NBER Innovation Research Boot Camp

Human Capital and Innovation

Benjamin F. Jones
July 18, 2022

Human Capital and Innovation

- Growth Viewpoints
- Micro Viewpoints
  - Collaboration
  - The Life Cycle
- Labor Supply & the Availability of Talent
  - Can we accelerate growth?
  - The role of stars and the "Lost Einsteins" question

Growth Viewpoints

The "Knowledge Production Function"

❖ <u>Question</u>: How do inputs to innovative activity map into innovations and new firms?

$$\frac{dA}{dt} = q(H, K, Z, A)$$

❖ Inputs

- Human capital (H), physical capital (K)
- Institutions (Z)
- Current state of ideas (A)

❖ How do we understand the role of human capital, especially in light of views/models of the creative process?

Endogenous Growth Viewpoints: Romer Approach

❖ Romer (1990) et cetera assumed

$$\frac{dA}{dt} = \delta A L_A$$

Inputs are (1) effort (L_A) and (2) current stock of ideas (A)

❖ Growth rate in economy (divide through by A) is

$$g_A = \delta L_A$$

Implication: growth rate follows the *level* of innovative effort

Recall Growth Theory's Second Approach

❖ C. Jones (1995) showed empirically that a constant growth rate appears consistent with *growing* innovative effort, not with a constant *level* of innovative effort

❖ Led to a generalization

$$\frac{dA}{dt} = \delta A^{\phi} L_A^{\lambda} \quad \Rightarrow \quad g_A = \frac{\lambda}{1-\phi} g_{L_A}$$

❖ Allows limited increase in idea production *per researcher* along the growth path ($0 < \phi < 1$); perhaps even declining idea production per researcher ($\phi < 0$).

❖ Now we need to grow labor supply

Digging Deeper:
Three Models of the Creative Process

❖ Fishing Out (Kortum 1997)

❖ Recombination (Weitzman 1998)

❖ Burden of Knowledge (B. Jones 2009)

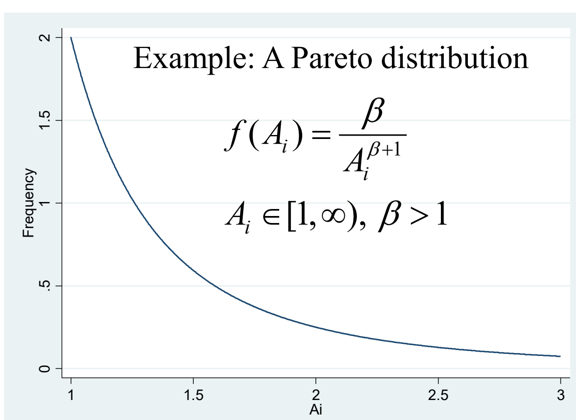
Thinking especially about how current stock of ideas influences further idea production, based on particular views of the *underlying creative process*.

This leads to insights (from simple to more complex) about the role of innovative labor supply / human capital.

Model #1: Creativity as Going Fishing

❖ Draw ideas from a stationary distribution of possible ideas. Implement idea if better than best draw in past.

❖ Fishing out: harder and harder to beat existing best draw



Cumulativeness and Fishing Out

$$\frac{dA_{max}}{dt} = \underbrace{\theta L_A}_{\{1\}} \times \underbrace{\Pr(A_i > A_{\max})}_{\{2\}} \times \underbrace{E[A_i - A_{\max} \mid A_i > A_{\max}]}_{\{3\}}$$

{1} Idea rate per person x no. of people

{2} Probability idea is better than best idea yet

{3} Expected size of innovative step

❖ Pareto distribution has nice properties

$$\{2\} = A_{\max}^{-\beta}$$

$$\{3\} = A_{\max}/(\beta - 1)$$

$$\frac{dA_{max}/dt}{A_{max}} = \frac{\theta}{\beta - 1} L_A A_{\max}^{-\beta}$$

$$\boxed{\Rightarrow g_A = \frac{1}{\beta} g_{L_A}}$$

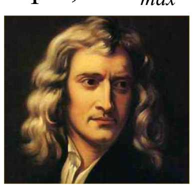
❖ In this model, $\phi = 1 - \beta < 0$ (and no crowding, $\lambda = 1$)

- ❖ *Negative* inter-temporal spillovers
- ❖ Need people growth for steady state productivity growth

Model #2: Creativity as Recombination

❖ The fishing out mechanism, taking the idea distribution as stationary, leads strongly towards $\phi < 0$

❖ Another approach imagines that the idea distribution shifts. For example, as A_{max} increases we see new possibilities



"If I have seen further it is by standing on ye sholders of Giants."
- Isaac Newton

❖ Many inventions rely on new ideas. As we have new ideas, we replenish the idea distribution.

❖ One version of 'new possibilities' emphasizes <u>recombinant</u> nature of creativity (Weitzman 1998)

Creativity as Combinations

- Darwin: evolution as random mutation (new) + selection (old: animal husbandry)
- Edison: Light bulb = candle (old) + electricity (new)
- Mullis: DNA replication technology = DNA (new) + polymerase enzyme (new)

Cumulativeness as an Improving Distribution

- Example: A Pareto distribution where new ideas are always better than current best idea

$$f(A_i) = \frac{\beta A_{\max}^{\beta}}{A_i^{\beta+1}}, \quad A_i \in [A_{\max}, \infty), \ \beta > 1$$

So idea production function is now

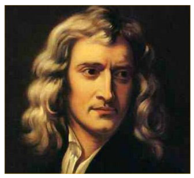
$$\frac{dA_{max}}{dt} = \theta L_A \times \underbrace{\Pr(A_i > A_{\max})}_{\{2\} = 1 \ now} \times \underbrace{E[A_i - A_{\max} \mid A_i > A_{\max}]}_{\{3\} = A_{\max}/(\beta - 1)}$$

$$\frac{dA_{max}/dt}{A_{max}} = \frac{\theta}{\beta - 1} L_A$$

We return to initial Romer world: growth follows effort *level*. But of course this does not appear valid empirically…

Model #3: Cumulativeness and Human Capital

❖ As a phenomenon, we do seem to draw from evolving idea distributions. Fields see new theories, facts, techniques, and other complementary inputs that allow new insights



"If I have seen further it is by standing on ye sholders of Giants."
- Isaac Newton

❖ But to use such knowledge, we first have to learn it. What happens if one must first climb up the Giants' backs?

❖ *Cumulativeness* may impose increasing *human capital* investment challenges on the young

The Burden of Knowledge

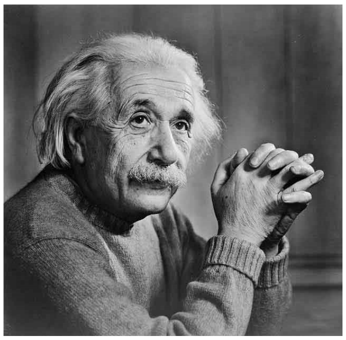
❖ What happens if new ideas, by creating new knowledge, impose an increasing educational burden on future innovators?

❖ Two margins of response

- Spend more time in training
- Choose narrower expertise

❖ Implications

- Individual innovators are less capable
  - Less time to innovate if more time in training
  - Harder to have broad impact if narrowing expertise
- Greater need for collaboration in research

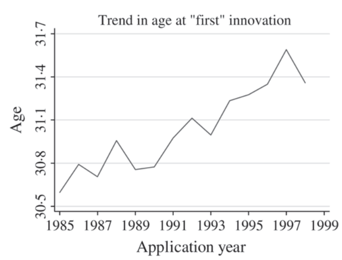


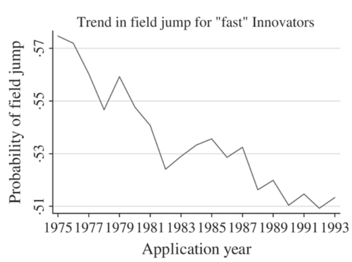
"…knowledge has become vastly more profound in every department of science. But the assimilative power of the human intellect is and remains strictly limited. Hence it was inevitable that the activity of the individual investigator should be confined to a smaller and smaller section…"

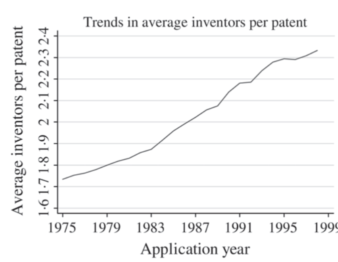
-- Albert Einstein (1932)

The Burden of Knowledge: Some Evidence

❖ Micro-evidence from patent data (B. Jones 2009)

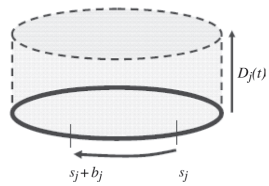






The Burden of Knowledge and Growth

❖ Focus on creativity effect of narrowing expertise. Consider a "circle of knowledge" with a continuum of knowledge types (indexed by *s* around circle) where depth of knowledge is *D(t)*



❖ Let educational attainment for innovator born at time *t* be their breadth (b) times the prevailing depth (D)

$$E(t) = b(t)D(t)$$

❖ Let creativity (for an individual) be

$$\dot{A} = A^{\chi} L_A^{-\sigma} b^{\beta}$$

The Burden of Knowledge and Growth

❖ Let the depth of knowledge follow the stock of existing ideas

$$D = A^{\delta} \Rightarrow g_D = \delta g_A$$

❖ In equilibrium, individuals choose educational expenditure as a constant fraction of lifetime income, *E(t)/y(t)* = *c*, implying

$$g_E = g_y = g_A$$

Thus individual educational attainment grows along the growth path at the same growth rate as the economy

❖ From educational attainment equation we then have

$$E(t) = b(t)D(t) \Rightarrow g_b = g_E - g_D \Rightarrow$$
$$g_b = (1-\delta) g_A$$

The Burden of Knowledge and Growth

❖ Growth rate of economy is

$$\dot{A} = A^{\chi} L_A^{1-\sigma} b^{\beta} \quad \Rightarrow \quad \dot{A}/A = A^{\chi-1} L_A^{1-\sigma} b^{\beta}$$

❖ For steady-state growth, take logs, differentiate with respect to time, plug in $g_b = -(\delta-1)g_A$ and rearrange:

$$\boxed{g_A = \frac{1-\sigma}{1-\chi+\beta(\delta-1)} g_{L_A}}$$

Relating to prior models, the crowding term is $\lambda = 1 - \sigma$, and the inter-temporal spillover term is $\phi = \chi - \beta(\delta-1)$

❖ With increasing specialization ($\delta>1$), we can now have rich idea possibilities (large χ) and yet still explain macro facts ($\phi<1$) because individual innovators see narrowing share

Human Capital and Innovation

❖ Growth Viewpoints

❖ **Micro Viewpoints**
- Collaboration
- The Life Cycle

❖ Labor Supply & the Availability of Talent
- Can we accelerate growth?
- The role of stars and the "Lost Einsteins" question

Cumulativeness and the Burden of Knowledge: Microeconomic Dimensions

- If knowledge accumulates as science advances, then training decisions naturally shift

(1) Extend training

Innovations less common at young ages

Life-Cycle Changes

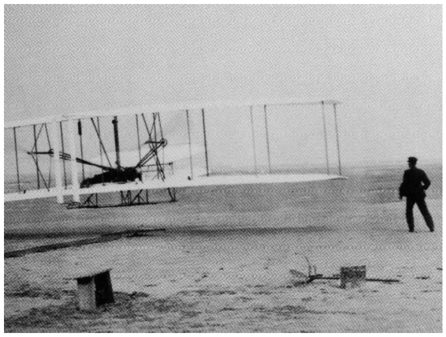
(2) Choose narrower expertise

Innovators increasingly work in teams

Organizational Changes

(Two dimensions of response)

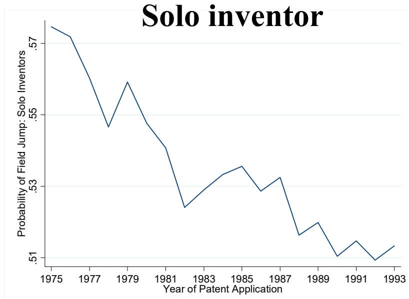
Collaboration



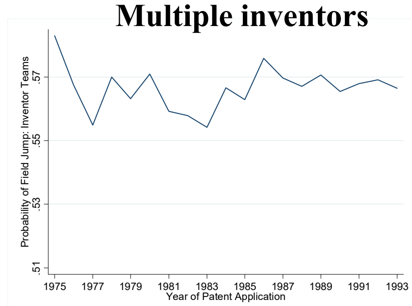
Specialization & Collaboration

❖ Do you switch fields between consecutive patents? (Jones 2011)

Solo inventor



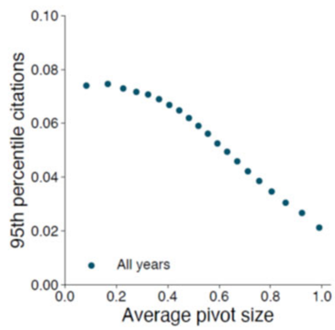
Multiple inventors

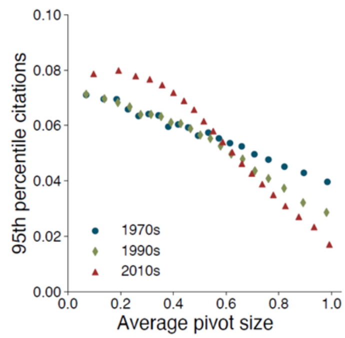


- Solo inventors appear increasingly narrow
- Teamwork is associated with sustained breadth

Specialization & the Pivot Penalty

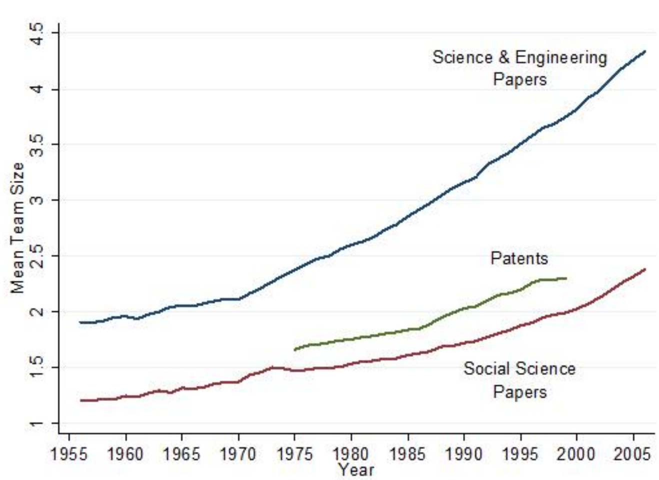
❖ Measure "pivot size" as how far you move in a given paper or patent from your recent work (Hill et al. 2022)





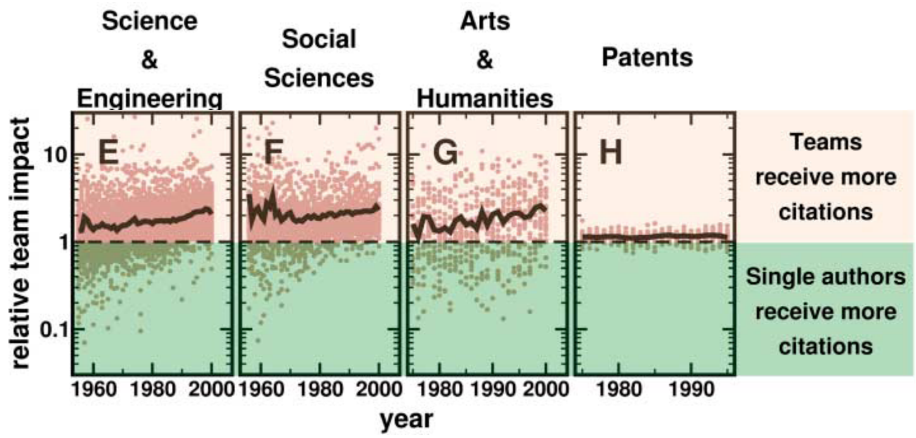
- Find that work has lower impact the further you pivot.
- And this pivot penalty is getting steeper with time.

The Ubiquitous Rise in Teamwork



Data: Web of Science, 19 million articles (Wuchty et al. 2007)

The Rising Team Impact Advantage



Data: Web of Science, 19 million articles (Wuchty et al. 2007)

The Team Advantage Today

	Mean Citations Received			Probability > 100 citations		
	Team	Solo	Team/Solo	Team	Solo	Team/Solo
Science and Engineering	11.95	4.55	2.63	1.21%	0.28%	4.25
Social Sciences	8.74	3.31	2.64	0.59%	0.13%	4.57
Patents	6.66	5.64	1.18	0.025%	0.015%	1.65

❖ Teams have a large and increasing advantage in producing the highest impact ideas

Question: Team Organization

❖ How should you organize individuals into teams?

How do you allocate pilots across planes?

How do you allocate musicians across quartets?

Question: Individual Assessment

How should you credit the individual member?

e.g., if team output is good, is it due to the best person (pilot-copilot) or must all the team members be good (string quartet)?

Method: Key Idea

$$y = \beta_n \left[\frac{1}{n} \sum_{i=1}^{n} a_i^{\rho} \right]^{\frac{1}{\rho}}$$

- The "Hölder Mean" (a.k.a. "Generalized Mean", CES)
  - y is a measure of impact, a_i is productivity of individual i, and n is team size

- Special cases
  - max ($\rho \to \infty$)
  - min ($\rho \to -\infty$)
  - arithmetic mean ($\rho = 1$)
  - geometric mean ($\rho = 0$)
  - harmonic mean ($\rho = -1$)

Potential Team Advantage

$$y = \beta_n \left[\frac{1}{n} \sum_{i=1}^{n} a_i^{\rho} \right]^{\frac{1}{\rho}}$$

- β_n captures impact benefit associated with team of size n, incl. advantages of aggregating effort, skill, marketing, or disadvantages via coordination costs (Wuchty et al. 2007; NAS 2015).
- Normalize by setting $\beta_1 = 1$ for solo-authored work.
  - Thus $y = a_i$ for solo-authored work => individual productivity measured on the scale of outcome metric.
  - $\hat{\beta}_n$ interpreted as the impact advantage of teamwork over solo-work for individuals that share a common individual productivity level.

Data Sets

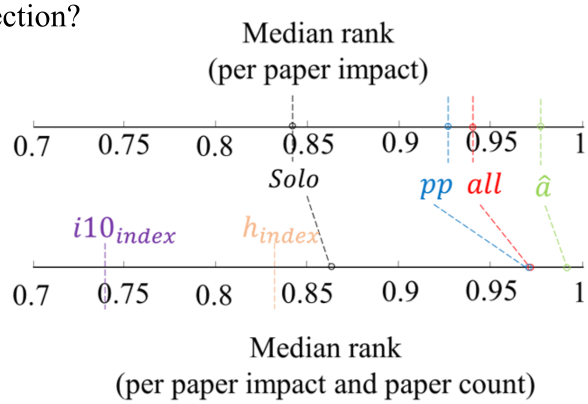
- Web of Science, 1945-2005
  - All 185 science & engineering and social science fields with $\geq$ 500 papers
  - Author name disambiguation from WOS (Bai 2016)
- USPTO, 1975-2006
  - All 384 tech classes with $\geq$ 500 patents
  - Inventor name disambiguation from Li et al. (2014)
- Restrict to papers/patents with <= 8 team members
  - 97% of papers and 99% of patents
  - 24 million journal articles, 13 million authors (WOS)
  - 3.9 million patents, 2.6 million inventors (USPTO)

Results: Team Production Function Parameters

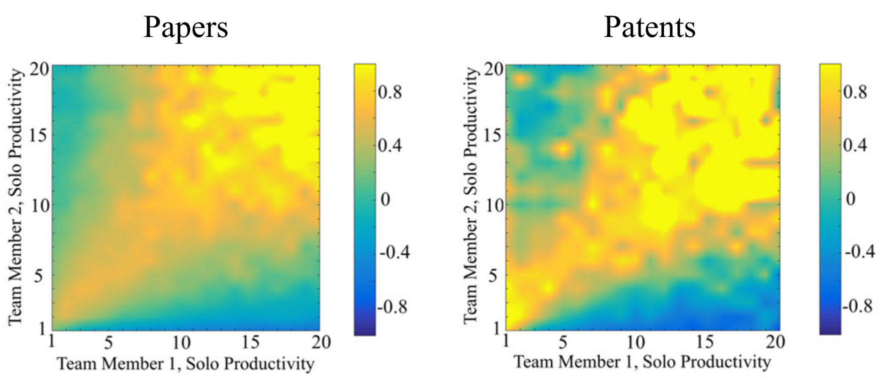
- $\hat{\rho} < 1$
  - Universal across fields
  - Centers between geometric and harmonic averages
  - Implication: Greater influence of <u>lower-productivity team members</u>
- $\hat{\beta}_n > 1$
  - For all WOS fields and 94% of patenting fields
  - $\hat{\beta}_2 = 1.85$ (papers) and $\hat{\beta}_2 = 1.44$ (patents)
  - Implication: <u>team advantage</u> exists conditional on quality of individual team members.

Career Metrics: Election to the NAS

- Consider capacity to predict who is elected to the National Academy of Sciences
- Take all individuals in a given field and cohort. Where do the NAS members rank in their cohort at time of election?



Final Implication: Matching in Team Assembly



- ❖ Find strong tendency toward positive assortative matching

Summary: Collaboration

- Big dynamics
  - People increasingly work in teams in all fields
  - Highest impact ideas increasingly come from teams
  - Researchers experience increasing impact penalties when moving into new areas

    (Burden of knowledge reasoning may explain patterns)

- Team production: Innovation teams appear like "string quartets." Consistent with "specialist" teams.
  - Positive assortative matching

- Individual assessment: How we credit individuals is essential to career progression, incentives, etc.
  - "Decoding teams" method to confront teamwork challenge

<u>Human Capital and Innovation</u>

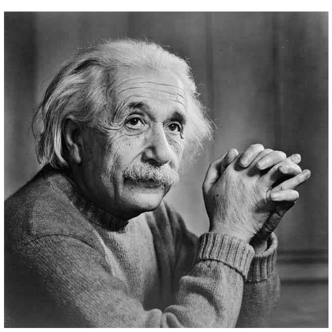
- Growth Viewpoints

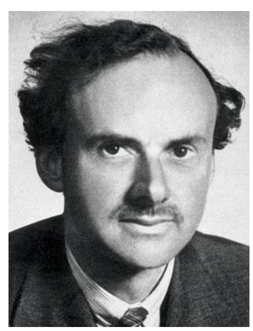
- Micro Viewpoints
  - Collaboration
  - **The Life Cycle**

- Labor Supply & the Availability of Talent
  - Can we accelerate growth?
  - The role of stars and the "Lost Einsteins" question

When in life is one most innovative?

Common Views





"A person who has not made his great contribution to science before the age of thirty will never do so."

(Einstein)

"Age is, of course, a fever chill
that every physicist must fear.
He's better dead than living still
when once he's past his 30th year."

(Dirac)

Why These Views?

- Young people sometimes thought to have advantages in:

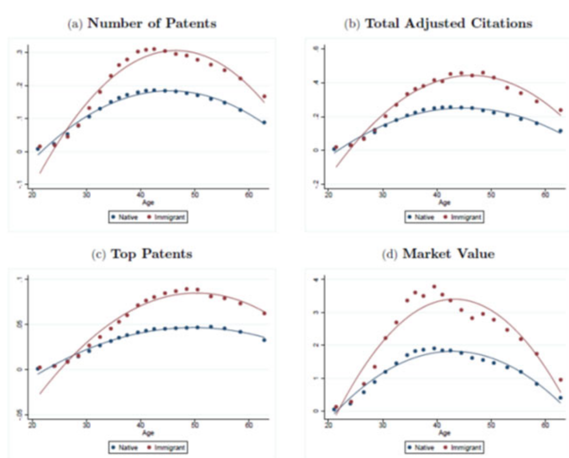
1) Deductive reasoning (e.g., Galenson and Weinberg 2005)
2) Transformative thinking (e.g., Planck 1949, Weinberg 2007)
3) Energy / Time (e.g. Jones et al. 2014)

…Yet key resources may accumulate with age

- Human capital, Financial capital, Social capital (e.g., Lazear 2004, Chatterji 2009, Jones 2009, Evans and Jovanovich 1989, etc.)

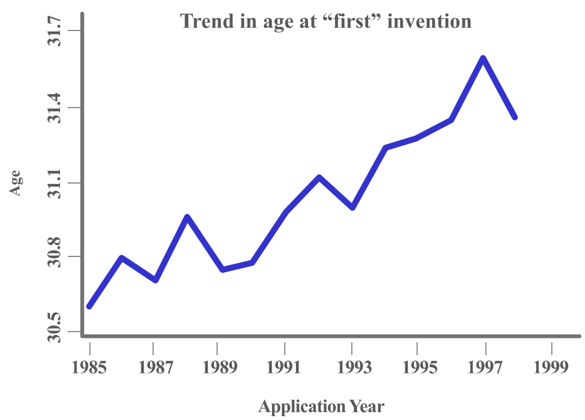
Science, Invention, and the Life-Cycle Peak

- Bernstein et al. (2019): U.S. patent data, virtually all U.S. inventors



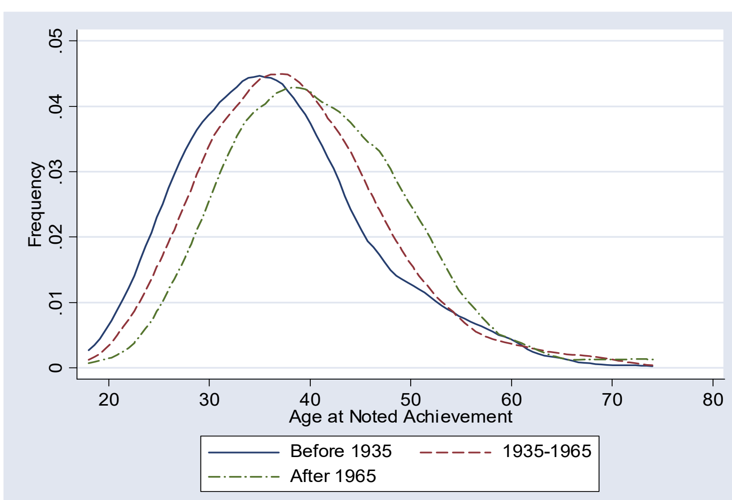
- Jones, Reedy, Weinberg (2015): Review literature on scientists. *Middle age peak is a universal finding*.

But Dynamic in Age



- ❖ Age at first patent is going up (Source: Jones 2009)
- ❖ Return to cumulativeness in understanding life-cycle creativity

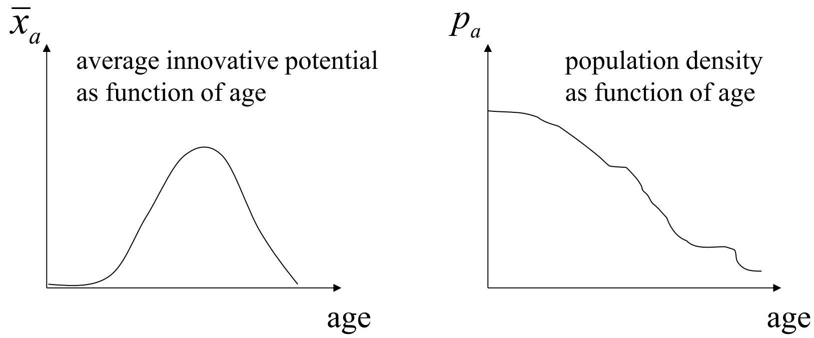
The Shifting Life Cycle Peak



Data: (1) Nobel Prize winners in Physics, Chemistry, Medicine, and Economics; (2) Great technological achievements over 20th Century. (Jones "Age and Great Invention" 2010)

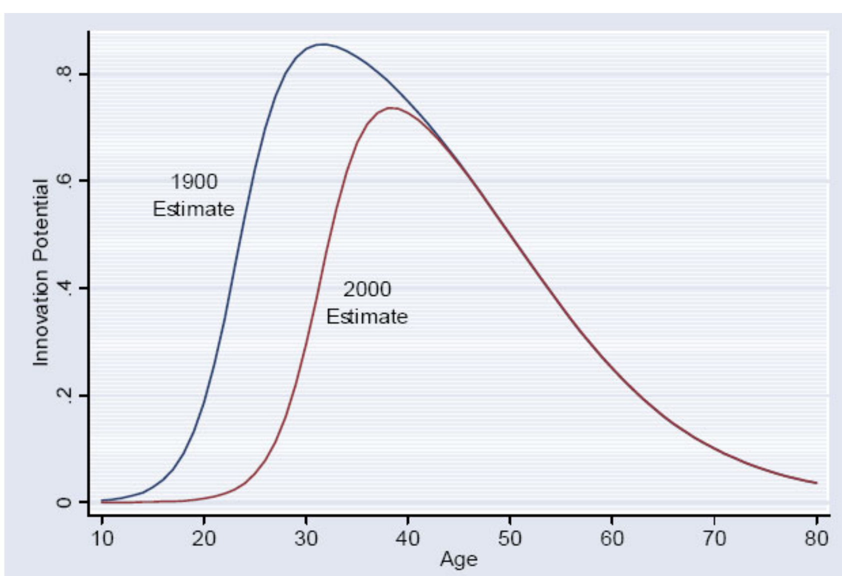
The Shifting Age Distribution of Great Invention

- Why this aging pattern?
  - Hypothesis #1: Shift in life cycle productivity
  - Hypothesis #2: Aging population

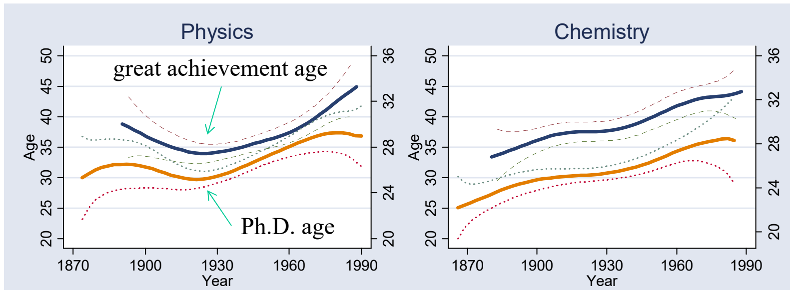


- If there is a shift in life-cycle productivity? If so, does it come early in life-cycle, late in life-cycle, or both?

Age: Estimated Shift in Innovation Potential



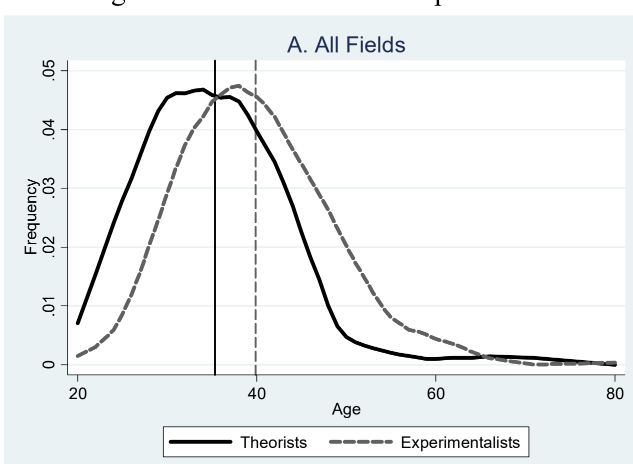
The Physics Experiment



- Early 20th century physics experienced the quantum mechanics revolution, a broad shift in foundational knowledge
- The age at Ph.D. and great achievement in physics, and only in physics, fell during that time

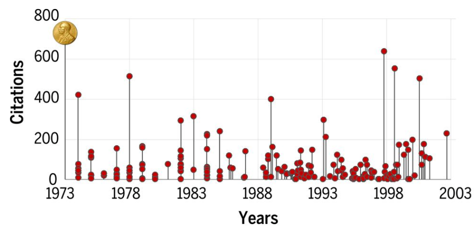
Age and Theoretical vs. Empirical Contributions

- Galenson and Weinberg have emphasized the distinction between "conceptual" and "experimental" reasoning, where the former favors the young (e.g., think mathematicians vs historians)
- Operationalizing this distinction for Nobel prize winners:

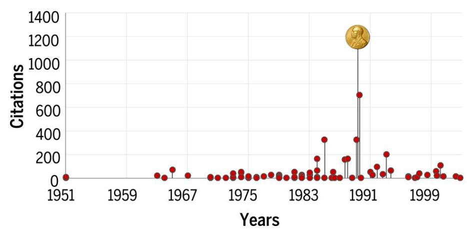


Random Impact Rule & Hot Streaks

- Despite strong tendency toward middle age peak, it appears that your single very best work may appear anywhere in the *sequence* of your work with uniform probability (Sinatra et al. 2016)



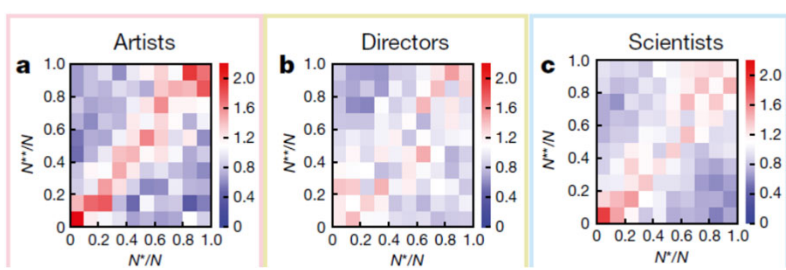
Wilczek
(Physics Nobel)



Fenn
(Chemistry Nobel)

Random Impact Rule & Hot Streaks

- Moreover, there are "hot streaks" where second or third best work come near your best work (Liu et al. 2018)



- Putting literature together:
  - It appears that the quantity of your work bunches in middle age (middle age peak)
  - But the quality of your work peaks randomly in sequence of your work (random impact rule) and tends to bunch up (hot streaks)

The Life Cycle:
Entrepreneurship

Beyond Technical Knowledge: Kline and Rosenberg (1986)

❖ "But technical success (or any purely mechanical measure of performance) is only a necessary and not a sufficient condition in establishing economic usefulness. Indeed, it is obvious from a casual examination of the proceedings in our bankruptcy courts that an excessive or exclusive preoccupation with purely technical measures of performance can be disastrous."

❖ "Successful innovation requires the coupling of the technical and the economic in ways that can be accommodated by the organization while also meeting market needs, and this implies close coupling and cooperation among many activities in the marketing, R&D, and production functions."

Again, a Common View

"Young people are just smarter."
(Zuckerberg)

"The cutoff in investors' heads is 32… after 32, they start to be a little skeptical."
(Graham)

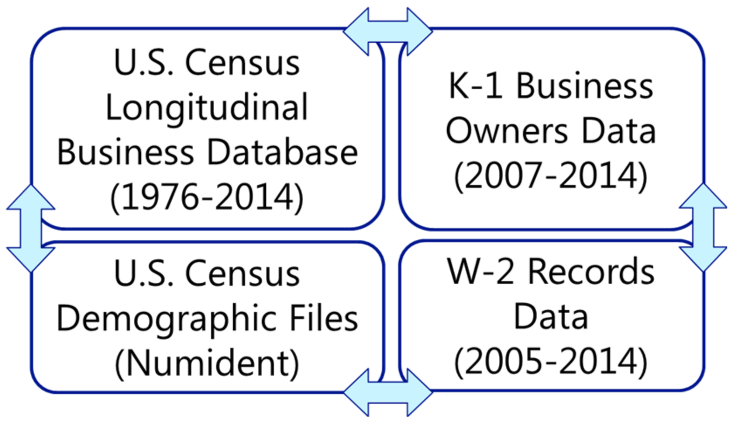
Perception: A Youth Advantage

❖ Consider media focus, and VC focus

	TechCrunch Awards	Inc. and Entrepreneur Magazines	Sequoia	Matrix Partners
Mean	31.0	29.1	33.9	36.5
Median	30	27	33	36
(St. Dev.)	(7.1)	(7.0)	(8.7)	(8.6)
Observations	232	51	415	246
Period	2008-2016	2015	1969-2014	1948-2014
Sectoral Focus (top 5)	Education, Software, Social Media, Consumer Electronics, e-Commerce	Technology, Retail, Media, Consumer Goods, Food Delivery	Semiconductors, Networks, Task Mgmt Apps, Website Compilers, Cloud	Networks, Applications, Commerce, Platform/ Infrastructure, Semiconductors/ Materials

Azoulay, Jones, Kim, Miranda (2020)

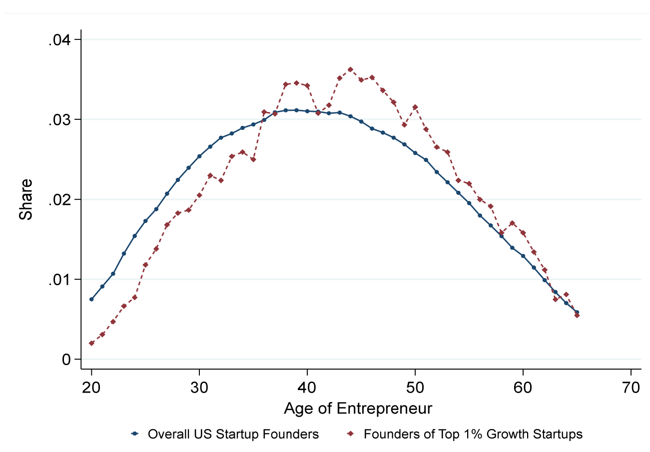
❖ Data: U.S. Administrative Databases



\+ Patents via Longitudinal Linked Patent Business Database
\+ Venture Capital data via VentureXpert & PCRI

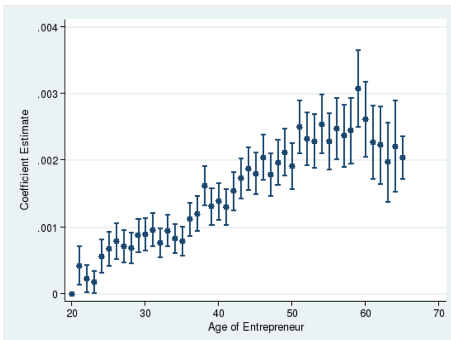
Azoulay, Jones, Kim, Miranda (2020)

❖ Mean age at founding for high-growth firms: 45!

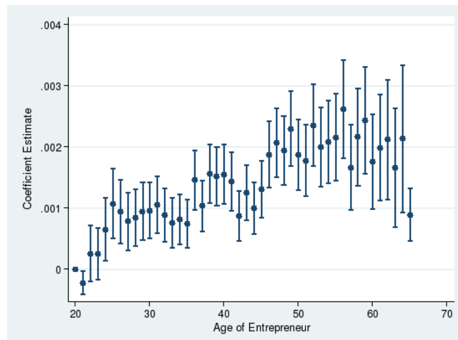


Azoulay, Jones, Kim, Miranda (2020)

- And probability of success, conditional on starting firm, increases with age

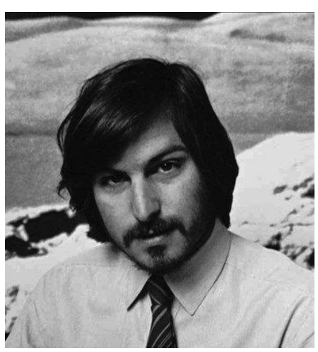


Probability of Successful Exit (IPO or acquisition)

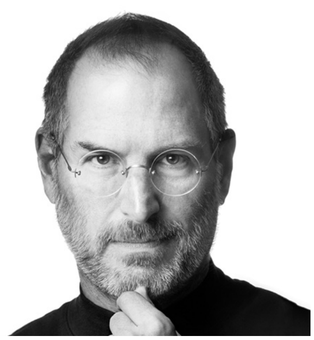


Probability of Top 0.1% Employment at 5 Years

What about Steve Jobs?



(Jobs – Age 23)



(Jobs – Age 51)

Summary: The Life Cycle

- Middle age peak. Common view of youth advantage is wrong. Middle age peak is found quite generally.

- Dynamics. In sciences and patenting, scientific and technological breakthroughs by very young people are increasingly limited. Burden of knowledge may explain.

- Subtleties.
  - Theoretical contributions (somewhat) favor younger people
  - Advantages of youth may increase when burden of knowledge is light
  - "Random impact rule" suggest the quantity of output peaks in middle age, rather than the quality
  - Hot streaks…. Explore vs. exploit pattern?

Human Capital and Innovation

- Growth Viewpoints

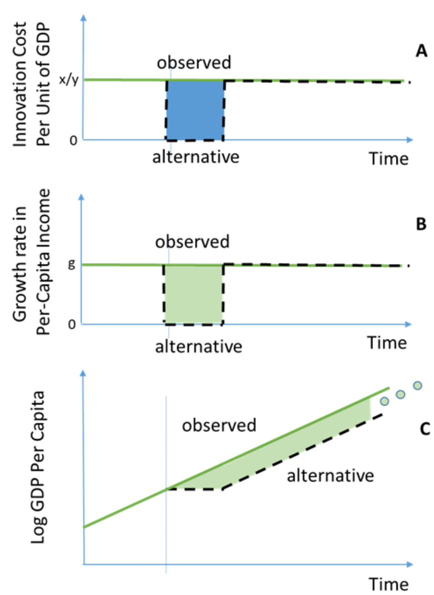
- Micro Viewpoints
  - Collaboration
  - The Life Cycle

- **Labor Supply & the Availability of Talent**
  - Can we accelerate growth?
  - The "Lost Einsteins" question

Labor Supply & the Availability of Talent

Can we greatly increase the rate of innovation?

Recall Social Returns Calculation (Jones and Summers 2021)



Baseline Calculation

The *average* social return (benefit cost ratio) is then

$$\rho = \frac{g/r}{x/y} \approx \mathbf{10}?$$

This calculation suggests that the average social returns to innovation investments are really, really high.

But policy is interested in the marginal return. Would we achieve a high return for *additional* effort at innovation?

The Average vs. the Margin: Growth Model #1

Consider initial class of endogenous growth models (Romer 1990, Aghion and Howitt 1992)

$$g_A = \gamma L_R \qquad (1)$$

Lemma 1: For the knowledge production function, (1), the marginal social rate of return to R&D is

$$\rho_{marginal} = \frac{g/r}{x/y}$$

Here there are no diminishing returns to R&D effort and large intertemporal spillovers. The average and social returns are the same! But this model leads to the "scale effects" problem.

The Average vs. the Margin: Growth Model #2

Consider endogenous growth models where growing effort is need to drive steady-state growth (Jones 1995, Kortum 1997, Jones 2009, Bloom et al. 2020)

$$g_A = \delta A(t)^{\theta-1} L_R(t)^{\sigma} \qquad (2)$$

Lemma 2: For the knowledge production function, (2), the marginal social rate of return to R&D is

$$\rho_{marginal} = \frac{\sigma}{1-(\theta-\sigma)(g/r)} \frac{g/r}{x/y}$$

Here we can have diminishing returns to R&D effort (σ) and various degrees of intertemporal spillovers (θ).

The Average vs. the Margin: Growth Model #2

So, in principle, marginal return to more innovative effort could be quite low. One reason would be that talent is limited, giving steep diminishing marginal returns to increasing $L_R(t)$.

Where does the talent come from? Are there steep diminishing returns?

Channels for more "innovative human capital"

- Immigration
- Domestic creation

Counter view

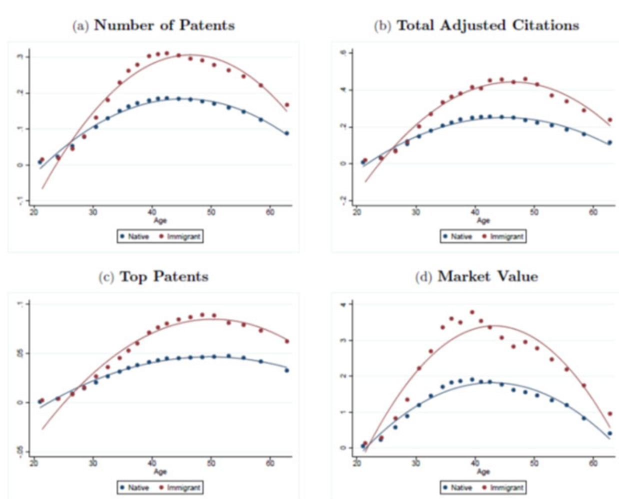
- High end talent essential, and this is fundamentally limited

“Lost Einsteins”?

Are there large, untapped sources of additional innovative talent?

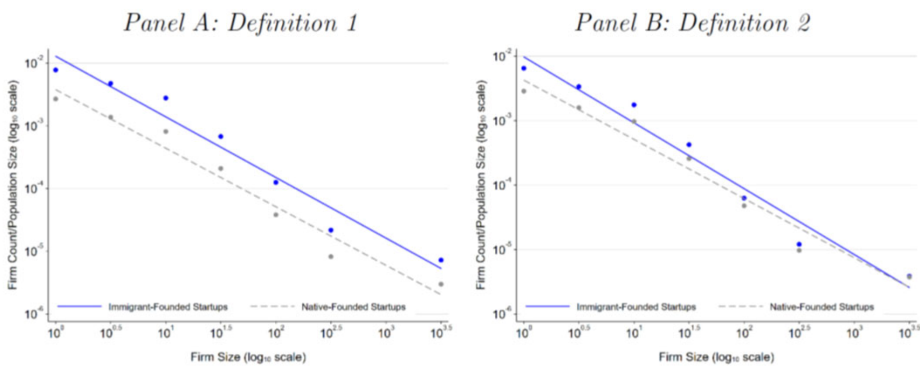
Immigration Channel

- Bernstein et al. (2019): Lots of inventive talent from abroad



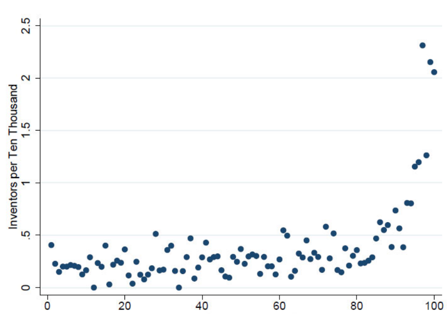
Immigration Channel

❖ Azoulay et al. (2022): Ditto for entrepreneurship



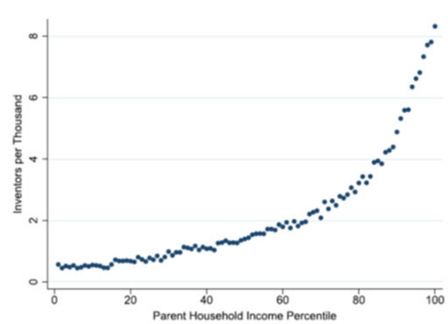
Domestic Talent

❖ Inventors come from high income households in U.S., and always have



Akcigit et al (2016)

early 20th century

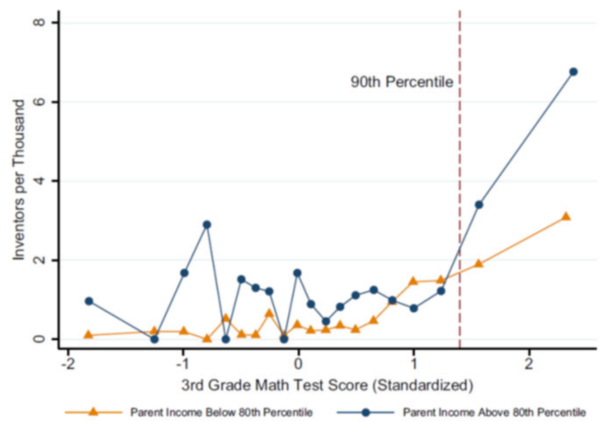


Bell et al (2019)

early 21st century

Lost Talent? (Bell et al. 2019)

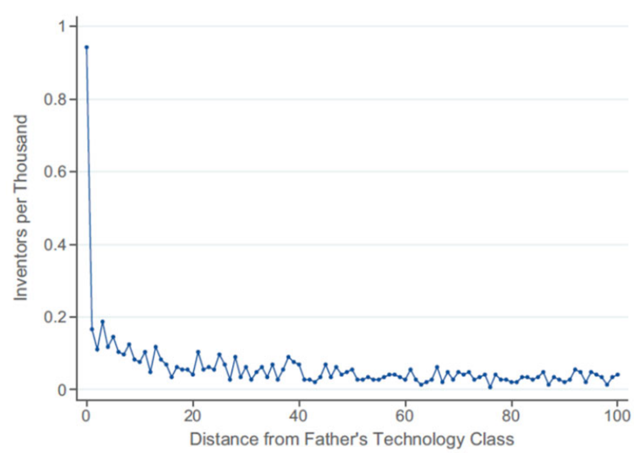
❖ Complementarity between early math skills and household income



❖ Implication: lots of kids with equivalently strong math skills don't proceed to inventive careers

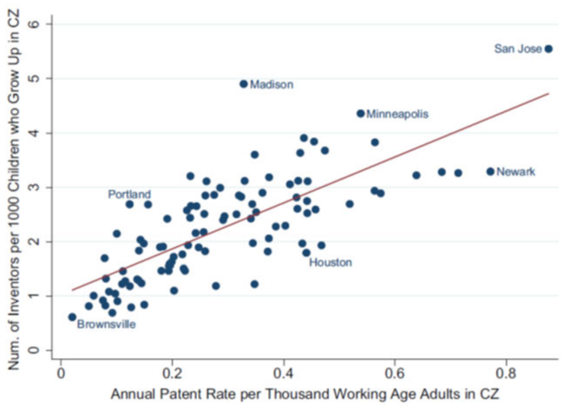
The Exposure Hypothesis

❖ Children tend to patent in exactly the same extremely narrow technology class as their father



The Exposure Hypothesis

- ❖ You are far more likely to become an inventor if that is common in the geographic area where you grow up



- ❖ Also works for movers. The greater the share of childhood spent in a high invention area, the higher the likelihood you become an inventor

Summary: Labor Supply & the Availability of Talent

- ➢ Childhood context appears to have huge implications for whether people enter innovative careers
- ➢ The "people part" of innovation appears highly constrained by contextual factors, including career exposure. Immigration policy is also constraining. Looking globally, national institutional and cultural features are plausibly highly limiting.
- ➢ The logic of growth models suggests there may be substantial room on the margin to expand innovative labor supply and achieve high returns, and more rapid growth
- ➢ Even if stars mainly matter (which is debatable), it seems like there may be very many "lost Einsteins"

END