ECONOMICS OF SCIENCE: ROBUST FINDINGS, OPEN QUESTIONS AND LINES FOR NEW RESEARCH

Paula Stephan

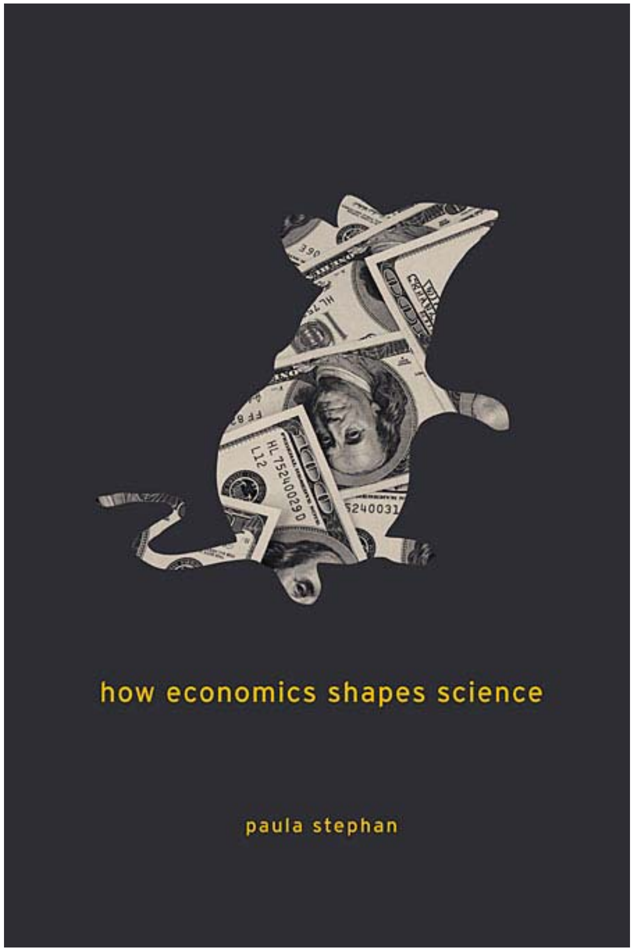
Georgia State University and NBER

ZEW May 2012

Introduction

- During past few years I've had opportunity to think about the economics of science
- One way of summarizing these thoughts is to focus on what I see as robust findings and then look at open questions/lines for new research
- That's what I will do today

Much of discussion based on book



Robust Findings

- Economics is about incentives and costs
- It is also about the study of production—how inputs are used to create output
- The most robust findings are that incentives and cost matter at both the level of the individual scientist and at the institutional level
- Costs also matter, as we will see
- When it comes to the production of scientific knowledge and underlying production function of knowledge, we know that the concept is important but we know considerably less about actual production function

Incentives Matter

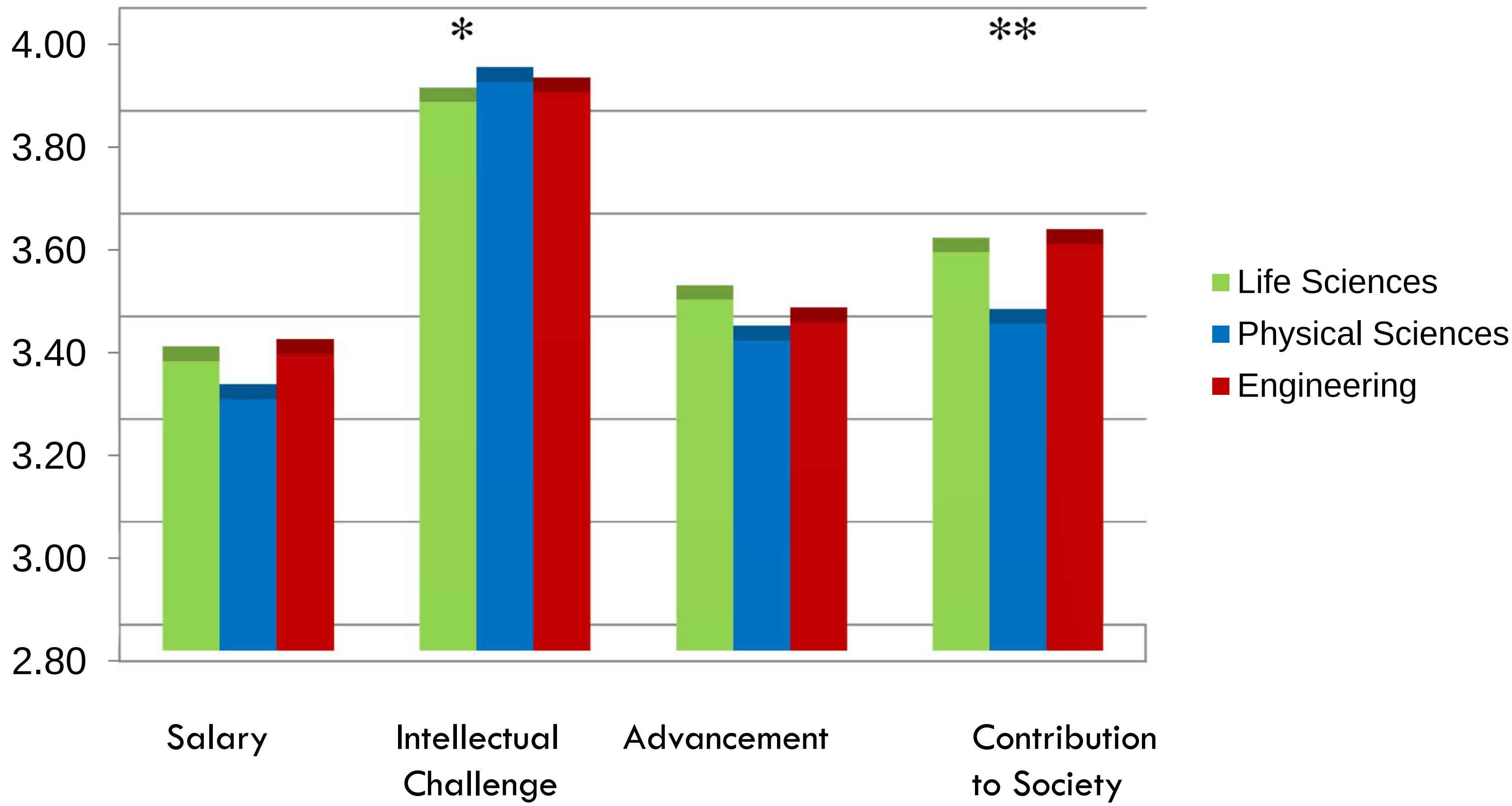
- Puzzle
  - One reason for doing science is the pure “pleasure of finding things out” to quote Richard Feynman. Scientists are clearly motivated by an interest in puzzle solving. For many, it is this interest that attracted them initially to science.
- Ribbon
  - Scientists value the recognition awarded by their peers for being first to make a discovery—to establish priority of discovery
- Gold
  - Scientists are not uninterested in money

Suggestive supporting evidence

- Scientists place highest weight on "challenge" when asked by NSF to score a number of job characteristics
- Scientists chronically argue over issues related to priority; only on rare occasions do they turn down honors associated with establishment of priority; scientists readily adapt to new measures—such as the *h*-index-- of reputation)
- In countries, such as the U.S., where academic salaries vary by institution, scientists move in response to more lucrative job offers

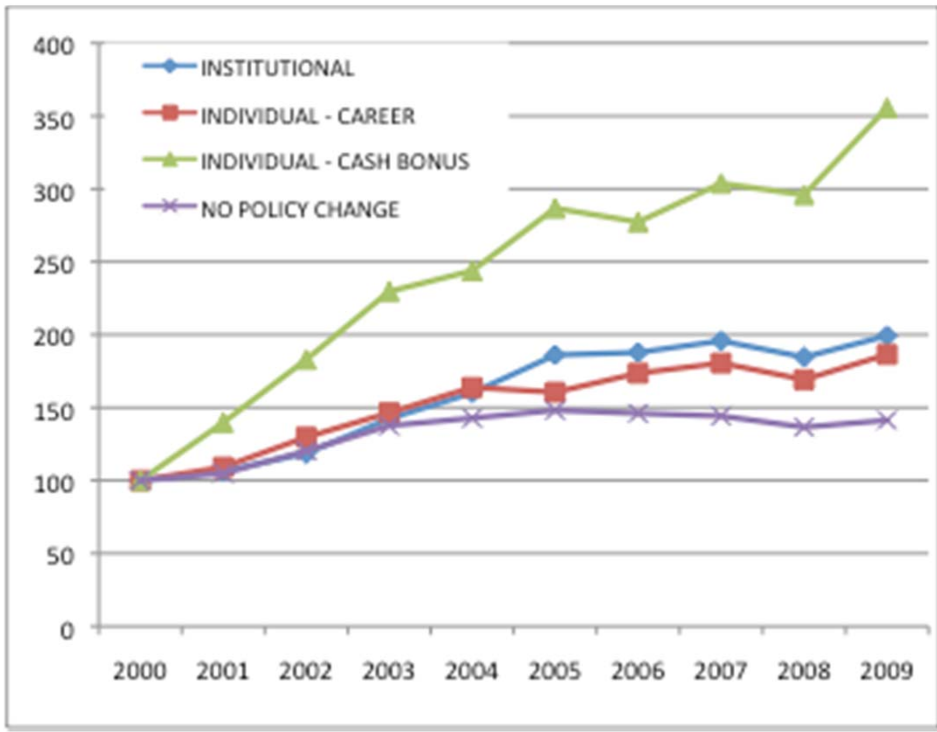
Motives

Motives: “When thinking about a job, how important is each of the following factors to you…” (4-point scale)




Zhong Lin (ZL) Wang's Nanoscience and Nanotechnology group
http://www.nanoscience.gatech.edu/zlwang/
Reader
lin zhong georgia tech
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News (308)
Popular
Home
News
Research
Group
Publications
Dr. Wang
Books
Patents
招聘
Contact
Hot Links:
• Highlights
• Nanotechnology
• Journal Covers
• Lab Tour
• CNCF
• Group Videos
• Group Meeting
• Group Memories
Search
by Google
Web
Nano Site
Z.L. Wang's publications have been cited over 40,000 times!
The publication H-factor is 95 !
"Video of the week: Z.L. Wang on piezoeletronics and piezophotoeletronics. (new!)
We are a leading group in nanoscience and nanotechnology in Georgia Institute of Technology.
Our current research focuses on the fundamental science in the physical and chemical processes in
Selected Current Pub
" High-Output Nanogenerator by Rational Unipolar Assembly of Conical Nanowires and Its Application for Driving a Small Liquid Crystal Display. (new!)
" Self-powered nanowire devices. (new!)
"Flexible High-Output Nanogenerator Based on Lateral ZnO Nanowire Array (new!)
" Strain-Gated Piezotronic Logic Nanodevices. (new!)
more...
Current News
Hot News: Prof. Wang has been elected as a foreign member of the Chinese Academy of Science (new!)
中国科学院
English Version

Response to Incentives to Publish in Top Journals: Submissions by country to *Science*

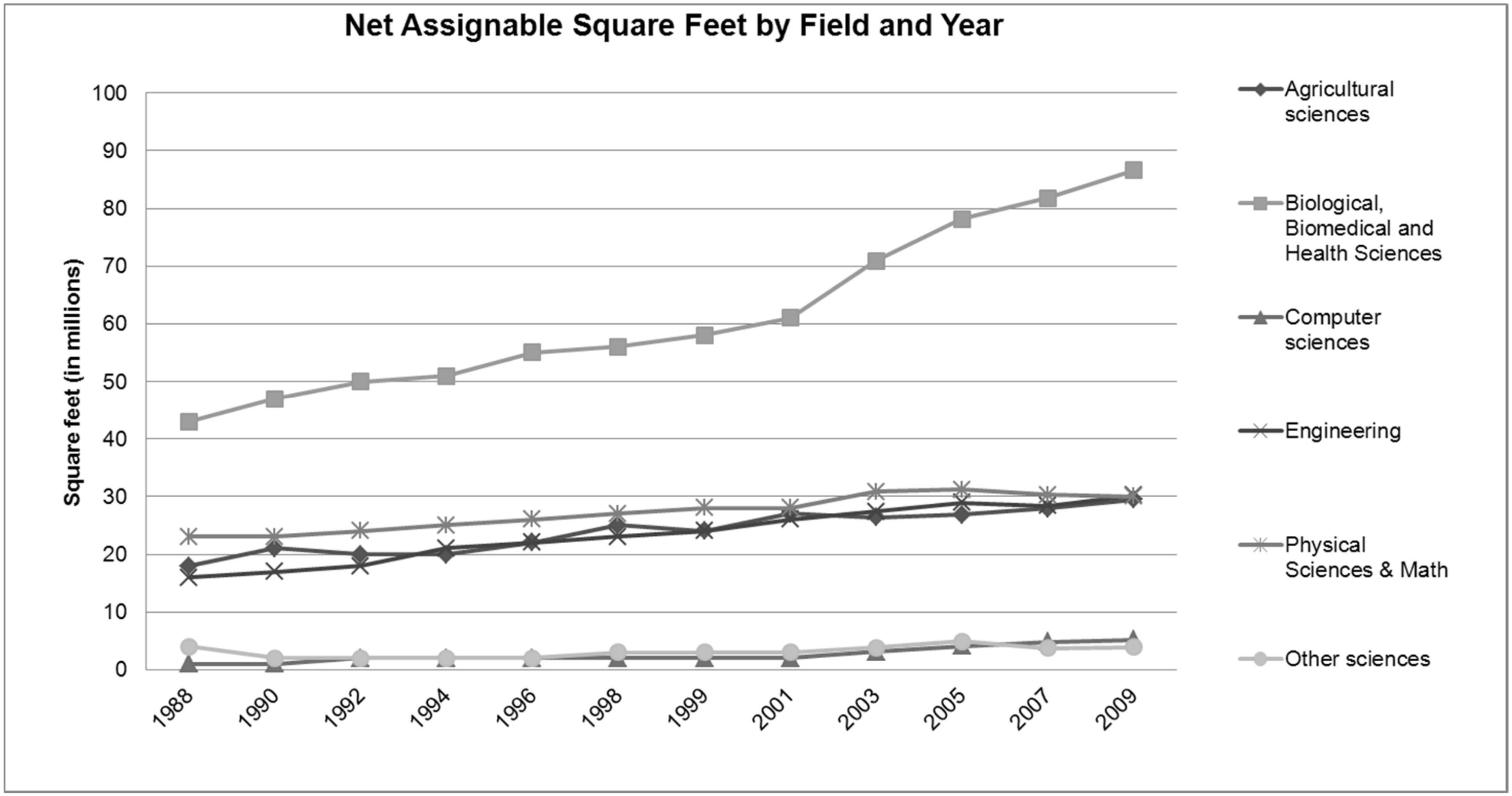


Source: Franzoni, Scellato, Stephan (2011)

Institutions Also Respond to Incentives

- Case of Australia
  - University funds were initially allocated partly on basis of quantity of ISI publications
  - Response: Publications grew considerably; largest increases were in the bottom quality quartile with exception of medical and health sciences where bottom two quartiles grew at a similar rate
- Just-in-time hires in UK in response to Research Assessment Exercise:
  - Between 2002-2006 number of faculty earning more than £100,000 grew by 169%
- Saudi Arabia following similar approach.
  - In an effort to move up in the Shanghai rankings Saudi universities are offering "most cited authors" $72,000 to list university as an affiliation and spend a limited amount of time on campus.

Building Boom in Biological, Biomedical and Health Sciences in U.S.



Not all about puzzle, ribbon and gold

- Other incentives matter, as well
- By way of example, best predictor of which faculty in the life sciences in the U.S. patent is value faculty member places on "contributing to society;"
  - a one standard deviation increase in importance a life scientist places on contributing to society increases expected patent count by almost 50%
- Do not find similar results for other fields—a reminder that incentives vary across fields

Costs Affect Practice of Science

- Important to recognize because cost of doing science—even “small” science-- is non-trivial and growing
- Examples:
  - Telescope can easily cost over 1 billion €
  - LHC cost in excess of $8 billion
  - Cost of researchers’ time: I estimate it costs more than $400,000 to staff a small lab with 8 researchers in the U.S.

Even Mice Costs Money

- Off the shelf mouse cost $17 to $60
- Mutant strains cost $40 to $500-plus
- Cost $1900 to recover a strain from cryopreservation—that's where 67% of lab mice come from
- Designer mice with disposition for such diseases as obesity, alcoholism, Alzheimer's, diabetes, cost considerably more—on the magnitude of $3500

Many Mice Are Used in Research

- Mice are king

- 90% of all animal models are mice
- At least 20 million mice in use in labs
- Johns Hopkins alone has 200,000

Keeping mice

- Costs per day: $.10 to $.18
- Can add up: one researcher was paying Stanford $800,000 a year for mouse upkeep
- At aggregate, spending about $1 billion a year keeping mice

Mouse equipment

- 6 million cages
- New area for innovation:
  - Mouse ultrasound: $150,000 to $400,000.
  - Cage enrichments

Mice continued

- Cost of mouse upkeep factor encouraging Tian Xu of Yale University to work at Fudan University 3 months each year
  - Fudan provides facilities for 45,000 mouse cages (usually 5 to a cage)
  - Could cost over $12,000,000 annually in U.S. to keep.
  - Also issue of where one could keep that many mice in US—more mice than all the mice at Johns Hopkins

Examples of How Costs Affect Practice of Science

- Europe had to “settle” for the E-ELT telescope (extremely large) after plans to build the OWL (overwhelmingly large) telescope proved too expensive and overly complex
- The LHC is shut down in the winter when the price of electricity, due to demand, increases
- Faculty began to substitute postdocs for graduate students in US: reason—they are cheaper, primarily because faculty member does not have to pay for tuition for postdocs and postdocs work more hours (incentives!)

Examples of Costs continued

- Cage rates, which vary considerably across institutions, can play a role in where scientists choose to work
- Costs affect whether researchers work with male or female mice (males turn out to be cheaper)

To Recap

- Considerable evidence that practice of science is affected by incentives and cost
- We have made considerable head way in understanding how these factors affect the practice of science—especially how incentives affect the practice of science, both at the individual as well as at the institutional level; policy makers are beginning to pay attention to these findings
- But scientific results do not just come out of a hat—they involve the combination of inputs—and we know considerably less about this production process

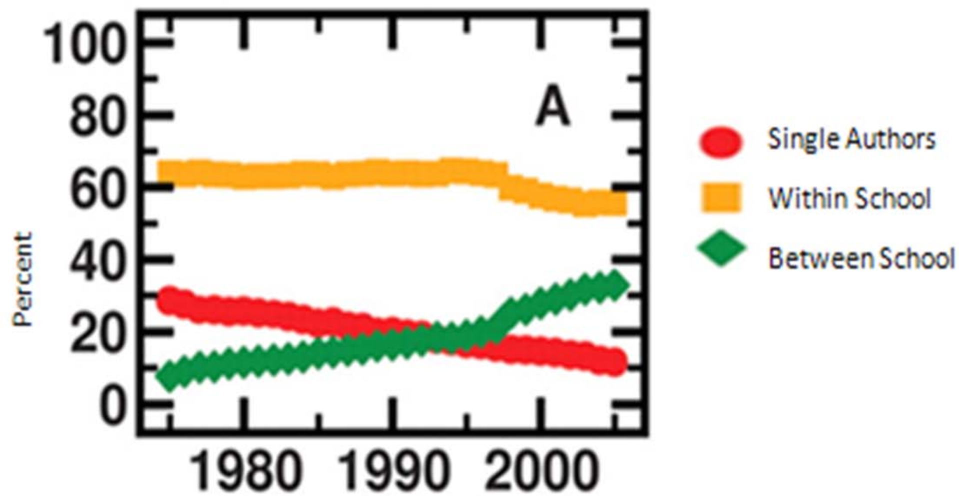
Production Function for Research

- Widely recognized that production of scientific research involves multiple inputs, including knowledge, time, materials and equipment
  - Q=f(k, t, m, e)
  - Some of these inputs, such as knowledge and time, are embodied in people
- Despite this understanding, almost all research in economics of science related to productivity focuses on examining relationship of output to people
  - Some examines individual productivity
  - Some examines patterns of collaboration among researchers overtime and how these change
  - Some examines location of collaborators and relationship to productivity

Robust Findings Here

- Collaboration is on the increase
- Collaborative science often produces better science
- Some evidence that collaboration across institutions is associated with "better" science

The Rise in Multi-University Collaboration, Science and Engineering



Source: B. F. Jones et al., Science 322, 1259 -1262 (2008)

But Numerous Areas of Ignorance When It Comes to Collaboration

- Know virtually nothing about how structure of teams relates to productivity
  - Age structure
  - Position—postdoc, graduate student mix
    - Particularly important to investigate given way labs are staffed in the U.S. by graduate students and postdocs and that such a staffing model results in overproduction of PhDs in terms of research jobs
  - Number of collaborators
  - Mix of fields
  - International collaboration vs. national collaboration
  - Role of equipment and materials in collaboration

More Ignorance Regarding Role of Equipment

- Types of questions that can be asked
  - What happens to capital-labor ratio in the lab?
  - How efficient are markets for scientific equipment? What is extent of price discrimination?
  - How quickly does new equipment diffuse?
  - How much does research depend upon having a monopoly on certain types of equipment?
  - How much does equipment dictate where research is performed, both in terms of number of research centers and distinction between private and public sector?
  - What role does equipment play in recruitment of scientists?

Examples of new equipment

- Large scale
  - Telescopes such as E-ELT (European Extremely Large Telescope or TMT (thirty meter telescope))
- Small scale
  - RatCap--wearable PET scanner—developed at Brookhaven
  - Digital ultra sound for mice: costs between $200,000 and $400,000. Market said to be brisk.

Small Scale

RatCap

Vevo 2100

Large Scale

Example of Role Equipment Plays

- Can be seen by examining what has happened in past 25 years in sequencing
- Provides good examples of importance of equipment and how it affects production of research and research questions that can be examined

Sequencing—two minute history

- Chain termination method of sequencing developed by Frederick Sanger (University of Cambridge) mid-1970s.
  - Shared Nobel Prize in Chemistry (his second) in 1980 for development.
- Procedure became less labor intensive with invention in 1986 of “DNA sequencer” by Leroy Hood and colleagues.
- During 1990s and early 2000’s improvements made to technology which greatly expedited speed of sequencing and lowered costs of base pair

Summary of Pattern over 18 Years

- When Human Genome Project (HGP) began in 1990, best equipped lab could sequence 1000 base pairs a day
- By January 2000, 20 labs involved in mapping were collectively sequencing 1000 base pairs a second, twenty-four/seven.
- Cost per finished base pair fell from $10.00 in 1990 to under $.05 in 2003 to $.01 by 2007. Ancient history.
- Measured in terms of base pairs sequenced per person per day, productivity of a researcher operating multiple machines increased more than 20,00 fold from early 1990s to 2007, doubling approximately every 12 months.
- Innovations in sequencing equipment and microprocessor technology led announcement of working draft of genome to be made 5 years early and within budget

New Generation of Machines

- Began coming on market in 2005
- Read millions of sequences at once although length of base that was read was shorter
- Jonathan Rothberg invented first of these—sold by his company 454
- Three other next generation machines—one from Helicos, one from Applied Biosystems, one from Illumina, entered market

Marketed Equipment Creatively and Aggressively

- 454 ran full page ads in prominent journals such as *Nature* and *Science* for its FLX system
- Got Broad Institute (directed by Eric Lander—first author of first draft of HG) to adopt their machines
- Mapped first million base pairs of Neanderthal genome (apparently with what is now known as errors)
- Mapped James Watson's genome

Price and market of second generation equipment

- First generation Applied Biosystems machines cost about $300,000
- Illumina's next generation machine cost about $470,000
- Estimated that approximately $1.5 billion a year for the hardware sold to scientists.
- Consumables not insignificant; make up largest portion of Illumnia's revenues

http://www.genomeweb.com//node/1061546?hq_e=el&hq_m=1260235&hq_l=3&hq_v=2d02897a45

Where Sequencing Stood in 2011

- In 2007, a single sequencing run produced maximum of around one gigabase of data
- By 2011, rate had nearly reached a terabase of data in a single run—nearly a 1000x increase
- In 2011 researchers could sequence more than 5 human genomes in a single run, producing data in about a week for a cost of less than $5,000 per genome
- By comparison, spent $3 billion and 10 years to sequence first published genome in 2003

(An Introduction to New-Generation Sequencing Technology, Illumina)

Third Generation Equipment Began Entered Market in late 2010

- March 2010 Rothberg introduced a silicon chip sequencer—analogy for some is when photography went from film to digital
- Manufactured by his latest company, Ion Torrent Systems, recently bought out by Life Technologies (7 page ad in recent issues of *Science*)
- Machine sold for $50,000 when came on market in January 2011
- Published an article in *Nature* describing it; sequenced Gordon Moore's genome of Moore's Law
- Consumables are expensive;
- Fits on desk top


personal genome machine
ion torrent
ion torrent

New Equipment Emerging in late 2011

- Introduced by number of companies, including Life Technologies (claims they will have a $1000 whole genome in one day by year's end) and Illumina, the leader in sequencing equipment
- In February of 2012 Oxford Nanopore introduced a device the size of a USB memory stick called a MinION, which will be sold for less than $900 and supposedly can deliver 150 megabases of DNA sequences per hour. Larger version will also be marketed
- High error rate of 4%

Implications for Way Science Is Done: Location

- Introduction of first sequencing machines led sequencing activity to shift to core facilities at biomedical research centers. Researchers sent material to the core facility to be sequenced. Required trained personnel and continuity
- Second generation equipment not cheap—but speed and lower unit costs means equipment has potential of being used more widely
  - Illumina markets its equipment with this in mind: "Genome Analyzer System enables even the smallest lab to have sequencing capabilities of the largest genome centers."
  - Despite this, estimated that half of the 1400 DNA-sequencing machines in world are at 20 big academic or government research centers. (Matthew Herper).
- Rothberg hopes third generation machine will be widely used—goal to open sequencing field to hundreds of small groups that lack access—including doctors'
- Clear intent of companies such as Oxford Nanopore is to vastly expand the market
- In past, companies like Illumina have grown by selling new equipment to same research centers

Implication for What Is Sequenced and Data Storage

- As price falls, can sequence almost anything
- Presents issue of what researchers will do with the data

Implications for Way Science Is Done: Capital-Labor Ratio

- Ratio of capital to labor depends on relative prices and technology—increase in relative price of labor should lead to substitution of capital for labor
- Amount of labor used also depends on scale of operation
- Currently substitution effect seems to be dominating scale effect:
  - Venter Institute eliminated 29 sequencing center jobs about 5 years ago
  - Broad Institute eliminated 24 three years ago
  - Difficult to track what has happened at "core" facilities at universities
  - Hard to know if new generation equipment will change this

Implications for Way Science Is Done: Business Model

- Who does the sequencing? Private or public sector?
- Some companies are betting that in end researchers will outsource sequencing needs: Complete Genomics of Mountain View, CA for example, recently sequenced material supplied by Lee Hood

Questions/issues apply not only to sequencing

- Many of these questions/issues apply to other fields of science:
  - Business model question
  - Capital/labor issue
  - Role of equipment in determining co-authorship
  - Role of equipment in affecting where scientist chooses to work
  - Role of equipment as it relates to issues of efficiency

Implications for Role of Equipment in Determining Co-authorship

- In many labs certain people dedicated to running equipment
- At certain large pieces of equipment, individuals work on site who run the equipment
- This plays a role in determining co-authorship
- In some large experiments—such as those at CERN or at telescopes—leads to extraordinarily large number of co-authors.
  - Easier to bestow authorship on all rather than determine who did and did not contribute.—Ice Cube
  - Relates also to large fixed cost of designing and building equipment and investments made by team.

Implications for Way Science Is Done: Efficiency

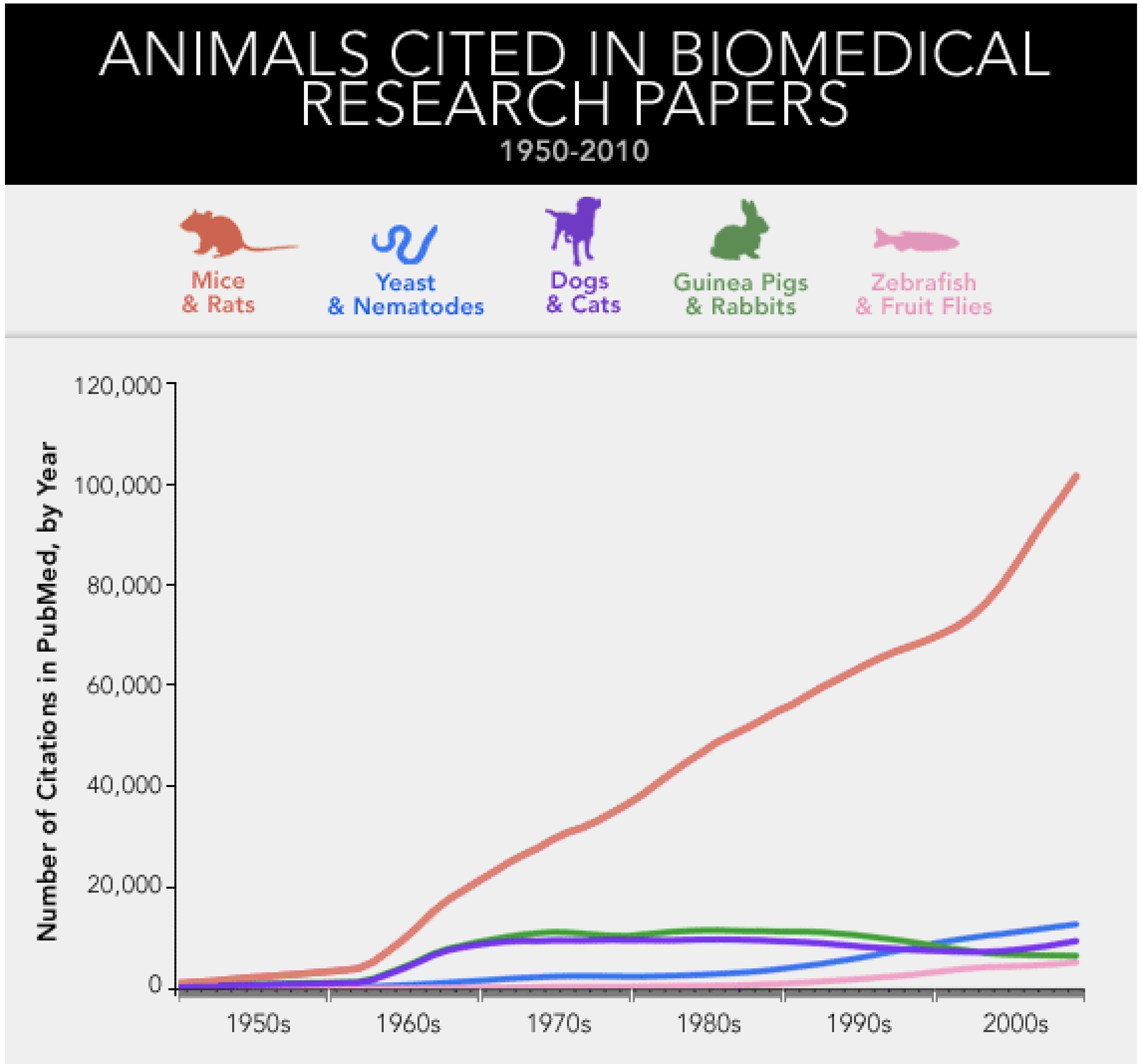
- Casual empiricism suggests market for scientific equipment is highly concentrated
- Market supplied by for-profit firms, unlike market for mice that is supplied primarily by a non-profit organization (JAX)
- Estimated that Illumina controls 2/3rds of sequencing market one year ago (another, more current estimate is 60% of next generation market); 90% of world's DNA sequencing output is being done on Illumina machines ("Five Reasons Illumina Should Fight Roche's Insulting Low-Ball Bid)
- Prices not advertised
- Clearly price discrimination in market—one-to-one negotiations—depends on "how hard you bargain"
- How much lack of efficiency is there? How much gained by price discrimination?
- How large must profits be to entice suppliers into these markets?

Implications for Way Science Is Done: Equipment Monopoly

- How much does a scientist's research depend upon having a monopoly on new types of equipment (or a research model, such as a mouse)?
- Or securing a monopoly on time slot on scarce resource such as a telescope or on a submergence vehicle such as Alvin?

Implications for Way Science Is Done: Diffusion

- How fast does new equipment diffuse?
- Reasonably tractable, although only person I know to be studying the question is Chiara Franzoni
- Tractable because for purposes of replication published papers generally note (by “brand” name) equipment used in the experiment
- Note that diffusion of animal models has been tracked through National Library of Medicine



Survey National Library of Medicine; Daniel Engber, Mouse Trap, *Slate,* Nov. 16, 2012

Implications for Way Science is Done: Rate of Productivity

- Technological advances in equipment explain part of the increase in number of articles scientists write
- It also explains change in focus of research—especially at dissertation stage
  - Used to be someone could map one gene for a dissertation
  - Or identify one protein

Implications for Way Science is Done: Location of Scientist

- Biophysicist Lila Gierasch was "wooed by an NMR machine" to the University of Texas Southwestern Medical Center after she repeatedly had difficulty obtaining funds to purchase a high-field NMR—which cost anywhere from $2 million to $16 million.
- Not only location within the non-profit sector.
  - Equipment plays a role in sector scientists choose to work in: "I have worked in some of the best-funded academic laboratories in the world and even these labs don't have access to the fancy next-generation machines in a way that large biopharmaceutical companies do."
  - To what extent is this occurring? How does it relate to productivity divide between two sectors? How does it affect ability of private sector to attract and recruit scientists? How does it affect collaboration across sectors?

Other Open Questions/Lines for New Research

- Major Efficiency Questions
  - Are we spending the “right” amount on R&D in the public sector?
  - Is current allocation of funding for R&D which—at least in U.S. gives about 2/3rds to the biomedical sciences--efficient?
  - Are grants structured in an efficient way in terms of
    - size, duration, criteria for evaluation and number of people?

Difficult but important questions

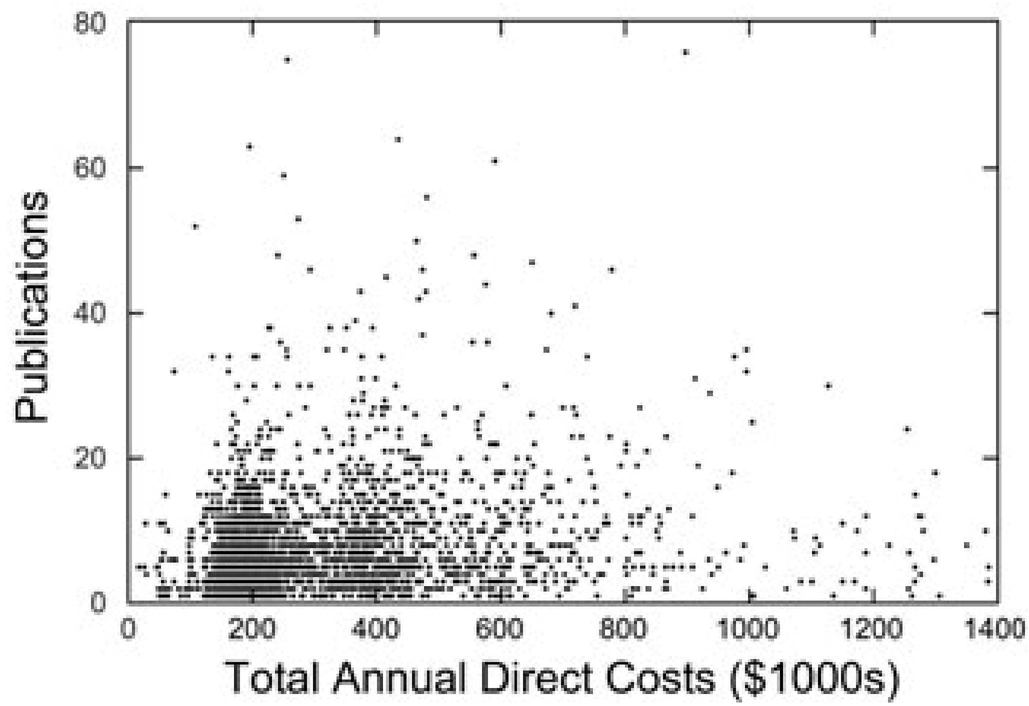
- Especially in an era of flat resources
- Some, due to measurement problems, may never be answerable
- With regard to amount…case can be made that we don't know the "right" amount but the research that has been done shows reasonable returns and suggests that we are underinvesting

What about Mix?

- There have been impressive returns from research in the biomedical sciences
- But is marginal benefit from another dollar spent in biomedical sciences equal to that in other sciences?
- Case could be made that it is lower
  - Spillovers and complementarities from other disciplines
    - lasers, MRI, etc. are important
  - Drug discovery model has produced few winners in recent years suggesting diminishing marginal productivity

Structure of Grants

- Fund people rather that proposals?
- Fund collaborative groups rather than individuals?
- Are rules—such as requirement of EU to have researchers from three or more countries—efficient?
- Large grants or small grants?
  - Did NIH use their funds efficiently during the doubling of its budget?
  - NIGMS study suggests the answer may possibly be “no.”
    - Found a correlation of only .14 between number of publications and total annual direct cost of grants



https://loop.nigms.nih.gov/index.php/2010/11/22/
another-look-at-measuring-the-scientific-output-and-impact-of-nigms-grants/
A plot of number of grant-linked publications from 2007 to mid-2010 for 2,938 investigators who held at least one NIGMS R01 or P01 grant in Fiscal Year 2006 as a function of the total annual direct cost for those grants.

Summarize

- Many robust findings—especially that incentives and cost play an important role
- Many open questions/lines for new research
  - Production of science
    - Structure of collaboration
    - Role of equipment and materials in production of new knowledge
  - Efficiency issues when it comes to funding of science

Comments/Questions?

- pstephan@gsu.edu