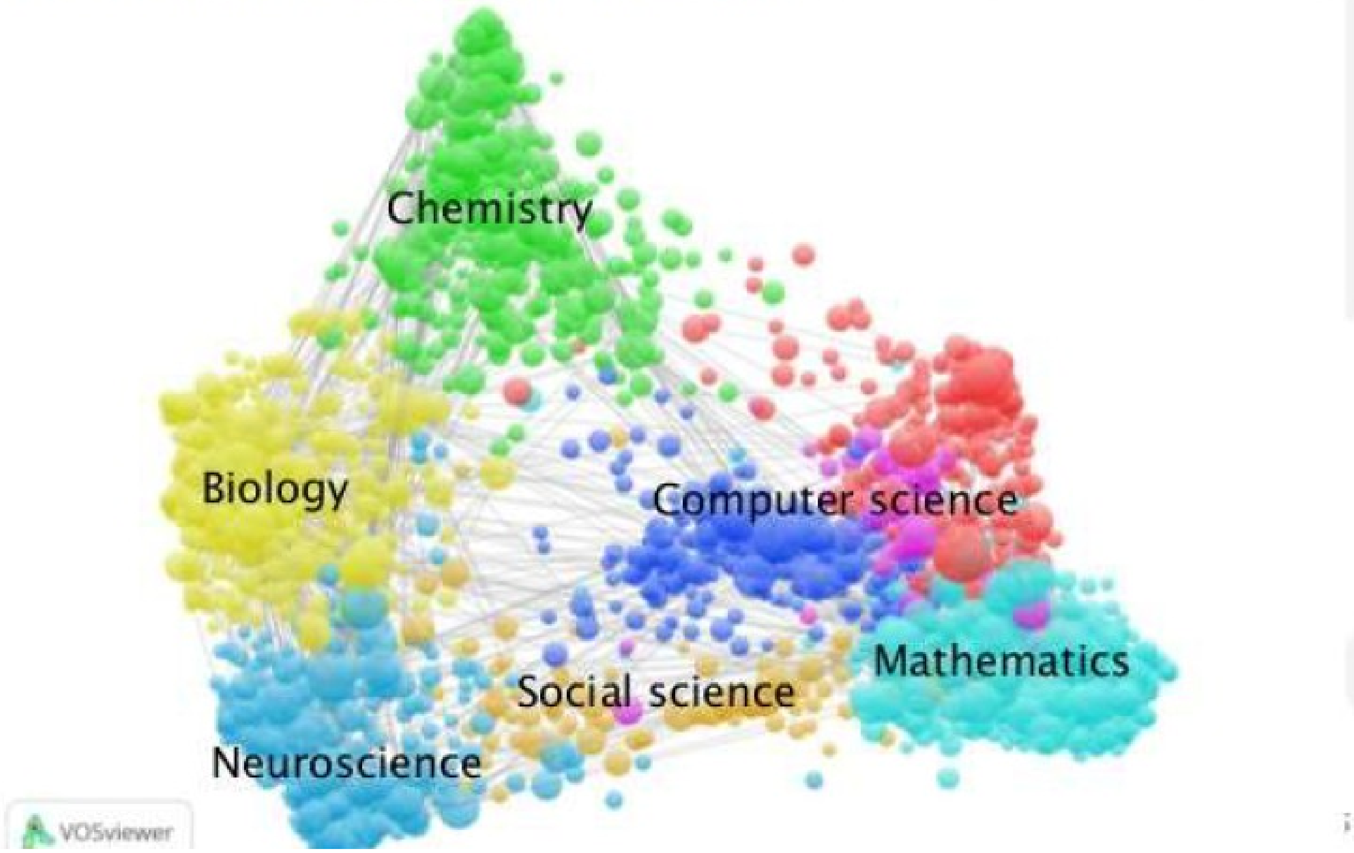
Times Cited – the forward citations – more recently published papers that cite this one.

Related Records – papers which share at least one cited reference in common with this paper. If they share citations, they're likely discussing similar topics.





Co-occurrence relations between terms in network publications



Lots of other data The following surveys feed into WebCASPAR database:

Survey of Earned Doctorates/Doctorate Records File;
Survey of Federal Funds for Research and Development;
Survey of Federal Science and Engineering Support to Universities, Colleges, and Nonprofit Institutions;
Survey of Research and Development Expenditures at Universities and Colleges/Higher Education
Survey of Science and Engineering Research Facilities;
NSF-NIH Survey of Graduate Students & Postdoctorates in Science and Engineering; and
National Center for Education Statistics Data Sources—Integrated Post-Secondary Education Data System (IPEDS)— IPEDS Completions Survey; IPEDS Enrollment Survey; IPEDS Institutional Characteristics Survey Tuition Data; and IPEDS Salaries, Tenure, and Fringe Benefits Survey.

SESTAT (http://www.nsf.gov/statistics/sestat/)is a comprehensive database on education, employment, work activities, and demographic characteristics of scientists and engineers in United States with at least a bachelor's. The SESTAT database includes data from 1993 to "present" from:

- The National Survey of College Graduates;
- The National Survey of Recent College Graduates;
- The Survey of Doctorate Recipients; and
- An integrated data file (SESTAT).

Data are available for download, or through the SESTAT Data Tool, which allows users to generate custom data tables. BUT LONGITUDINAL difficult to deal with. Ginther and Kahn (2015)[2] "SESTAT is collected by the NSF and is the most comprehensive database on the employment, educational, and demographic characteristics of U.S. scientists and engineers available. SESTATyincludes observations from the National Survey of Recent College Graduates (NSRCG), the National Survey of College Graduates (NSCG) and the Survey of Doctorate Recipients (SDR). From the NSCG respondents, SESTAT includes only those who received a degree in STEM or had ever worked in a STEM occupation. From the NSRCG, SESTAT includes recent bachelor's and master's degree recipients in STEM fields. The SDR samples US-awarded PhDs in STEM disciplines. SESTAT oversamples women and under-represented minorities to allow more accurate measures of gender and racial differences.

Within each decade, SESTAT followed individuals through the different waves, adding new people to represent more recent graduates (from the NSRCG). The 1990s panel includes 4 waves: 1993, 1995, 1997, and 1999. The 2000s panel includes 4 waves: 2003, 2006, 2008, and 2010. SESTAT thus includes as many as four observations on a single individual over a 7 or 8 year span in each decade (although for various reasons many people are seen for fewer than four surveys2). Note that there are primarily 2-year gaps between survey waves, although there is one 4-year and one 3-year gap."

Publicly available Census data Sets

American Community Survey – occupation and bachelor's major and higher degree, from Census long form
Current Population Survey – monthly survey to determine unemployment statistics
Decennial Census – detailed data on individuals economics and demographic characteristics

BLS Data Sets

- OES-- Occupational Employment Survey – establishment with ~600 occupations and earnings distribution for occupations, with great industry and city/MSA detail.
- **Census data center research data sets**: Censuses of Businesses, Surveys of Mfg, LEHD that follows workers through unemployment compensation data including income and establishment where they work. Critical to link to other attributes of workers and establishments and firms for link between workers and employers.

  International Data: OECD SCIENCE AND TECHNOLOGY : EUROSTAT;individual countries

2 https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01144/full