# The Lasting Effects of Government R&D

Jason Hall

Minnesota

Rice University
November 5, 2026

*Any views expressed are those of the author and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 3126. (CBDRB-FY26-P3126-R13126)*

# Motivation & Research Questions

- Federal government devotes about quarter of a percent of GDP to R&D at universities.
  1. Pays for ideas
  2. Pays for new PhDs

- Research Questions:
  1. What role do PhDs play in U.S. innovation?
  2. How does cutting federal funding for universities affect growth and welfare through this human capital channel?

# This Paper

- Measure PhD contributions to innovation: New data
  - Near census of U.S. PhDs and their patents
  - Observed patent output reflects both ability and selection into research
- New model of PhD training and growth
  - Students select into graduate school based on a public signal about their ability
  - Heterogeneous graduate programs admit students based on same public signal
  - Students learn their ability and select to continue as researchers or exit to production
- Calibrate the model to match new empirical evidence
  - Use model to evaluate proposed cuts to federal funding for university R&D
  - Use model to compare efficacy of R&D tax credits and university support

# Preview of Results

- Empirical:
  1. 44% of patents have at least one PhD inventor.
  2. A 1% increase in federal funding relative to GDP increases PhD yield by $\approx 0.55\%$ after 6 years.
  3. PhD institution conveys very little information about ex-post patenting.

- Quantitative:
  1. 3-year $\approx 40\%$ reduction in support reduces welfare 0.07%.
  2. 23-year $\approx 40\%$ reduction in support reduces welfare $\approx 1\%$.
  3. Moving money from R&D tax credits to university support raises welfare by 0.28%.

# Related Literature

- **Innovation, growth, and the allocation of talent**
  *Romer, 1990; Aghion and Howitt, 1992; Murphy et al., 1991; Aghion et al., 2017; Bell et al., 2019; Hsieh et al., 2019; Pearce, 2020; Akcigit et al., 2025*
  ↪ New patent dataset; model of endogenous PhD capacity → future talent distribution.

- **Education as experimentation and public learning**
  *Manski, 1989; Terviö, 2009; Chade et al., 2014; Hendricks and Leukhina, 2018; Wu, 2026*
  ↪ Infer ex ante uncertainty about ability → quantify the value of experimentation.

- **Returns to government R&D**
  *Azoulay et al., 2019; Myers and Lanahan, 2022; Babina et al., 2023; Tham et al., 2024; Dyevre, 2023; Fieldhouse and Mertens, 2025; Campbell et al., 2026*
  ↪ Quantify the growth effects of government R&D through PhD training capacity.

# Empirics

# Federal Funding, Doctoral Human Capital and Growth

$$\frac{\partial \ln(g)}{\partial \ln(f)} \approx \underbrace{\frac{\partial \ln(g)}{\partial \ln(H)}}_{\text{New Ideas}} \times \underbrace{\frac{\partial \ln(N)}{\partial \ln(f)}}_{\text{Quantity of PhDs Changes}} \times \underbrace{\frac{\partial \ln(H)}{\partial \ln(N)}}_{\text{Quality Changes}}$$

# Data

Goal: Measure the role of PhDs in U.S. innovation.

# Data

Goal: Measure the role of PhDs in U.S. innovation.
Challenge: Systematic data on U.S. inventor education is limited.

- Restricted use datasets
- Time-limited (typically post-2000)
- Limited institutional coverage.

# Data

Goal: Measure the role of PhDs in U.S. innovation.
Challenge: Systematic data on U.S. inventor education is limited.

- Restricted use datasets
- Time-limited (typically post-2000)
- Limited institutional coverage.

Match two datasets:

1. ProQuest Dissertations and Theses Global Coverage
   - PhD Dissertations from 1950 to the present.
2. PatentsView Coverage
   - Patents granted from January 1976 to December 2025
   - Focus on assigned patents with at least one US-based inventor

Ultimate sample:

- 3.83 million patents, 1.70 million inventors, 1.96 million Ph.D.s

Matching Example

# Fact 1: 44% of Patents have a PhD Inventor

| | Patents | PhD Share (%) |
|---|---|---|
| **All Patents** | 891,000 | 44.1 |
| AI/ML & Autonomy | 32,000 | 52.5 |
| HPC & Semiconductors | 34,000 | 69.9 |
| Quantum Information | 1,100 | 76.1 |
| Biotech & Medical | 125,000 | 55.4 |
| Advanced Materials | 16,000 | 69.8 |

Patents granted 2018–2023. PhD share = fraction with $\geq 1$ matched STEM PhD inventor.

Full table

# Government Funding Effects on New PhDs

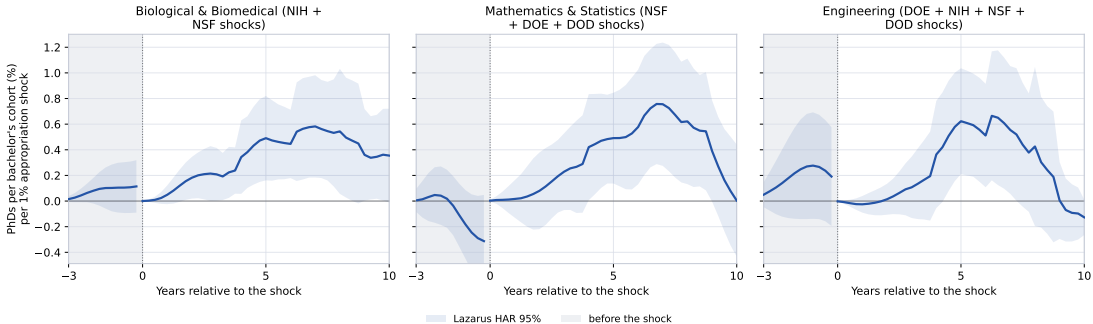
- Want to know how increases in government funding affect supply of PhDs
- OLS is biased because of contemporaneous productivity concerns.

# Government Funding Effects on New PhDs

- Want to know how increases in government funding affect supply of PhDs
- OLS is biased because of contemporaneous productivity concerns.
- Use Fieldhouse and Mertens (2025) narrative R&D account shocks:
  - Map R&D account to fields
  - Scale by university-linked obligation shares
  - 1974–present
- Estimate field-by-field local projections:

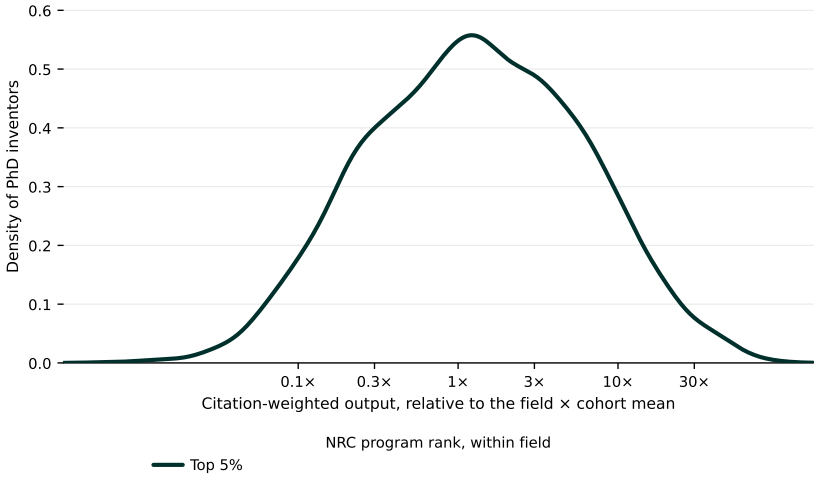
$$\overline{\ln \text{PhD}}_{f,t+h} = \alpha + \beta_h \cdot \text{shock}_{f,t} + \sum_{j=1}^{12} \gamma_j \ln A_{f,t-j} + \sum_{j=1}^{12} \delta_j \ln \text{PhD}_{f,t-j} + \sum_{j=1}^{12} \boldsymbol{\phi}_j \mathbf{X}_{t-j} + \varepsilon_{f,t+h}$$

# Fact 2: 1% Increase in $G/Y \implies \approx 0.55\%$ Increase in PhD Yield



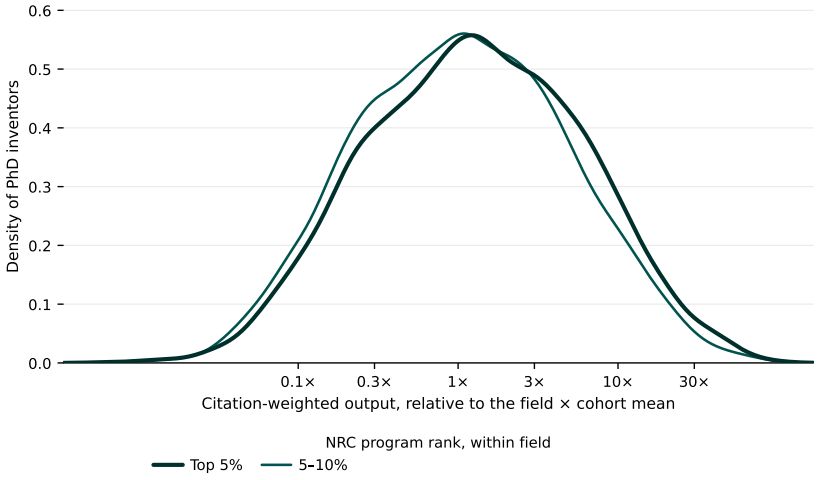

Specification | Funding Responses

# Fact 3:Institution Explains Little Variance in 10 Year Patenting



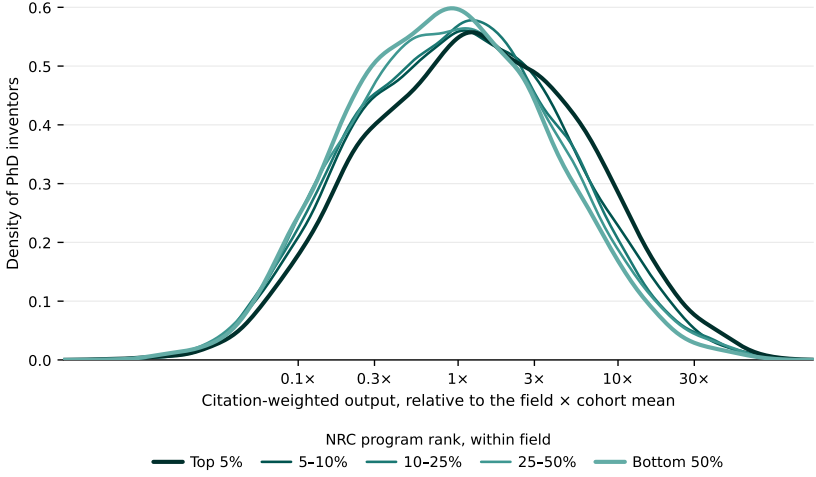

Variance Decomposition | Beats Statistics

# Fact 3:Institution Explains Little Variance in 10 Year Patenting




Variance Decomposition | Beats Statistics

# Fact 3:Institution Explains Little Variance in 10 Year Patenting




Variance Decomposition | Beats Statistics

# Three Facts

1. Individuals with doctorates appear on 44% of patents.
2. A 1% exogenous increase in appropriations to GDP increases PhD yield by $\approx 0.55\%$.
3. PhD institution conveys very little information about ex-post patenting.

# Three Facts

1. Individuals with doctorates appear on 44% of patents.
2. A 1% exogenous increase in appropriations to GDP increases PhD yield by $\approx 0.55\%$.
3. PhD institution conveys very little information about ex-post patenting.

- Model allows me to handle selection into research and patenting.
- Evaluate the effects of cuts in federal funding to R&D over the long-run.

# Model Overview

- Continuous time, infinite horizon, unit mass of agents.
- Workers:
  - Choose to either work in production or attend graduate school.
  - If they attend graduate school, they learn their research ability and choose to either work in research or production.
- Graduate schools:
  - Decide which students to admit.
  - Produce university services using graduate students.
- Research firms:
  - Produce new ideas when matched with researchers.
- Production firms:
  - Produce intermediate goods that are then aggregated.
- Government:
  - Funds graduate schools.
  - Provides a R&D subsidy.
  - Collects taxes to finance these programs.

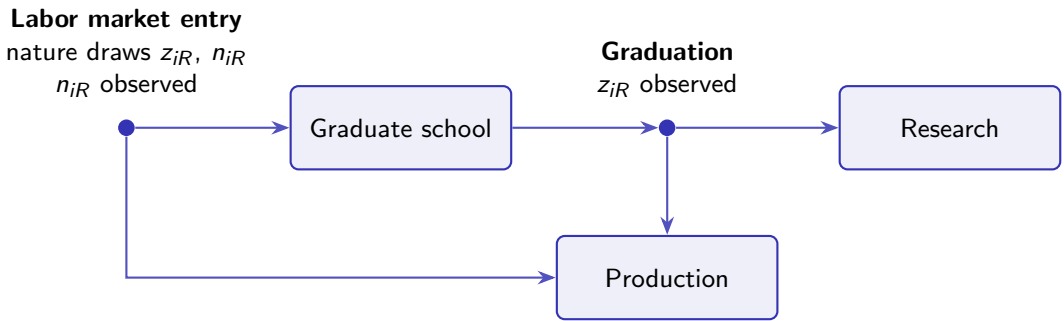
# Preferences and Endowments

- Time is continuous, indexed by $t$.
- Unit mass of agents with preferences over consumption given by:

$$\mathcal{U}(\{c_s\}) = \mathbb{E}_t\left[\int_0^\infty e^{-(\rho+\delta)(s-t)} ln(c_s) ds\right]$$

- $1-\pi$ people are born as production workers.
  - Have an endowment $\xi_1$ of production skill
- $\pi$ people can be researchers.
  - Have one unit of production skill.
  - Draw research skill $z_{iR}$ from a distribution $F_R$.

# Information Structure and Choices




All information is public and common knowledge.

# Technology

- Production technology:

$$y_{jt} = q_{jt}\ell_{jt}$$

- Final goods technology:

$$\ln(Y_t) = \int_0^1 \ln(y_{jt})dj$$

- University technology:

$$B_{jt} = \alpha_j N_{jt}$$

- University admissions cost:

$$K(x_{jt}) = \frac{\kappa x_{jt}^{1+\nu}}{1+\nu}$$

- Research technology:

$$I_t = \varphi_R z_R$$

# Graduate Programs

- Unit mass of universities of $J$ types, with productivity $\alpha_j$.
- Each university chooses how many people to admit: $a_{j,t}(n_R)$.
- Universities maximize ability of admitted students:

$$\int \mathbb{E}\left[z_R \mid n_R\right] a_{j,t}(n_R) dn_R,$$

- They are constrained to break-even on a cohort:

$$K(\int a_{j,t}(n_R) dn_R) = \int \left[A_{\lambda,t}(\alpha_j \mathbb{E}\left[z_R \mid n_R\right] - s_t(n_R)) + \mathcal{F}_t\right] a_{j,t}(n_R)\, dn_R$$

- Can borrow from abroad at interest rate $r^*$.

▸ Full Sequence Problem

# The Life of an Agent

1. Enter labor market with $n_{iR}$ and (unobserved) $z_{iR}$.
2. Choose to attend graduate school or enter production.
   - If enter production, have productivity 1.
   - If attend graduate school, recieve $s_t(n_{iR})$ and draw $z_{iR}$.
   - Graduate and learn $z_{iR}$.
   - Decide to continue as researcher, produce research and recieve $\mathcal{P}_t(z_{iR})$ or exit to production with productivity $\xi_2$.
3. Exit to production.
   - If continue as a researcher, match with a research firm and produce ideas.

# The Value of an Idea

- An idea randomly improves a product line by step-size $\lambda$
- Old productivity technology diffuses to the fringe.
- Within a product line, competition is undifferentiated bertrand $\implies$ limit pricing.
- Research firm auctions off idea to highest bidder.
- Value of a product line $V_t$ evolves according to:

$$(r_t + \iota_t)V_t - \dot{V}_t = \pi_t$$

# The Creation of Ideas

- In order to generate ideas, a research firm must match with a researcher.

$$I_t = \varphi_R z_R$$

- Perfectly elastic supply of research firm entrants.
- Flow surplus is divided via Nash bargaining, with researcher weight $\theta$
  - Outside option for researcher is **permanent** exit to production.
  - Outside option for research firm is exit.

# The Creation of Ideas

- In order to generate ideas, a research firm must match with a researcher.

$$I_t = \varphi_R z_R$$

- Perfectly elastic supply of research firm entrants.
- Flow surplus is divided via Nash bargaining, with researcher weight $\theta$
  - Outside option for researcher is **permanent** exit to production.
  - Outside option for research firm is exit.
- Rate of creative destruction is given by:

$$\iota_t = \underbrace{\chi}_{\text{Exogenous idea flow}} + \varphi_R \int_{z \geq z_t^*} z \, dG_R(z)$$

- Rate of aggregate growth is given by:

$$g_t = \ln(\lambda)\, \iota_t$$

Worker HJBs | Researcher HJBs

# Equilibrium Definition

Given the world interest rate $r$ and the policy path $\{f_t, s_{Rt}, \tau_t\}$, a perfect-foresight competitive equilibrium is prices $\{p_{jt}, w_t\}$, the value of a product line $V_t$, the stipend schedule $s_t(n_R)$, researcher payments $\mathcal{P}_t(z_R)$, the dividend $d_t$ and tax rate $\tau_t$; admission measures $a_{jt}(n_R)$; stocks of students, graduates and researchers; and output $Y_t$, growth $g_t$ and consumption $\{c_{it}\}$ such that:

- Workers choose optimally given $(n_R, z_R)$.
- Firms maximize profits.
- Research firms maximize profits and $\mathcal{P}_t(z_R)$ solves the Nash bargain.
- $a_{jt}(n_R)$ solves each program's problem.
- The government's budget balances.
- Stocks are consistent.
- Labor and goods markets clear.

# Calibration

# Critical Parameters

- **Goal**: Estimate effects of changes in federal funding on talent pool.

# Critical Parameters

- **Goal**: Estimate effects of changes in federal funding on talent pool.
- Crucial objects:
  1. Cost of expansion of graduate programs: $\nu$
  2. How much surplus researchers capture: $\theta$
  3. What does the distribution of research ability look like: $G_R$
  4. How much ex-ante uncertainty there is: $\sigma_\varepsilon$

# Critical Parameters

- **Goal**: Estimate effects of changes in federal funding on talent pool.
- Crucial objects:
  1. Cost of expansion of graduate programs: $\nu$
  2. How much surplus researchers capture: $\theta$
  3. What does the distribution of research ability look like: $G_R$
  4. How much ex-ante uncertainty there is: $\sigma_\varepsilon$
- $\ln(z_{iR})$ is distributed exponentially modified Gaussian (EMG) with parameters $(\sigma_R, \tau)$:

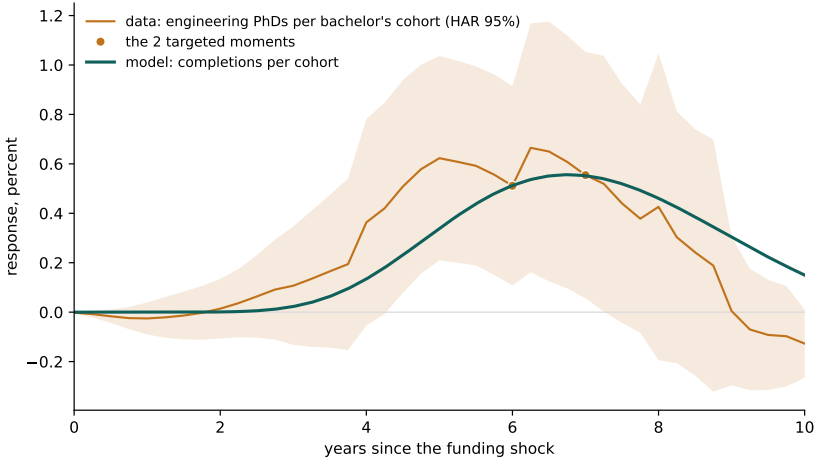
$$\ln(z_{iR}) = \underbrace{X}_{\text{Normal}(0,\sigma_R^2)} + \underbrace{Y}_{\text{Exp}(\tau)}$$

- $\ln(n_{iR})$ is distributed normal with variance $\sigma_\varepsilon^2$:

$$\ln(n_{iR}) = X + \underbrace{\varepsilon}_{\text{Normal}(0,\sigma_\varepsilon^2)}$$

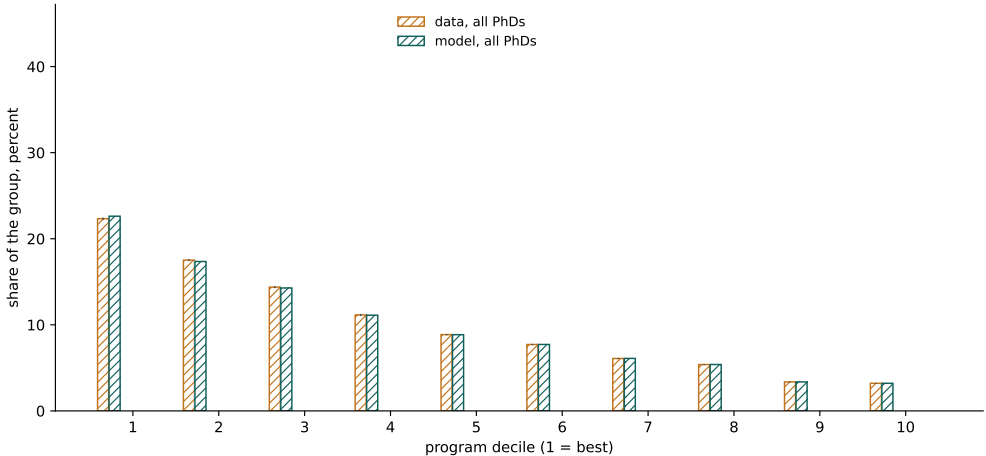
- Patents are produced by a negative binomial process with dispersion $\rho$ and mean $A\, z_{iR}$.

# Identifying the Graduate Cost Elasticity



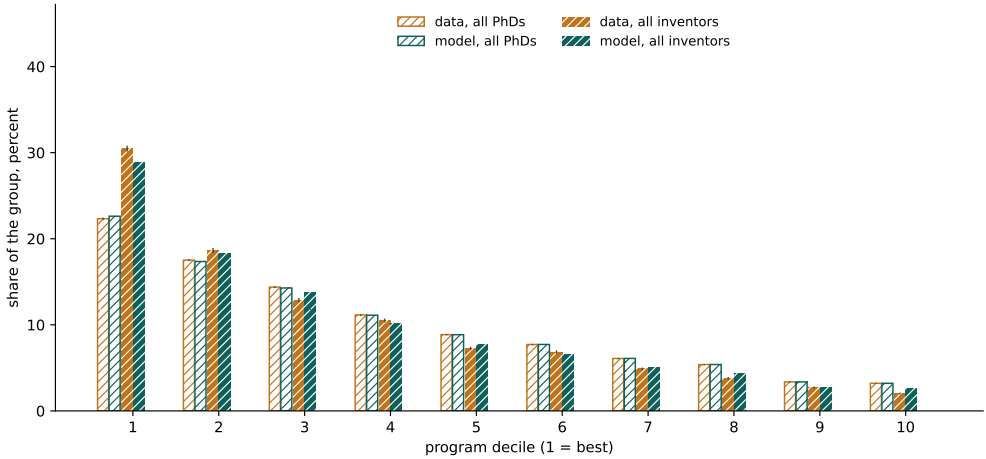

Funding Shock

# Identifying the Tail



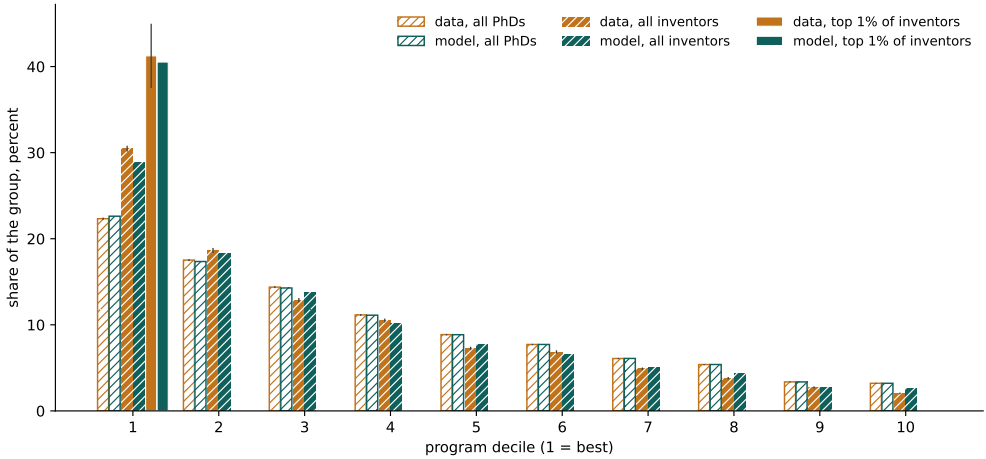

▸ Identifying the Tail ▸ Tail Index by Decile

# Identifying the Tail



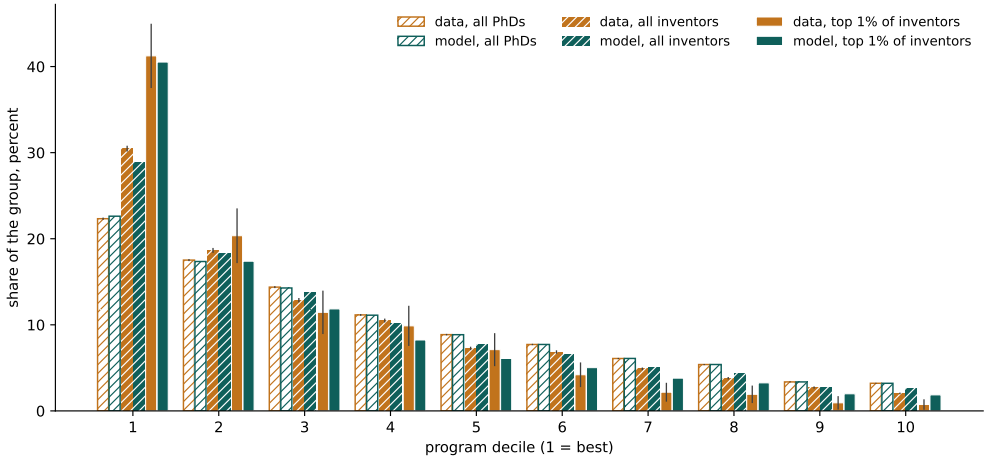

Identifying the Tail | Tail Index by Decile

# Identifying the Tail



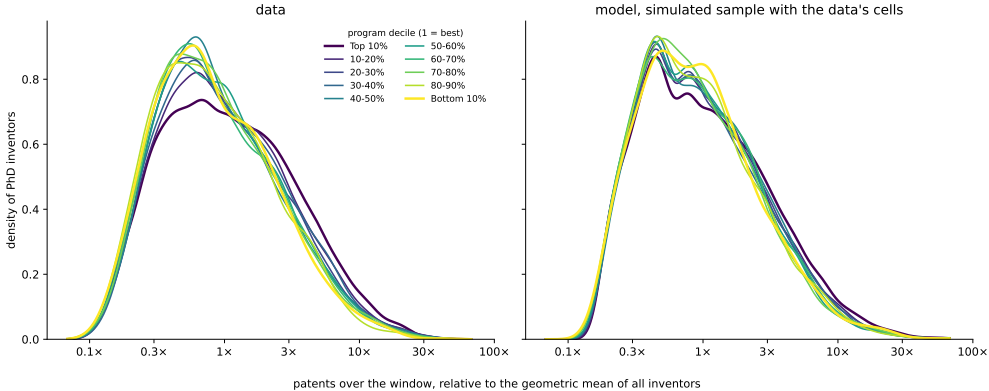

▸ Identifying the Tail ▸ Tail Index by Decile

# Identifying the Tail




▸ Identifying the Tail ▸ Tail Index by Decile

# Identifying Uncertainty





▸ Calibration: Aggregates and Programs ▸ Calibration: Ability and Information

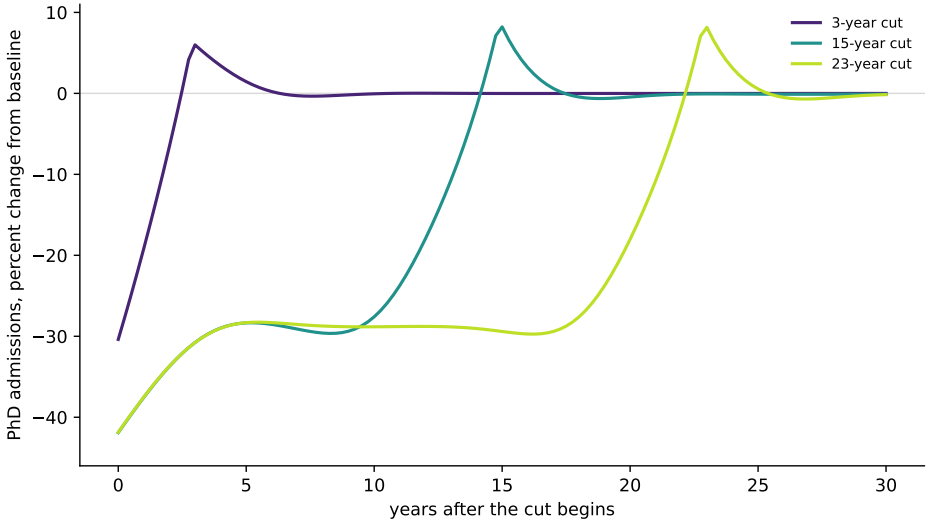
▸ Means ▸ Variances ▸ Cell Rate Adjusted ▸ By Decile ▸ Untargeted Moments

# Results

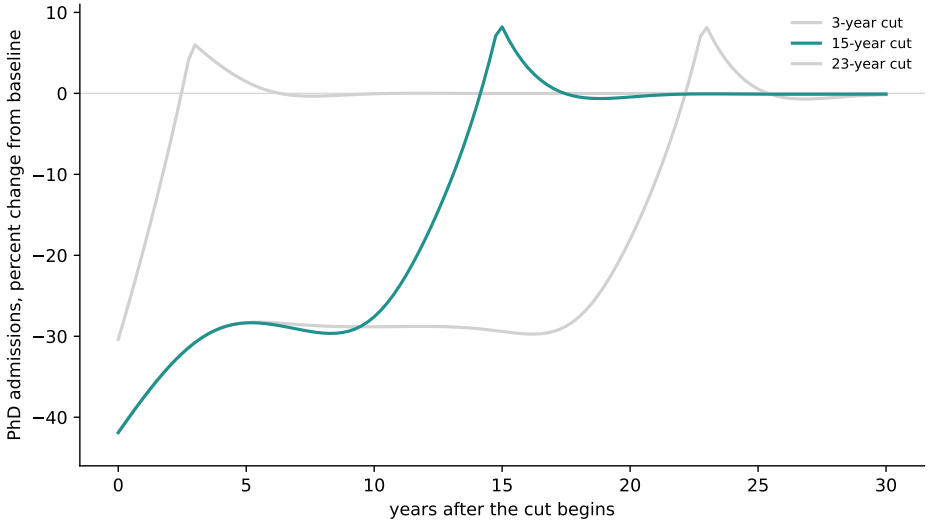
# The Proposed Cuts

- FY2026 President's Budget Request proposed reduction in R&D support for universities of 41.4%.
- Appropriations are very sticky - how long Congress will take to reverse the cut?
  - Potentially very important because of stock-flow dynamics.
- Consider three scenarios:
  1. 3 years
  2. 15 years
  3. 23 years

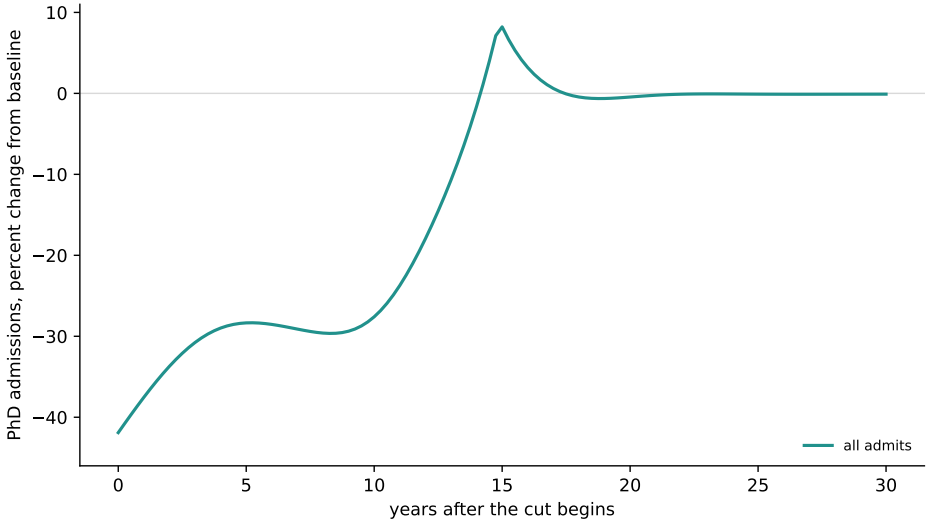
# Admissions Response



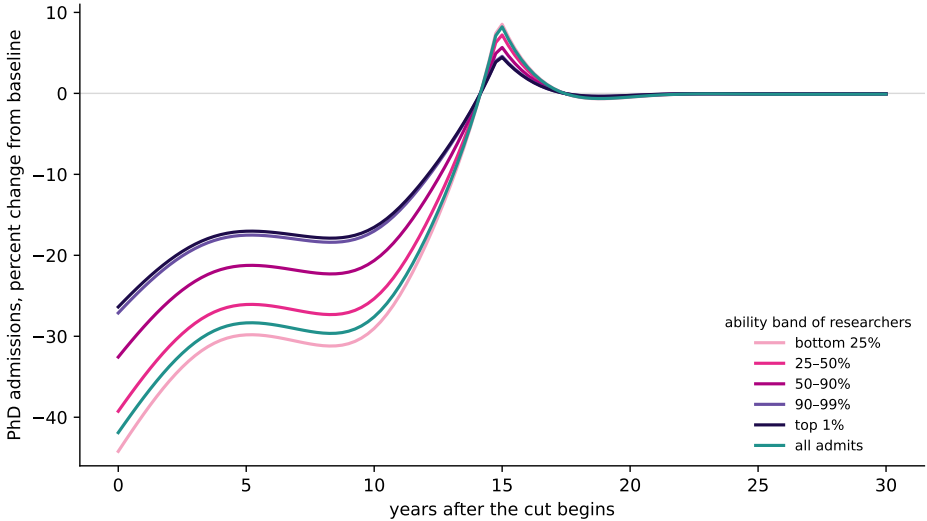
# Admissions Response



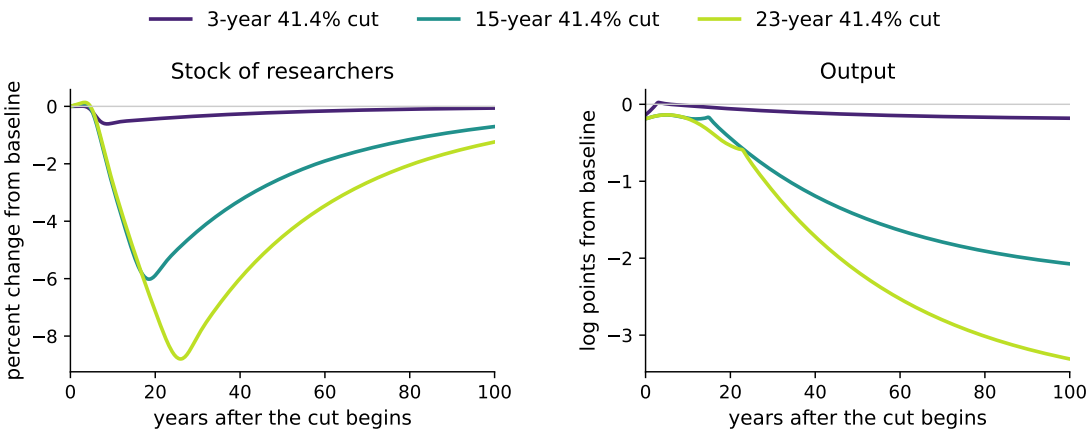
# Lost Ability



# Lost Ability



# Researcher Stocks and Output




Growth Rates

# Subsidizing Innovation

- In the paper: implementing first-best requires both:
  1. Subsidizing research: R&D tax credits (IRC §41)
  2. Subsidizing universities: direct support for R&D (e.g., NSF, NIH, DOE)

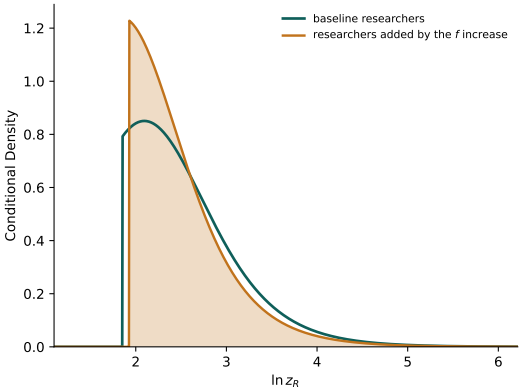
  .
- In practice, governments have limited fiscal space
- Natural question: how should we subsidize innovation given a fixed budget?

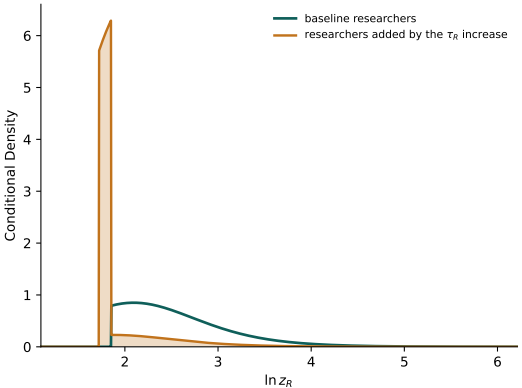
# Subsidizing Innovation

- In the paper: implementing first-best requires both:
  1. Subsidizing research: R&D tax credits (IRC §41)
  2. Subsidizing universities: direct support for R&D (e.g., NSF, NIH, DOE)

  .
- In practice, governments have limited fiscal space
- Natural question: how should we subsidize innovation given a fixed budget?
- Answer: subsidize universities only.
  - 0.28% increase in welfare, 6 basis point increase in long-run growth.
  - BGP comparison is substantially overstated: 2.1% welfare gain.

# Selection Under Different Instruments



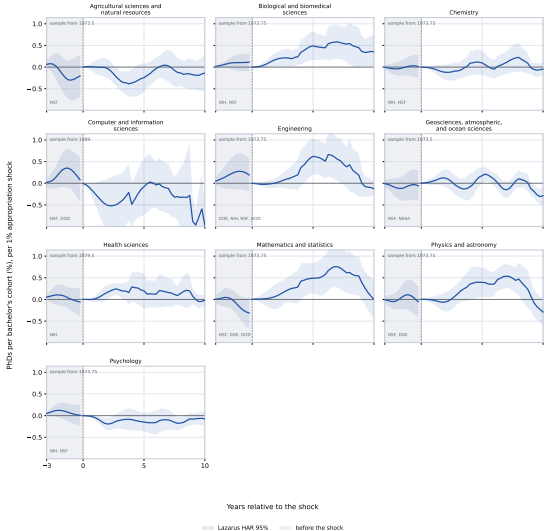



# Conclusion

- Doctoral education is important for economic growth.
- It is hard to identify research talent ex-ante.
- This means the marginal entrant is much closer to the average all equal.
- Short-term cuts reduce welfare, but long-term cuts reduce welfare substantially more.
- To increase growth, subsidize universities, not firms.

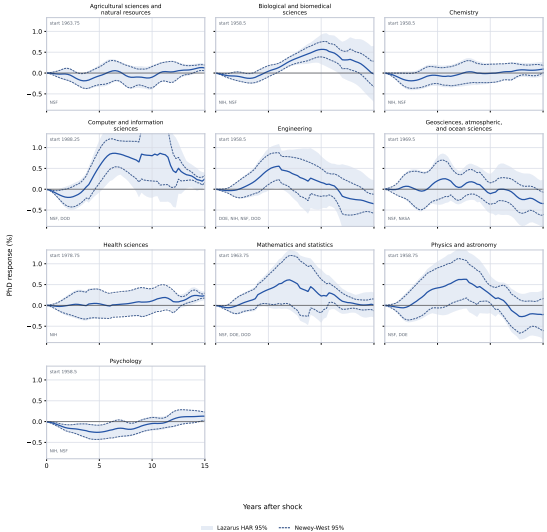
# Appendices

# All Fields — Cyclical Stock Controls



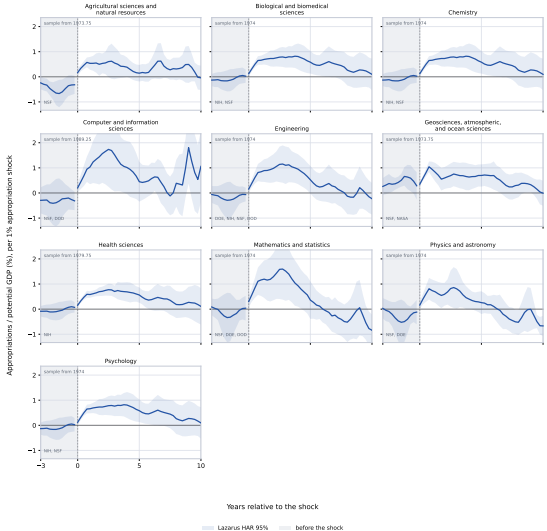

Back

# All Fields — Basic Specification



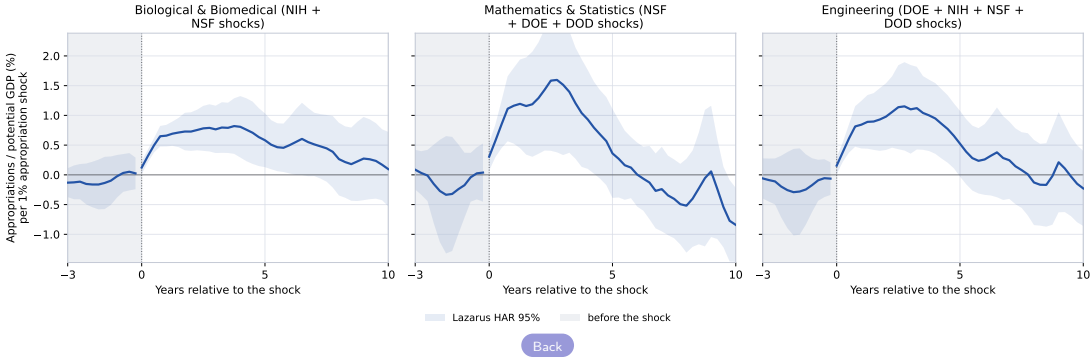

Back

# First Stage: Appropriation Responses

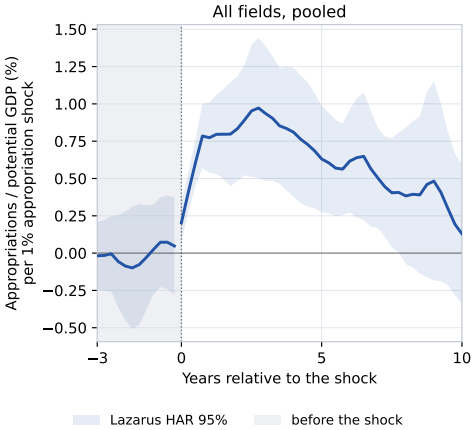


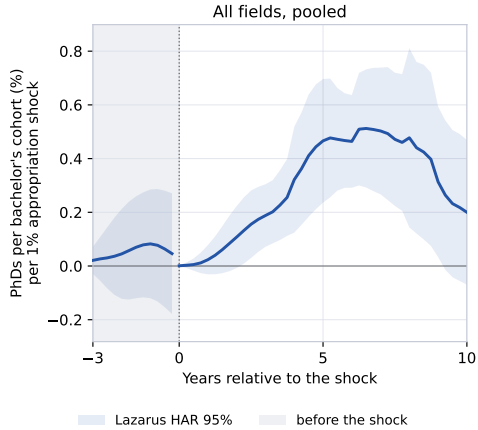

Back

# First Stage: Appropriation Responses



# All Fields Pooled

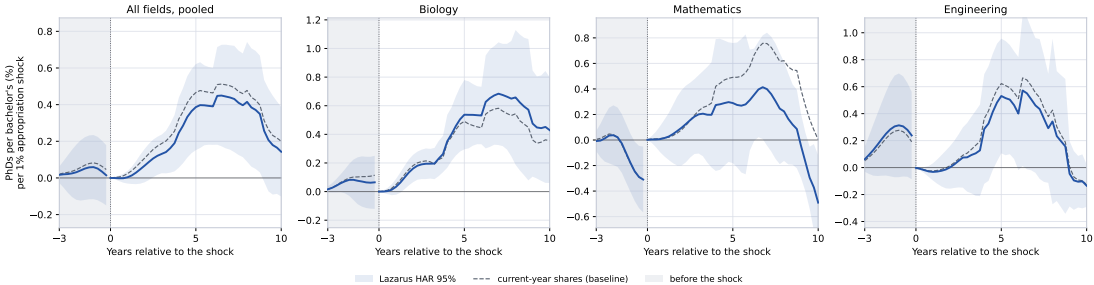



The ten fields stacked on the same sample as the field panels: field fixed effects, one shock coefficient, common lag and control coefficients.

Back

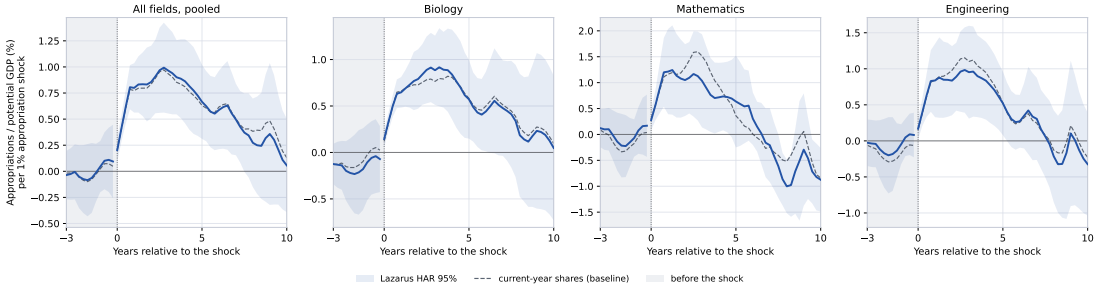
# Predetermined Agency Shares




Agency weights from fiscal years before the shock's (two years earlier); dashed: current-year weights (baseline). Sample from 1975.

Funding Back

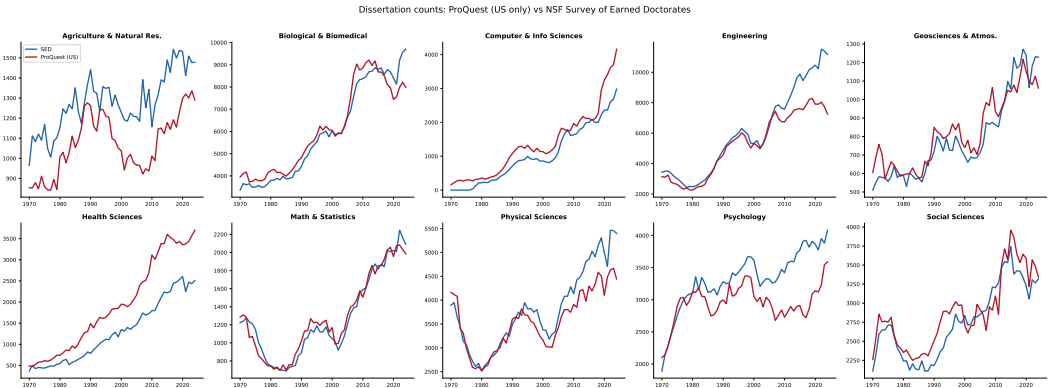
# Predetermined Agency Shares: Funding




Appropriations / potential GDP. Agency weights from fiscal years before the shock's (two years earlier); dashed: current-year weights (baseline). Sample from 1975.

PhDs per Bachelor's Back

# ProQuest Coverage




Back

# PatentsView

- Restrict to patents with $\geq 1$ currently domiciled U.S. inventor.
  - 4,391,716 patents.
- Further restrict to assigned patents.
  - 3,832,919 patents.

Back

# Sample by Decade

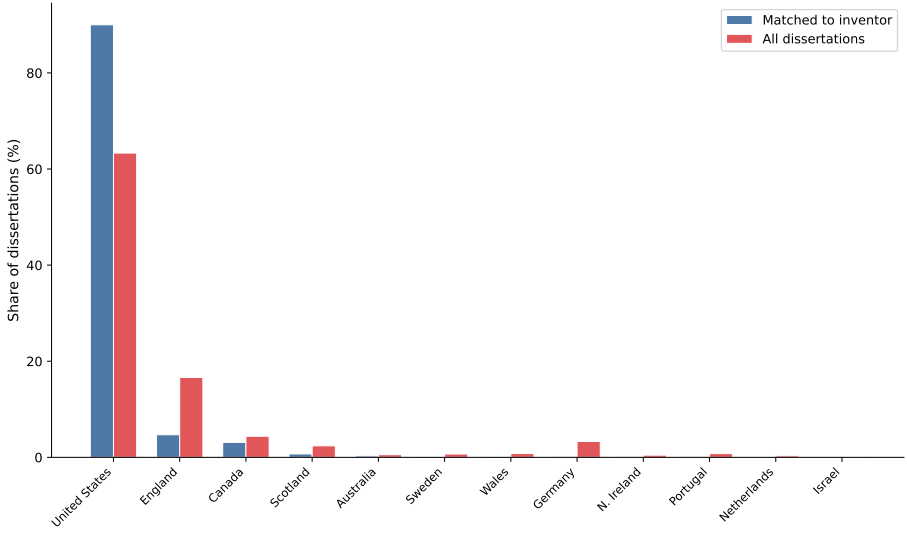
| | Patents | | | Inventors | | |
|---|---|---|---|---|---|---|
| Decade | Total | w/ PhD | % | Total | PhD | % |
| 1970s | 120,904 | 30,650 | 25.4% | 100,088 | 14,037 | 14.0% |
| 1980s | 305,681 | 84,739 | 27.7% | 223,504 | 32,762 | 14.7% |
| 1990s | 485,380 | 168,060 | 34.6% | 379,543 | 62,805 | 16.5% |
| 2000s | 745,716 | 300,225 | 40.3% | 533,813 | 93,236 | 17.5% |
| 2010s | 1,307,905 | 562,084 | 43.0% | 777,951 | 139,420 | 17.9% |
| 2020s | 867,333 | 388,978 | 44.8% | 633,466 | 120,667 | 19.0% |
| Total | 3,832,919 | 1,534,736 | 40.0% | 1,706,516 | 264,996 | 15.5% |

Assigned patents with $\geq 1$ US-located inventor, by grant year.

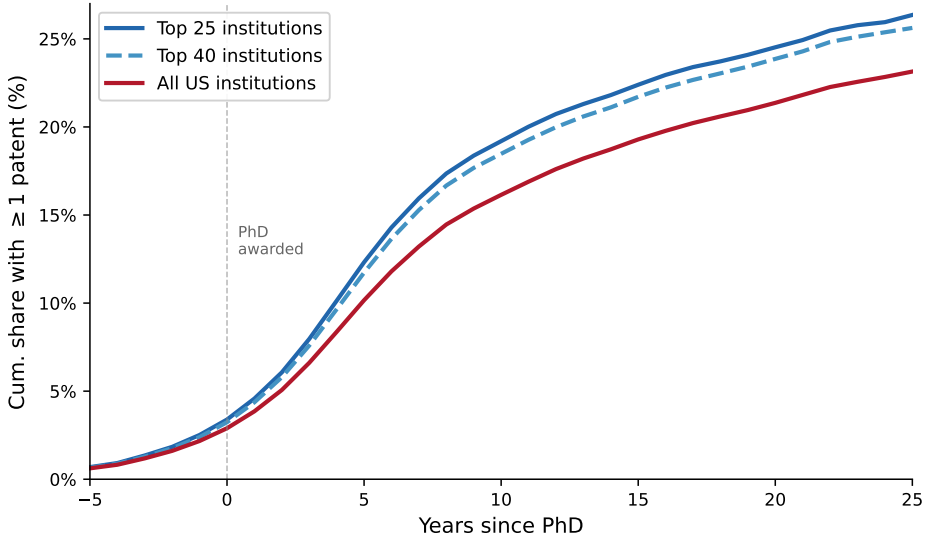
Inventors in a decade are counted once per decade; the total counts each once.

Back

# Dissertations by Country



# Cumulative Patenting Rate of STEM PhDs



# Example CPC Classes

## CPC Classes

A01 Agriculture
A61 Medical / pharma / devices
B33 Additive manufacturing / 3D printing
B60 Vehicles
B81 MEMS / microstructural tech
B82 Nanotechnology
C07 Organic chemistry
C08 Polymers
C12 Biochem / genetic engineering
C22 Metallurgy / alloys
C23 Thin-film coating (CVD/PVD)
C30 Crystal growth
G01 Measurement / sensing
G02 Optics / photonics / displays
G05 Control / automation / autonomous
G06 Computing / AI / image
G11 Information storage
G16 ICT (bio/health informatics)
G21 Nuclear (fission + fusion)
H01 Basic electric elements / semiconductors
H02 Power
H03 Electronic circuitry
H04 Communications
H05 Electronics NEC
H10 Semiconductor devices

## CPC Subclasses

B25J Robotics / manipulators
F03D Wind power
G03F Photolithography
G10L Speech analysis / synthesis

# Matching: An Example

**PatentsView: matched patents**

**Method for determining a border in a complex scene with applications to image masking**

Gregg D. Wilensky; **Scott Cohen**
Filed 2000-05-08 · Granted 2002 · Adobe
G06T7/12 edge segmentation (2 CPC)

**Classifying colors of objects in digital images**

**Scott Cohen** and 3 co-inventors
Filed 2019-07-22 · Granted 2022 · Adobe
G06T7/90 colour characteristics (9 CPC)

**ProQuest dissertations**

**Finding color and shape patterns in images**

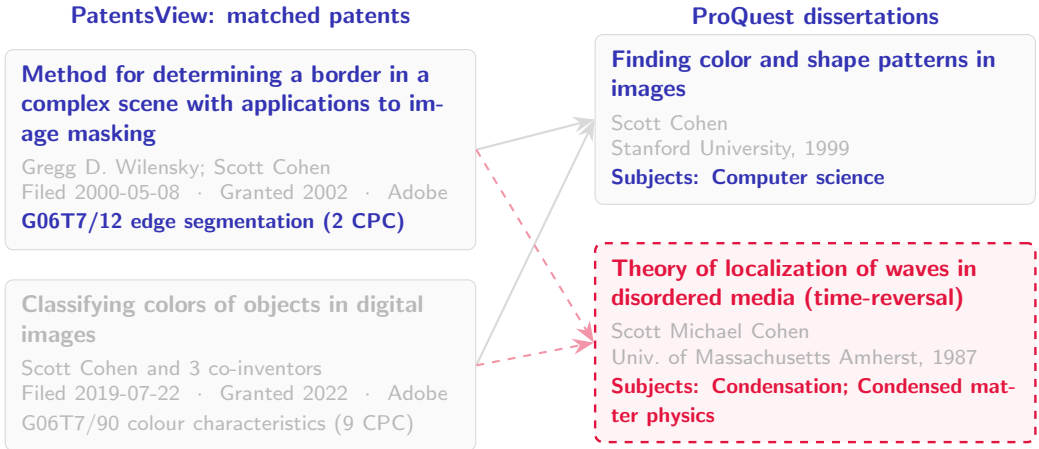
**Scott Cohen**
Stanford University, 1999
Subjects: Computer science

**Theory of localization of waves in disordered media (time-reversal)**

**Scott Michael Cohen**
Univ. of Massachusetts Amherst, 1987
Subjects: Condensation; Condensed matter physics

Algorithm Details · Back

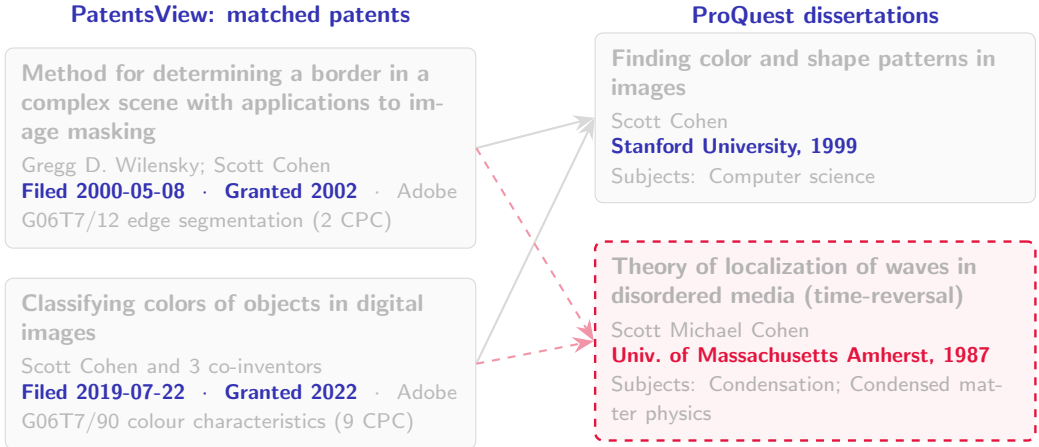
# Matching: An Example




**Text.** SPECTER2 cosine 0.891 mean vs. 0.795 for the alternative.

Algorithm Details · Back

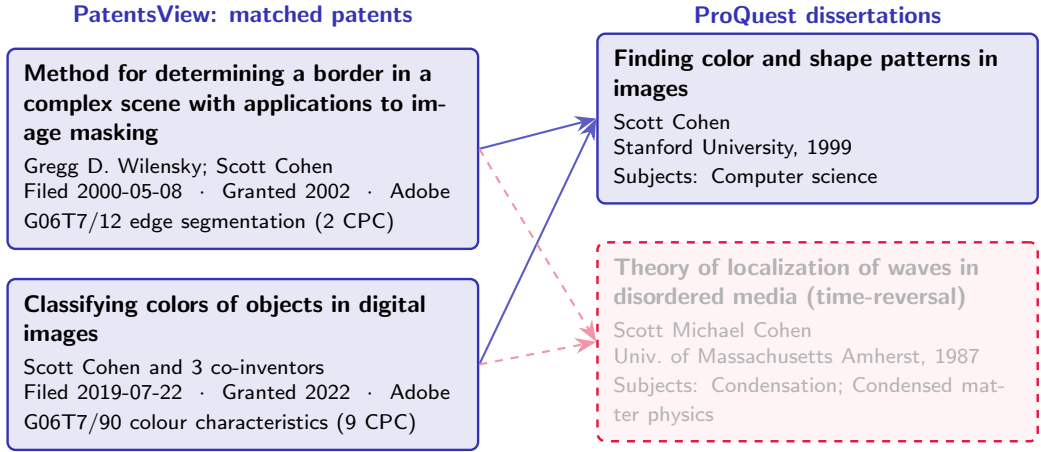
# Matching: An Example




**Timing.** Degree conferred 1999 (alternative: 1987); first filing 2000-05-08.

Algorithm Details · Back

# Matching: An Example




**Corroboration.** 148 patents related to dissertation. Validation

Algorithm Details · Back

# Matching Algorithm

1. Embed (dissertation titles + subjects) and (patent titles + abstracts) using SPECTER
2. Block on inventor surname + given name variants (4 stages, strict $\rightarrow$ loose)
3. Within each block, compute cosine similarity of embeddings
4. Score = text similarity + name quality + temporal proximity + advisor co-invention
5. Greedy 1-to-1 assignment (each inventor $\leftrightarrow$ one dissertation)
6. Filter: mean text similarity $\geq 0.80$, corroboration ratio $\geq 0.80$ or $\geq 8$ corroborating patents

488,775 raw matches $\rightarrow$ 264,996 pass threshold. STEM fields only.

Back

# ORCID Precision

| Stage | TP | FP | N | Precision |
|---|---|---|---|---|
| 1 (name + corroboration) | 311 | 12 | 323 | 96.3% |
| 2 (name only) | 328 | 38 | 366 | 89.6% |
| 3 (advisor) | 6 | 1 | 7 | 85.7% |
| **Total** | **645** | **51** | **696** | **92.7%** |

696 ORCID PhD holders linked to analysis sample via patent + name.
School matching determines TP/FP.

Back

# Constructed Match vs. Administrative Data

| | | Inventor Share (%) | | | Patent Share (%) | | |
|---|---|---|---|---|---|---|---|
| CPC | Description | Admin | ProQuest | Diff | Admin | ProQuest | Diff |
| H10 | Semiconductors | 29.5 | 23.1 | −6.4 | 61.1 | 60.7 | −0.4 |
| H01 | Basic electronics | 17.6 | 14.9 | −2.7 | 38.3 | 36.6 | −1.7 |
| A61 | Medical/pharma | 14.4 | 12.3 | −2.1 | 25.2 | 26.9 | +1.7 |
| G01 | Instruments | 16.7 | 14.7 | −2.0 | 32.2 | 32.8 | +0.6 |
| G06 | Computing | 9.0 | 7.6 | −1.4 | 22.5 | 23.4 | +0.9 |
| H04 | Telecom | 8.0 | 7.6 | −0.4 | 25.9 | 27.6 | +1.7 |
| B60 | Vehicles | 5.7 | 5.4 | −0.3 | 12.5 | 14.9 | +2.4 |
| H02 | Power generation | 10.0 | 9.7 | −0.3 | 22.6 | 23.2 | +0.6 |
| A01 | Agriculture | 9.1 | 9.7 | +0.6 | 20.0 | 24.1 | +4.1 |
| **Mean abs. diff** | | | | **1.8pp** | | | **1.6pp** |
| Rank corr. $\rho$ | | | | **0.98** | | | |

Admin: SED–PatentsView linkage (Goldschlag et al.). US STEM PhDs 2000–2017, patents 2018–2023.

Back

# University Specialization

| CPC Class | University | Share (%) | Patents | Total |
|---|---|---|---|---|
| C08 Polymers | SUNY ESF | 17.8 | 309 | 1,738 |
| | Southern Miss | 23.4 | 1,062 | 4,531 |
| | Akron | 23.7 | 3,034 | 12,816 |
| | *All universities avg.* | *1.7* | | |
| G02 Optics | Kent State | 19.0 | 675 | 3,555 |
| | U. Arizona | 8.4 | 2,807 | 33,244 |
| | UCF | 7.8 | 482 | 6,201 |
| | Rochester | 7.9 | 2,460 | 30,985 |
| | *All universities avg.* | *1.2* | | |
| A01 Agriculture | North Dakota State | 16.5 | 708 | 4,299 |
| | Nebraska | 10.9 | 1,079 | 9,918 |
| | Iowa State | 9.2 | 3,392 | 36,672 |
| | *All universities avg.* | *1.8* | | |
| B60 Vehicles | U. Detroit Mercy | 15.4 | 170 | 1,106 |
| | Wayne State | 5.1 | 785 | 15,425 |
| | Michigan | 2.8 | 2,106 | 75,985 |
| | *All universities avg.* | *1.4* | | |

Share = % of university's matched PhD-inventor patents in that CPC class.

# P(PhD Field | CPC Subclass)

| Subclass | CS/IT | Chem | EE | Bio | Biomed | Phys |
|---|---|---|---|---|---|---|
| G06F Computing | **42** | | 24 | | | 5 |
| C07D Heterocyclic | | **77** | | 9 | 10 | |
| A61K Pharma | | **41** | | 22 | 24 | |
| H04L Digital transm | 29 | | **48** | | | |
| C07K Peptides | | 30 | | **43** | 17 | |
| H04W Wireless | 14 | | **70** | | | |
| G01N Material anal. | | **28** | 7 | 21 | 11 | 12 |
| H01L Semiconductors | | 13 | 29 | | | **35** |
| C12N Genetic eng. | | 22 | | **52** | 10 | |
| A61B Diagnosis | 6 | 8 | 16 | 8 | **26** | 8 |

Each row sums to 100% across all 12 fields (6 largest shown). Bold = modal field. Cells $< 5\%$ omitted.

# Additional Tables

# PhD Share of Critical & Emerging Technology Patents

| | Technology Area | Patents | All PhD | US-Trained | |
|---|---|---|---|---|---|
| 1 | AI/ML & Autonomy | 31,731 | 52.5% | 49.5% | CPC |
| 2 | HPC & Semiconductors | 33,678 | 69.9% | 67.6% | CPC |
| 3 | Quantum Information | 1,119 | 76.1% | 73.2% | CPC |
| 4 | Robotics & Adv. Manufacturing | 15,077 | 42.8% | 40.2% | CPC |
| 5 | Disaster Prevention | 3,808 | 30.7% | 27.2% | CPC |
| 6 | Communications & Immersive | 148,660 | 44.2% | 41.3% | CPC |
| 7 | Biotech & Medical | 124,735 | 55.4% | 52.0% | CPC |
| 8 | Data Storage & Cybersecurity | 63,858 | 33.2% | 30.6% | CPC |
| 9 | Energy & Batteries | 19,179 | 51.8% | 49.2% | CPC |
| 10 | Advanced Materials | 15,740 | 69.8% | 66.7% | CPC |
| | All Critical Tech | 443,339 | 49.3% | 46.3% | |
| | All Patents | 891,018 | 44.1% | 41.1% | |

Technology areas from CHIPS & Science Act §10331. Patents granted 2018–2023, first-listed CPC code only.

# (1) AI/ML & Autonomy — 52.5% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| G06N | ML/Neural Networks | 8,132 | 63.4% | 60.4% |
| G01S | Radio Navigation/Radar | 6,540 | 53.6% | 50.3% |
| G06V | Image Analysis | 6,114 | 51.7% | 48.6% |
| B60W | Vehicle Control Systems | 3,734 | 47.1% | 44.0% |
| G01C | Navigation/Surveying | 3,609 | 41.3% | 38.8% |
| G05D | Control Systems | 3,602 | 43.9% | 41.1% |
| | **Total** | 31,731 | 52.5% | 49.5% |

Summary

# (2) HPC & Semiconductors — 69.9% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| H01L | Semiconductor Devices | 11,078 | 67.5% | 65.0% |
| H10D | Semiconductors (reclassified) | 7,991 | 83.0% | 80.9% |
| G11C | Static Stores/Memories | 6,289 | 49.8% | 47.5% |
| H10F | Photovoltaic Devices | 2,238 | 72.3% | 70.2% |
| H10N | Electric Solid-State Devices | 1,836 | 80.2% | 77.9% |
| H10P | Semiconductor Mfg. Processes | 1,255 | 72.8% | 70.8% |
| H10H | Semiconductor Light Sources | 1,127 | 78.8% | 76.1% |
| H10K | Organic Semiconductors | 944 | 79.9% | 78.7% |
| H10W | Semiconductor Packaging | 920 | 71.8% | 69.2% |
| | **Total** | 33,678 | 69.9% | 67.6% |

Summary

# (3) Quantum Information — 76.1% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| G06N10 | Quantum Computing | 523 | 79.3% | 73.8% |
| H10N60 | Superconducting Devices | 249 | 81.5% | 80.7% |
| H04L9/0852-58 | Quantum Key Distribution | 109 | 49.5% | 49.5% |
| H10D48 | Quantum Effect Devices | 98 | 68.4% | 68.4% |
| B82Y10 | Nano Info Processing | 77 | 87.0% | 84.4% |
| H01F6 | Superconducting Magnets | 40 | 60.0% | 60.0% |
| H10N69 | Superconducting Integrated | 23 | 95.7% | 95.7% |
| | **Total** | 1,119 | 76.1% | 73.2% |

Grey: fewer than 100 patents.

Summary

# (4) Robotics & Adv. Manufacturing — 42.8% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| B29C | Shaping of Plastics | 6,144 | 40.8% | 38.3% |
| G05B | Control/Regulating Systems | 4,096 | 41.6% | 37.9% |
| B25J | Manipulators/Robots | 2,408 | 53.8% | 51.9% |
| B23K | Soldering/Welding | 2,380 | 39.1% | 36.9% |
| B33Y | Additive Manufacturing | 49 | 42.9% | 40.8% |
| | **Total** | 15,077 | 42.8% | 40.2% |

Grey: fewer than 100 patents.

Summary

# (5) Disaster Prevention — 30.7% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| G08B | Signalling/Alarm Systems | 3,542 | 28.9% | 25.6% |
| G01W | Meteorology | 266 | 54.9% | 48.9% |
| | **Total** | 3,808 | 30.7% | 27.2% |

Summary

# (6) Communications & Immersive — 44.2% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| H04L | Digital Transmission | 63,269 | 38.6% | 35.6% |
| H04W | Wireless Networks | 28,363 | 52.3% | 49.1% |
| H04N | Pictorial Communication | 21,835 | 41.0% | 38.6% |
| G06T | Image Processing | 14,889 | 50.1% | 47.6% |
| H04B | Transmission Systems | 11,162 | 62.6% | 59.0% |
| H04M | Telephonic Communication | 4,219 | 26.4% | 23.8% |
| H04R | Loudspeakers/Microphones | 4,185 | 36.6% | 32.1% |
| H04J | Multiplex Communication | 738 | 62.3% | 60.0% |
| | **Total** | 148,660 | 44.2% | 41.3% |

Summary

# (7) Biotech & Medical — 55.4% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| A61B | Diagnosis/Surgery | 38,018 | 43.9% | 40.7% |
| A61K | Pharmaceuticals | 25,990 | 68.5% | 64.8% |
| A61F | Prostheses/Implants | 12,564 | 33.3% | 30.2% |
| A61M | Devices for Body Cavities | 10,701 | 36.9% | 33.3% |
| C12N | Genetic Engineering | 8,308 | 80.5% | 77.2% |
| C07K | Peptides | 7,497 | 80.3% | 76.9% |
| A61N | Electrotherapy/Radiation | 6,491 | 57.2% | 54.0% |
| C12Q | Measuring/Testing (Bio) | 5,424 | 80.0% | 75.2% |
| G16H | Health Informatics | 5,359 | 43.5% | 39.6% |
| C12P | Fermentation Processes | 1,257 | 78.1% | 75.7% |
| G16B | Bioinformatics | 852 | 77.1% | 74.8% |
| C12M | Bioreactors | 848 | 67.3% | 65.4% |
| A61P | Therapeutic Activity | 840 | 84.3% | 81.0% |
| C07H | Sugars/Nucleosides | 586 | 79.2% | 73.4% |
| | **Total** | 124,735 | 55.4% | 52.0% |

Summary

# (8) Data Storage & Cybersecurity — 33.2% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| G06Q | Data Processing/Business | 34,979 | 28.4% | 25.6% |
| G06F21 | Security Arrangements | 11,718 | 33.8% | 30.8% |
| G11C | Static Stores/Memories | 6,289 | 49.8% | 47.5% |
| H04L9 | Cryptography/Security Protocols | 5,042 | 33.7% | 32.1% |
| G11B | Recording/Reproducing | 2,855 | 50.4% | 47.8% |
| G06K | Graphical Data Reading | 2,855 | 34.7% | 32.3% |
| H04K | Secret Communication | 120 | 45.8% | 40.8% |
| | **Total** | 63,858 | 33.2% | 30.6% |

Summary

# (9) Energy & Batteries — 51.8% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| H01M | Batteries/Fuel Cells | 6,259 | 60.7% | 58.1% |
| H02J | Power Distribution | 5,047 | 37.5% | 34.8% |
| H02K | Electric Machines | 2,003 | 39.0% | 36.7% |
| H10N | Electric Solid-State Devices | 1,836 | 80.2% | 77.9% |
| H02S | Solar Electric Generators | 840 | 39.4% | 37.0% |
| H01G | Capacitors | 728 | 59.9% | 58.4% |
| F03D | Wind Motors | 662 | 41.4% | 39.1% |
| G21C | Nuclear Reactors | 552 | 49.8% | 43.3% |
| C10L | Fuels | 473 | 56.9% | 53.7% |
| F24S | Solar Heat Collectors | 237 | 43.5% | 43.0% |
| G21F | Radiation Protection | 177 | 36.2% | 34.5% |
| G21K | Nuclear Techniques | 109 | 67.0% | 63.3% |
| G21G | Nuclear Transmutation | 87 | 57.5% | 56.3% |
| G21B | Fusion Reactors | 80 | 58.8% | 58.8% |
| G21D | Nuclear Power Plant | 54 | 59.3% | 55.6% |
| G21H | Nuclear Energy Sources | 35 | 71.4% | 71.4% |
| | **Total** | 19,179 | 51.8% | 49.2% |

Grey: fewer than 100 patents.

Summary

# (10) Advanced Materials — 69.8% PhD

| CPC | Description | Patents | All PhD | US-Trained |
|---|---|---|---|---|
| C09D | Coating Compositions | 2,117 | 67.5% | 64.0% |
| C23C | Coating Metallic Material | 2,087 | 69.2% | 66.6% |
| C08G | Polycondensation Polymers | 1,839 | 77.7% | 75.8% |
| C08F | Addition Polymers | 1,771 | 81.1% | 77.3% |
| C08L | Polymer Compositions | 1,444 | 67.0% | 62.9% |
| C04B | Lime/Ceramics | 1,237 | 57.5% | 54.6% |
| C01B | Non-Metallic Elements/Compounds | 1,195 | 69.8% | 66.9% |
| C08J | Polymer Processing | 1,123 | 65.4% | 63.2% |
| C08K | Polymer Additives | 637 | 66.4% | 64.1% |
| C22C | Alloys | 564 | 65.2% | 62.4% |
| C30 | Crystal Growth | 384 | 70.8% | 69.3% |
| B81B | Microstructural Devices | 342 | 70.2% | 61.1% |
| B81C | Microstructural Fabrication | 279 | 74.6% | 72.4% |
| C22B | Production/Refining of Metals | 228 | 53.5% | 49.1% |
| C22F | Working Metallic Powder | 212 | 65.1% | 61.3% |
| C01G | Metallic Compounds | 140 | 77.9% | 75.0% |
| B82Y | Nanotechnology Applications | 105 | 86.7% | 83.8% |
| B82B | Nanostructures | 36 | 97.2% | 97.2% |
| | **Total** | 15,740 | 69.8% | 66.7% |

Grey: fewer than 100 patents.

Summary

# Ranking Programs

- Use National Research Council 2010 ranking of graduate programs, aggregated to 11 broad fields.
  - Ranking is survey based and covers $\approx$ 5000 programs
  - Survey was run from 2005 to 2006 (restrict to 1990-2010 cohorts to minimize rank drift)
  - Aggregate by averaging over percentile rank of programs within an institution.
  - Robust to other aggregations and running on singletons.
- This ranking predicts:
  1. The intensive margin Show
  2. The extensive margin Show
  3. The exit margin Show

Back

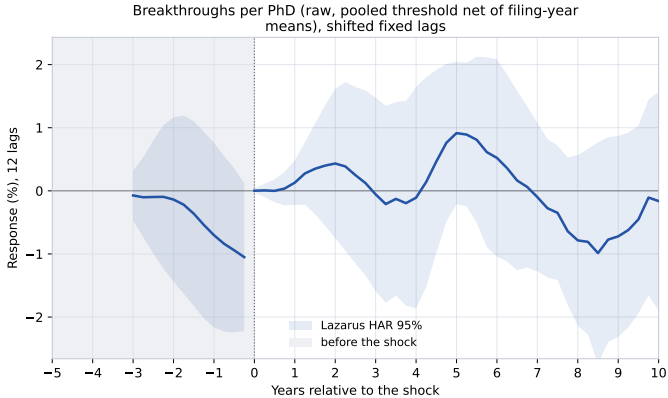
# Validation

**Precision**: ORCID → out of 696 patenting PhDs, 645 correctly matched (92.7%)

**Recall**: Census match of patents to 2000–2017 PhDs → share of patents differ by 1.6pp

ORCID Census Specialist Universities Fields ◂ Back

# Breakthroughs per PhD after a Funding Shock




Breakthrough: top decile of KPST importance among all patents, net of filing-year means.

Quality LPs ◂ Back

# Quality Local Projections

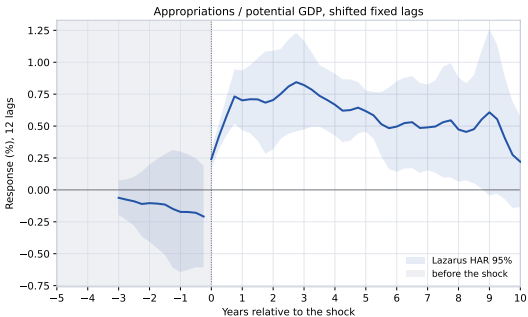
- Patent outcomes need a 10-year window ⇒ only the 1980–2010 cohorts.
- Stack the nine fields: field fixed effects, common coefficients on the shock, lags and controls, 12 lags:

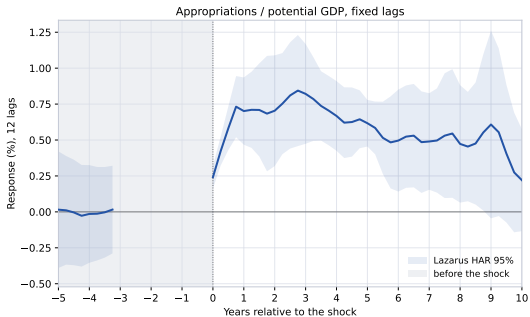
$$\overline{\ln Y}_{f,t+h} = \alpha_f + b_h \varepsilon_{f,t} + \sum_{k=1}^{12} \beta_{h,k} \ln Y_{f,t-k} + \sum_{k=1}^{12} \gamma_{h,k} \ln G_{f,t-k} + \sum_{k=1}^{12} \delta'_{h,k} X_{t-k} + u_{f,t+h}$$

- Patenting boomed in the mid-1990s to early 2000s ⇒ year-adjusted patents per PhD: cohort effects from a Poisson of a cohort's patents on cohort and filing-year effects.
- Outcomes: funding / potential GDP, PhDs per bachelor's, patents per PhD and breakthroughs per PhD (Kelly et al., 2021), each year-adjusted and raw.
- Breakthrough = top decile of KPST importance: ranked within filing year (baseline), net of filing-year means (breakthroughs slide), KPST's released flag, or one threshold on the unadjusted score.
- Breakthroughs counted over each cohort's first 8 filing years (KPST scores cover the latest cohorts only in part).
- Before the shock: controls fixed at $t-27, \ldots, t-16$ for the last 3 years (left) and at $t-12, \ldots, t-1$ (right, de Fraisse et al. (2026)); Lazarus HAR 95% bands.

Breakthroughs | Funding | PhDs per Bachelor's | Patents per PhD (Adjusted) | Patents per PhD (Raw)
Fixed Lags | Breakthroughs (Adjusted) | Breakthroughs (Raw) | No Year Adjustment | KPST Flag (Adjusted) | KPST Flag (Raw)
◂ Back

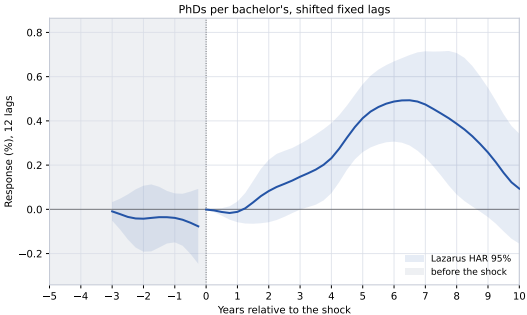
# Quality Sample: Funding

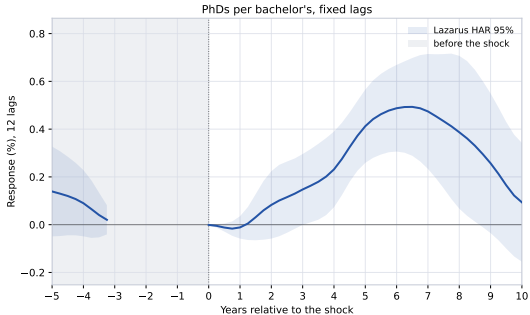



Back

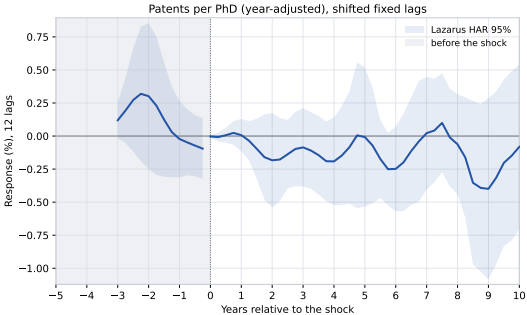
# Quality Sample: PhDs per Bachelor's

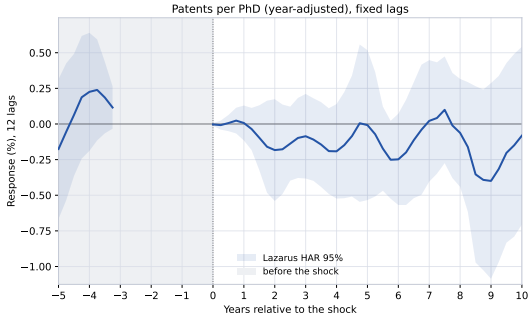



Back

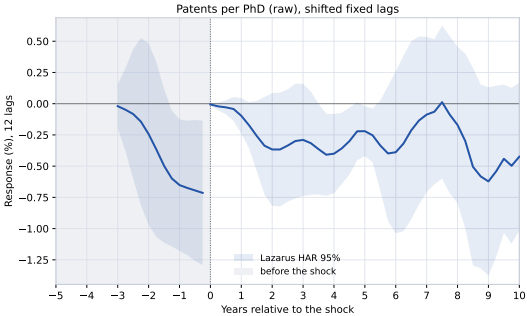
# Patents per PhD (Year-Adjusted)

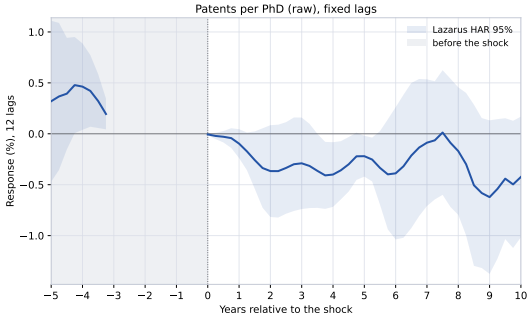



Back

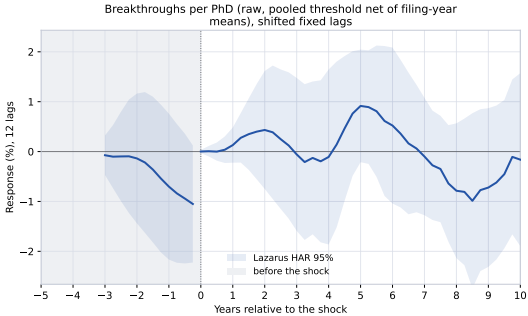
# Patents per PhD (Raw)

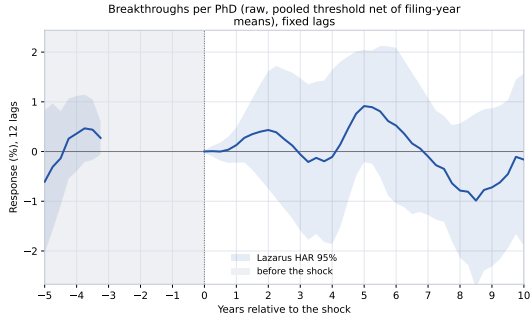



Back

# Breakthroughs per PhD (Raw, Net of Filing-Year Means)

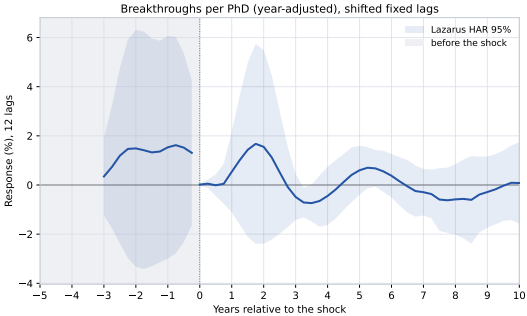


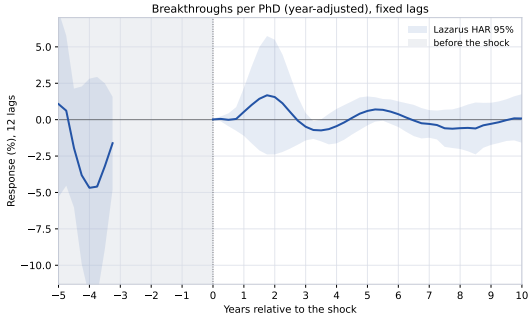




Breakthrough: top decile of KPST importance among all patents, net of filing-year means.

Back

# Breakthroughs per PhD (Year-Adjusted)

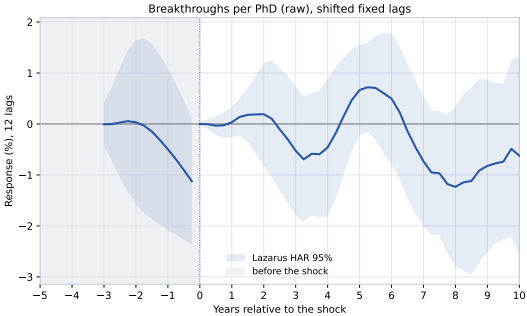


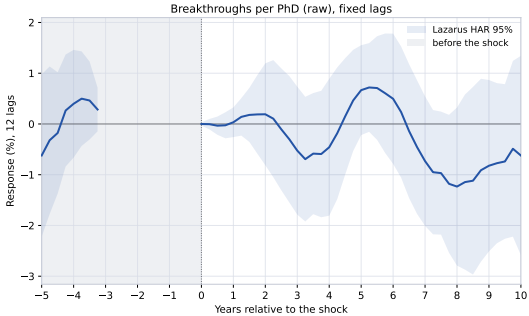




Breakthrough: top decile of KPST importance within filing year.

Back

# Breakthroughs per PhD (Raw)

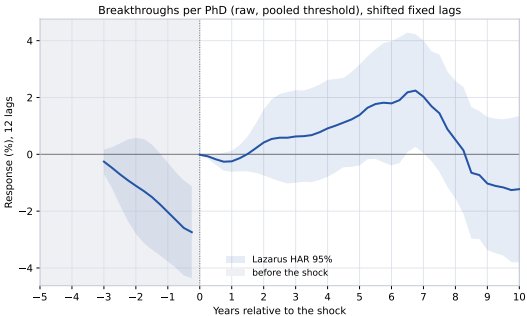


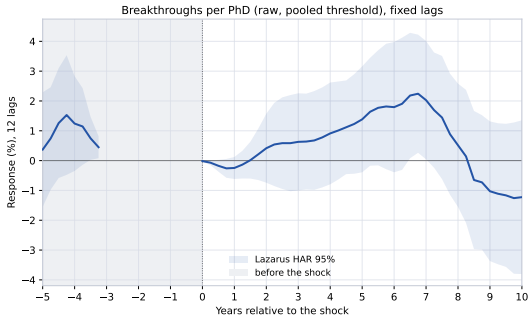




Breakthrough: top decile of KPST importance within filing year.

Back

# Breakthroughs per PhD (Raw, No Year Adjustment)

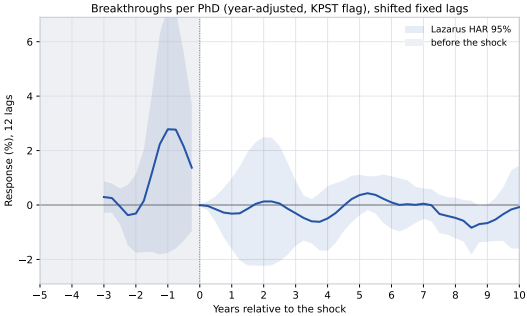


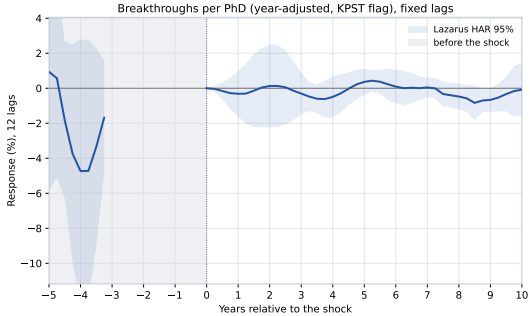




Breakthrough: top decile of KPST importance among all patents, one threshold for every filing year (not ranked within filing year).

Back

# Breakthroughs per PhD (KPST Flag, Year-Adjusted)

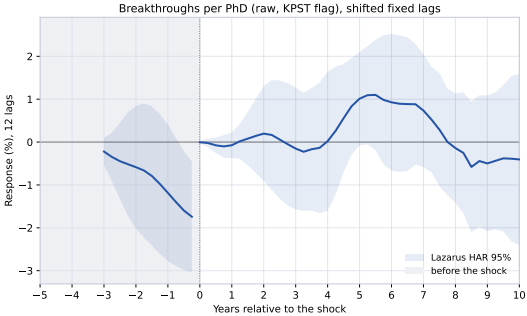


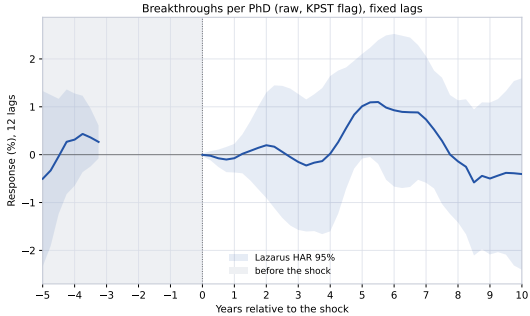




Breakthrough: KPST's released flag (top decile of importance net of issue-year means).

Back

# Breakthroughs per PhD (KPST Flag, Raw)

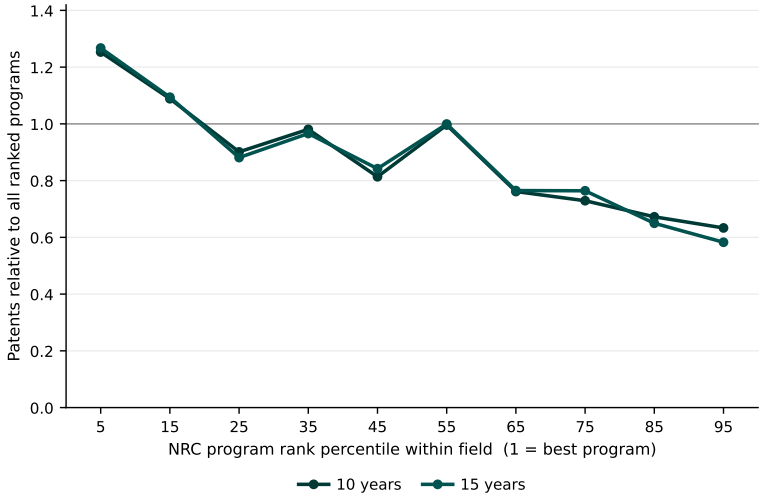


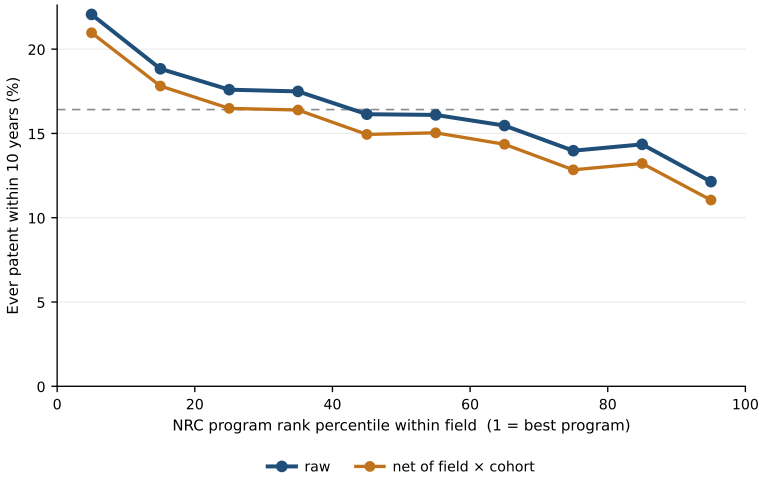

Breakthrough: KPST's released flag (top decile of importance net of issue-year means).

Back

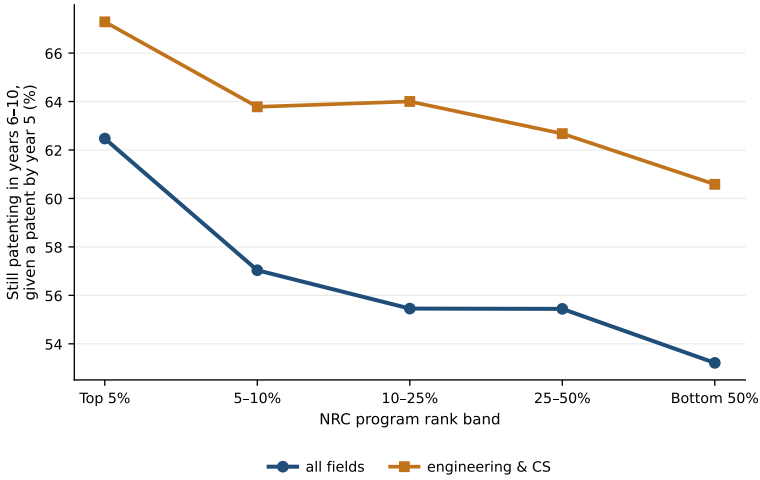
# NRC Rank Predicts the Intensive Margin



# NRC Rank Predicts the Extensive Margin

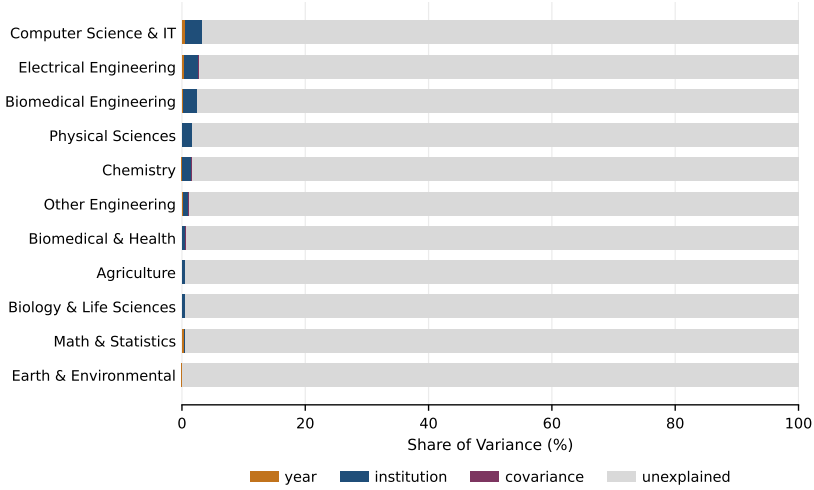


# NRC Rank Predicts Exit from Patenting



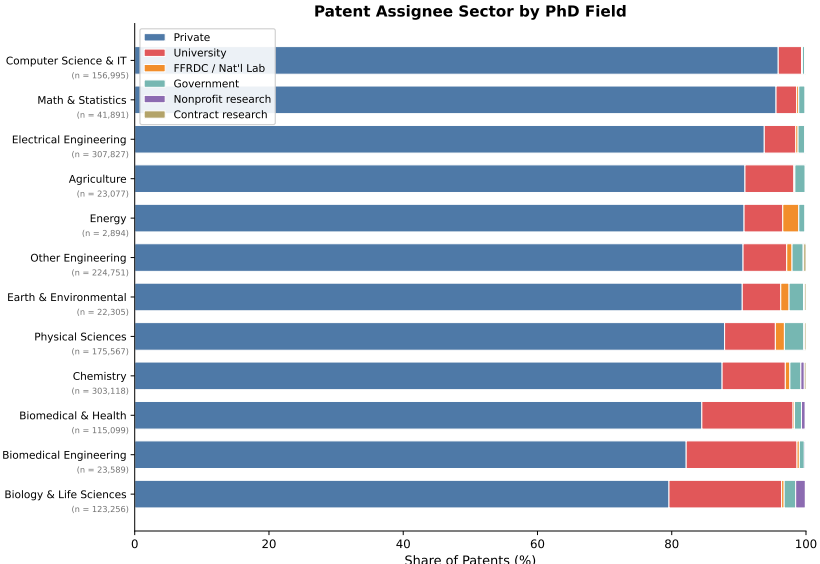
# Empirical Robustness

# Institutions Explain 1.5% of the Variance in 10 Year Patenting

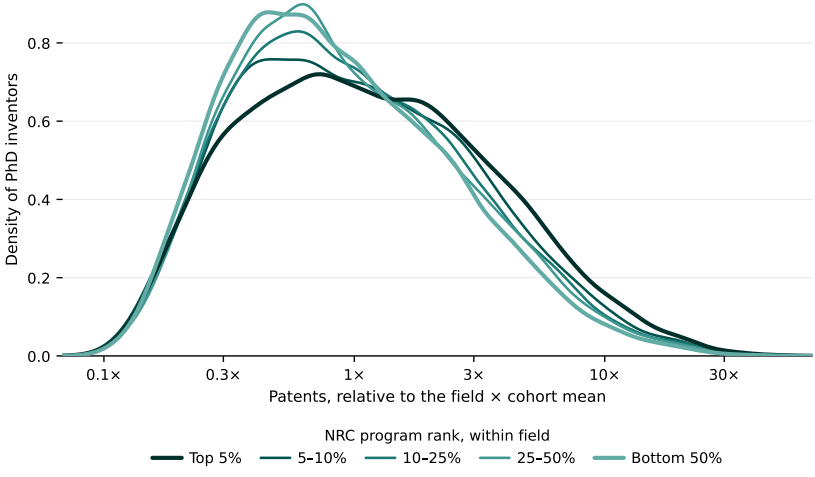



Heteroskedasticity | Capture Curves | Tail Index by Tier | Back

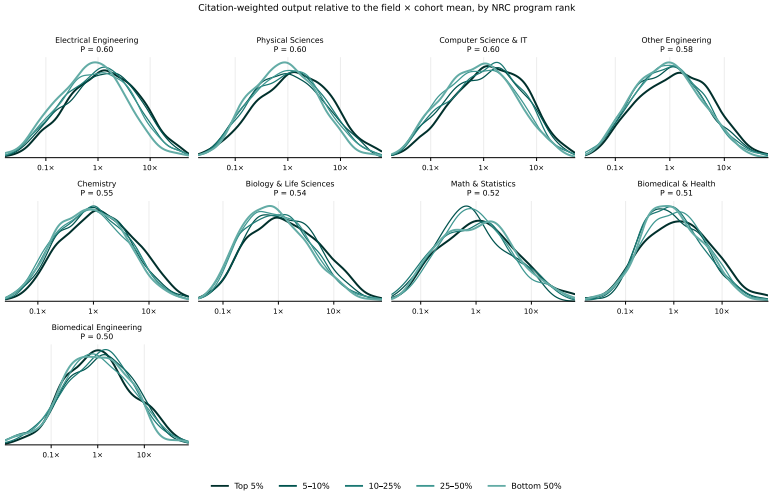
# Most PhD Patents Are Assigned to Private Sector Firms



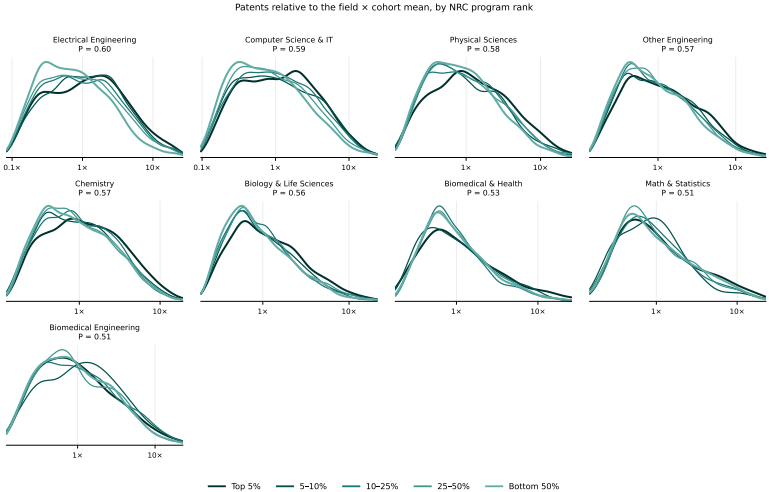
# Raw Patent Density



# Raw Density by Field



# Raw Density by Field



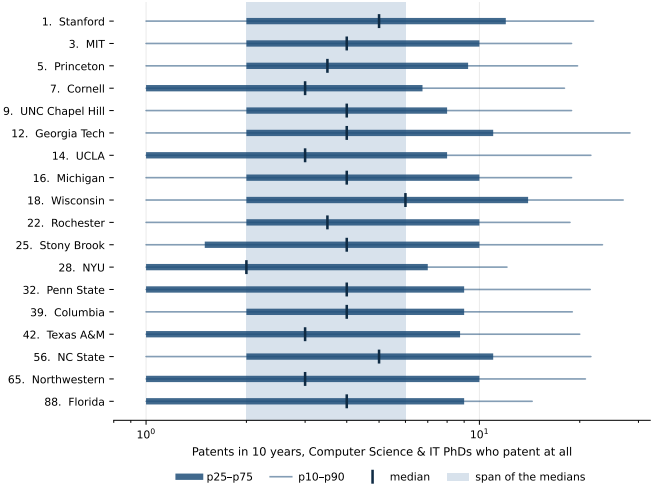
# How Often Does a Bottom-50% Program Beat a Top-5% Program?

| | Inventors | | Mann–Whitney | | | van Elteren | | |
|---|---|---|---|---|---|---|---|---|
| Field | Bottom 50% | Top 5% | P(beats) | 95% CI | $p$ | P(beats) | 95% CI | $p$ |
| All fields | 13,680 | 11,430 | 0.420 | [0.402, 0.437] | $<0.001$ | 0.426 | [0.413, 0.440] | $<0.001$ |
| Biology | 1,752 | 1,128 | 0.458 | [0.426, 0.484] | 0.005 | 0.454 | [0.421, 0.484] | 0.004 |
| Biomed. & health | 792 | 365 | 0.486 | [0.417, 0.544] | 0.661 | 0.467 | [0.392, 0.534] | 0.367 |
| Biomed. eng. | 592 | 323 | 0.497 | [0.443, 0.561] | 0.918 | 0.500 | [0.445, 0.566] | 0.991 |
| Chemistry | 2,924 | 2,326 | 0.447 | [0.424, 0.470] | $<0.001$ | 0.435 | [0.412, 0.460] | $<0.001$ |
| CS & IT | 1,127 | 1,520 | 0.401 | [0.369, 0.431] | $<0.001$ | 0.404 | [0.374, 0.431] | $<0.001$ |
| Elec. eng. | 1,657 | 2,201 | 0.396 | [0.345, 0.448] | $<0.001$ | 0.395 | [0.346, 0.447] | $<0.001$ |
| Math & stat. | 328 | 217 | 0.477 | [0.429, 0.534] | 0.386 | 0.477 | [0.427, 0.528] | 0.364 |
| Other eng. | 1,979 | 1,930 | 0.417 | [0.395, 0.454] | $<0.001$ | 0.422 | [0.401, 0.459] | $<0.001$ |
| Physical sci. | 1,811 | 1,338 | 0.401 | [0.369, 0.431] | $<0.001$ | 0.400 | [0.367, 0.431] | $<0.001$ |

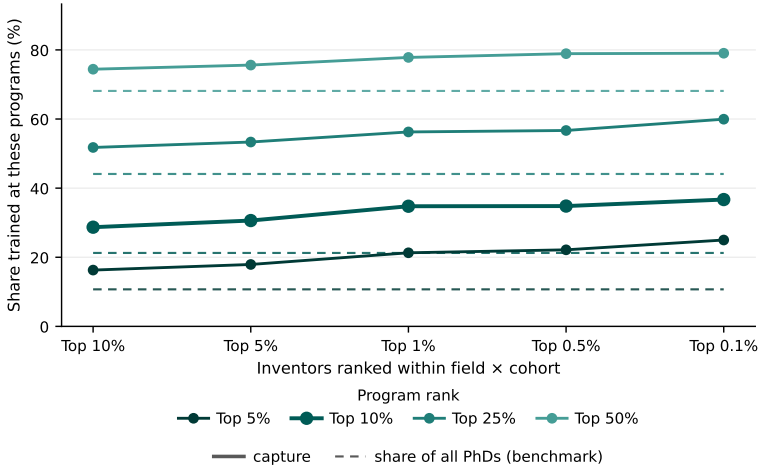
P(beats): the probability that a Ph.D. inventor from a bottom-50% program has more quality-adjusted patents filed within ten years of the degree than one from a top-5% program, output relative to the field $\times$ degree-year geometric mean, ties counted one half. Mann–Whitney compares every pair in the row, van Elteren only pairs in the same field and degree year. Intervals and $p$-values (P(beats) $= 0.5$) resample whole programs, 9,999 draws. Fields with fewer than 100 inventors in either group are omitted: Agriculture, Earth & Environmental.

Back

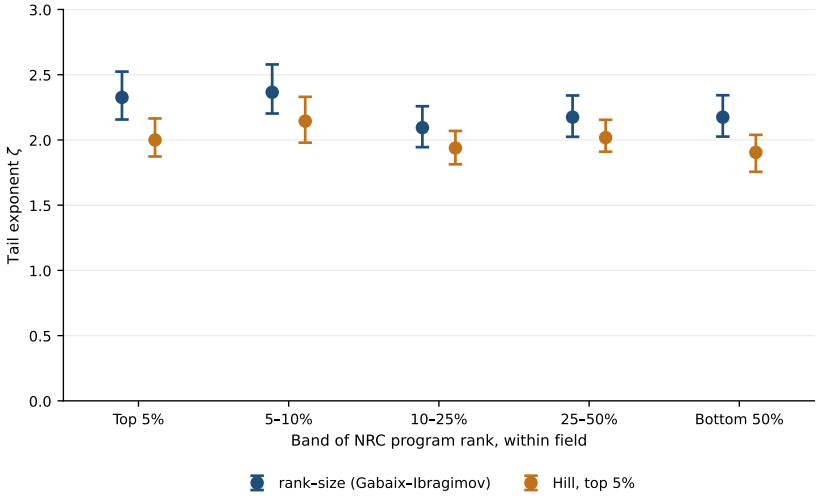
# Within Dispersion is not Concentrated at Top Programs



# The Tail is not Captured by the Top

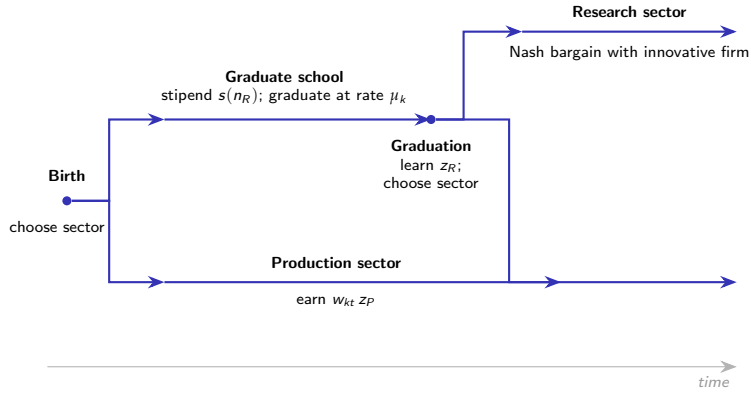


# Tail Index by Tier



# Baseline Model

# Timeline




Back

# University Problem

$$\rho U_{j,t} - \dot{U}_{j,t} = \max_{a_{j,t}(\cdot)} \int \mathbb{E}\left[z_R \mid n_R\right] a_{j,t}(dn_R) + U_N \dot{N}_{j,t} + U_P \dot{P}_{j,t}$$

Subject to:

$$K(x_{j,t}) \leq P_{j,t} + f_t N_{j,t},$$

Back

# Individual HJBs

- Production Worker

$$(\rho + \delta) W_t - \dot{W}_t = \ln\left((1 - \tau_W)(w_t z_P + \Pi_t)\right)$$

- Graduate Student

$$(\rho + \delta + \mu) G_t(n_R) - \dot{G}_t = \ln\left((1 - \tau_W)(s_t(n_R) + \Pi_t)\right) + \mu \mathbb{E}_{z_R}\left[\max\left\{W_t, R_t(z_R)\right\} | n_R\right]$$

Back

# Research Details

- Researcher

$$(\rho + \delta)R_t(z_R) - \dot{R}_t = \ln((1 - \tau_W)(\mathcal{P}_t(z_R) + \Pi_t))$$

- Matched innovative firm has value:

$$(r_t + \delta)\mathcal{V}_t(z_R; P_t(z_R)) - \dot{\mathcal{V}}_t = V_t\phi_r z_R - \mathcal{P}_t(z_R)$$

- Bargaining at every instant takes future path of payments as given.

Back

# Proposition 1

**Proposition 1**: *Suppose that $K(\cdot)$ is strictly convex and $K(0) = 0$. Then, along the balanced growth path, the equilibrium exhibits a positive ranking on expected research skill: if $\alpha_j > \alpha_{j'}$, then $\bar{m}_j \geq \bar{m}_{j'}$, where $\bar{m}_j$ is the average of $\mathbb{E}[z_R|n_R]$ among program $j$'s admits. If the distribution of $n_R$ is atomless, then this ranking is strict.*

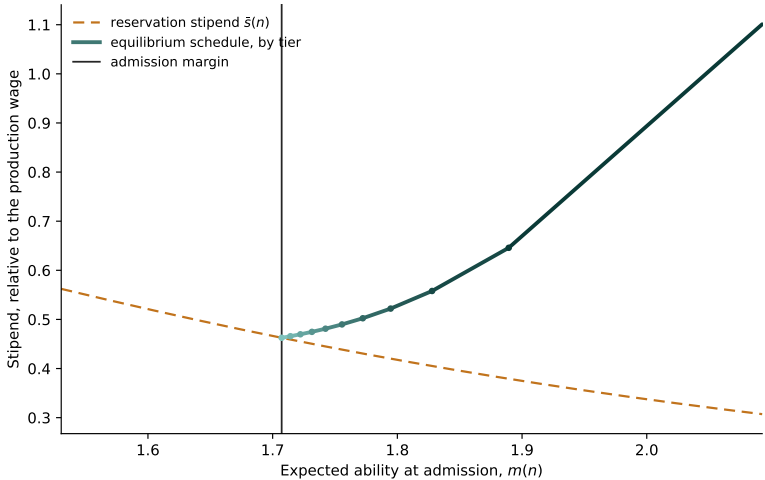
▸ Affine Schedule Lemma ▸ Schedule ◂ Back

# Characterizing the Stipend Schedule

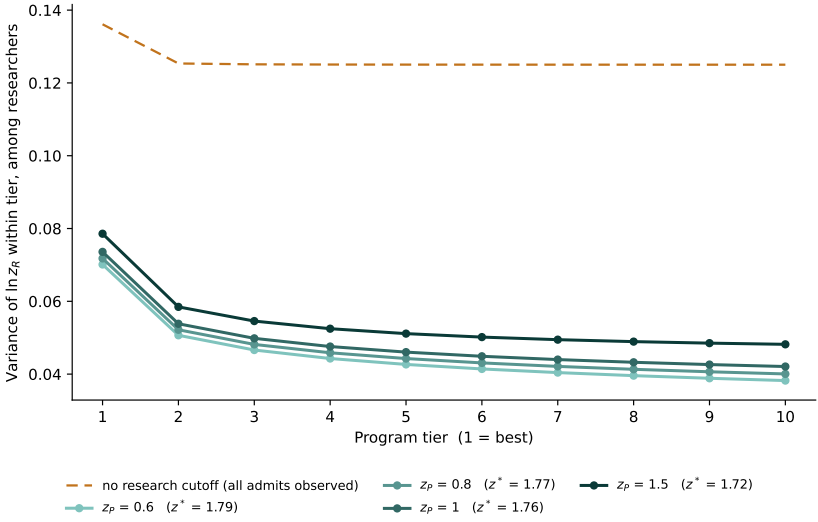
**Lemma 1**: *Order university tiers by productivity so that* $\alpha_1 > \alpha_2 > \cdots > \alpha_J$. *Then the equilibrium stipend schedule at any* $t$ *is piecewise affine in posterior expected productivity,* $\mathbb{E}[z_R|n_R]$, *continuous, and convex.*

Back

# The Stipend Schedule



# $z_P$ performance



# Quantitative Model

# Internally Calibrated Parameters: Aggregates and Programs

| Parameter | Description | Value | Targeted moment | Model | Data |
|---|---|---|---|---|---|
| **Aggregates** | | | | | |
| $\lambda$ | Innovation step size | 1.0231 | growth rate $g$ | 0.0200 | 0.0200 |
| $\varphi_R$ | Research productivity | 8.154 | PhD research payroll / GDP | 0.0060 | 0.0060 |
| $\theta$ | Researcher bargaining weight | 0.246 | incumbent rent passthrough (KPZW) | 0.61 | 0.60 |
| **Graduate programs** | | | | | |
| $\kappa$ | Admission cost, level | 59 | PhD holders / labor force (SDR) | 0.0041 | 0.0040 |
| $\nu$ | Admission cost, curvature | 0.072 | completions IRF, years 6 and 7 | 0.3% / 0.4% | |
| $\alpha_1$ | Program productivity, top tier | 0.386 | university services / GDP | 0.0035 | 0.0035 |
| $\alpha_1, \ldots, \alpha_{10}$ | Program productivity profile | 0.39–0.30 | relative decile sizes | 0.4% / 1.4% | |
| **Policy instruments** | | | | | |
| $f$ | Federal subsidy per student | 2.724 | federal R&D to universities / GDP | 0.0020 | 0.0020 |
| $s_R$ | R&D tax credit rate | 0.101 | R&D credit tax expenditures / GDP | 0.0006 | 0.0006 |
| **Luck and measurement** | | | | | |
| $\varrho$ | Within-person dispersion of output | 0.13 | patenting share, pooled | 0.178 | 0.179 |
| $\bar{A}$ | Patents per product line | 0.0101 | mean patents among patenters | 5.41 | 5.60 |
| | | | mean log count among patenters, by decile | 3.5% / 7.4% | |

▸ Externally Calibrated Parameters ◂ Back

# Internally Calibrated Parameters: Ability and Information

| Parameter | Description | Value | Identifying moments | Fit |
|---|---|---|---|---|
| **Research ability** | | | | |
| $\sigma_R$ | Dispersion of the normal component | 0.84 | Within-decile variance of log output | 3.7% / 6.5% |
| $\tau$ | Scale of the star component | 0.48 | Pooled tail index of patent counts (Hill, top 5%) | 1.99 / 2.06 |
| | | | Elite capture at four depths | 4.5% / 6.3% |
| **Information at admission** | | | | |
| $\sigma_\varepsilon$ | Signal noise (reliability $\beta = 0.62$) | 0.66 | Research share among PhD holders (SDR) | 0.688 / 0.681 |
| | | | Between-decile gradient of the patenting share | 4.8% / 17.4% |

Back

# Externally Calibrated Parameters

| Parameter | Description | Value | Source |
|---|---|---|---|
| **Preferences and demography** | | | |
| $\rho$ | Discount rate | 0.03 | Standard |
| $\delta$ | Death rate | 0.025 | Expected working life of 40 years |
| $r^*$ | World interest rate | 0.03 | Small open economy |
| **Graduate education and the labor market** | | | |
| $\pi_R$ | Eligible cohort share | 0.46 | CPS, labor force aged 25+ |
| $\xi_1$ | Relative productivity of ineligible workers | 0.62 | BLS median weekly earnings, 2024 |
| $\xi_2$ | Relative productivity of PhDs who exit research | 1.20 | PSEO, engineering: master's earnings premium |
| $1/\mu$ | Mean time to degree (years) | 5.5 | Survey of Earned Doctorates |
| $\mathcal{O}$ | Erlang stages (sd of time to degree 1.5 years) | 14 | Survey of Earned Doctorates |
| **Innovation** | | | |
| $\chi \ln \lambda$ | Growth from non-PhD sources (p.p.) | 1.12 | $2\% \times (1-$Patent Share$)$ |
| **Normalizations** | | | |
| $\mathbb{E}[\eta]$ | Mean of log research productivity | 0 | |
| $z_P$ | Baseline production productivity | 1 | |

Back

# Researchers: Ability

- Assume that

$$\ln(z_R) = \underbrace{\eta}_{\mathcal{N}(0,\sigma_R^2)} + \underbrace{\omega}_{Exp(\tau)}$$

$$\ln(n_R) = \eta + \underbrace{\varepsilon}_{\mathcal{N}(0,\sigma_\varepsilon^2)}$$

- In levels, this is the product of a Pareto and log normal.

# Researchers: HJBs

- Let $o$ denote the stage of an $O$-stage Erlang.
- HJB:

$$(\rho + \delta + O\mu)G_t(o, n_R) - \dot{G}_t = \ln((1 - \tau_W)(s_t(n_R) + \Pi_t)) + O\mu G_t(o + 1, n_R)$$
$$G_t(O + 1, n_R) = \mathbb{E}\left[\max\left\{R_t(z_R), W_t(\eta z_P)\right\} | n_R\right]$$

Back

# Graduate Program: sequence problem

- University of type $j$ chooses admissions $a_{j,t}(n_R)$ for each signal:

$$\max_{\{a_{j,t}(\cdot)\}} \int_0^\infty e^{-r^* t} \int \mathbb{E}\left[z_R \mid n_R\right] a_{j,t}(n_R)\, dn_R\, dt, \qquad m(n_R) \equiv \mathbb{E}\left[z_R \mid n_R\right]$$

- Subject to cohort break-even:

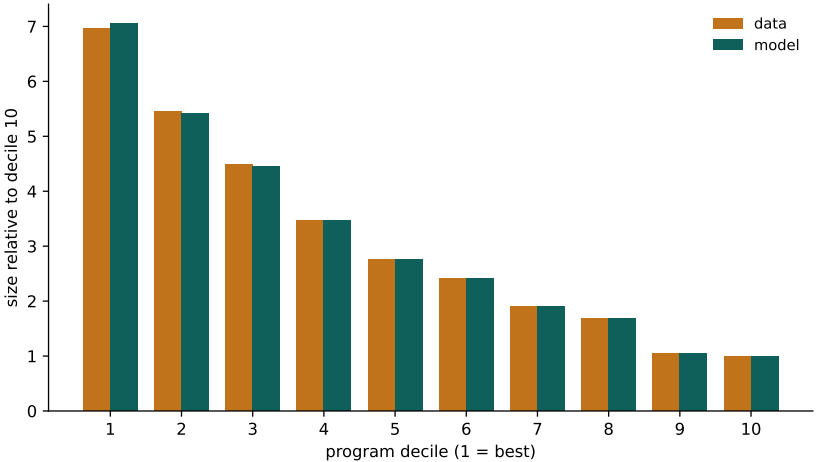
$$K\Big(\int a_{j,t}(n_R)\, dn_R\Big) = \int \Big[A_{\lambda,t}\big(\alpha_j m(n_R) - s_t(n_R)\big) + \mathcal{F}_t\Big]\, a_{j,t}(n_R)\, dn_R$$

$$A_{\lambda,t} = \int_0^\infty e^{-\int_0^u (r^*+\delta-g_{t+v})\, dv}\, S(u)\, du, \qquad \mathcal{F}_t = \int_0^\infty e^{-\int_0^u (r^*+\delta-g_{t+v})\, dv}\, S(u)\, f_{t+u}\, du$$

- Functional form: $K(x) = \kappa x^{1+\nu}/(1+\nu)$

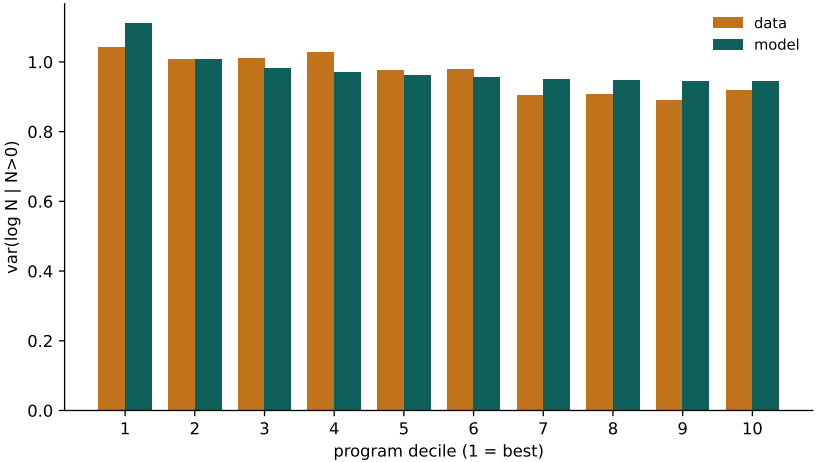
Back

# Program Sizes



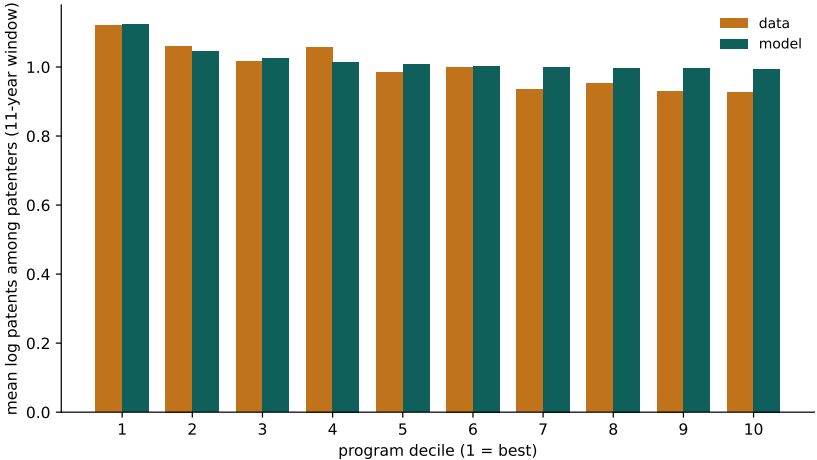

Back

# Within-Decile Variance



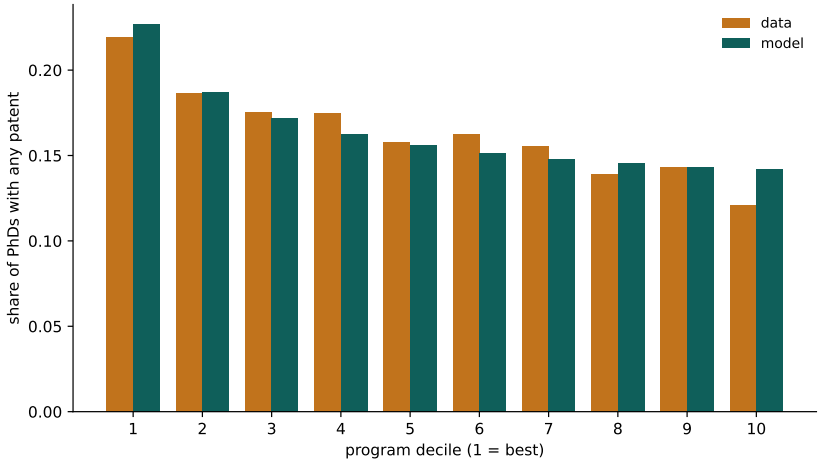

Back

# Mean Log Count



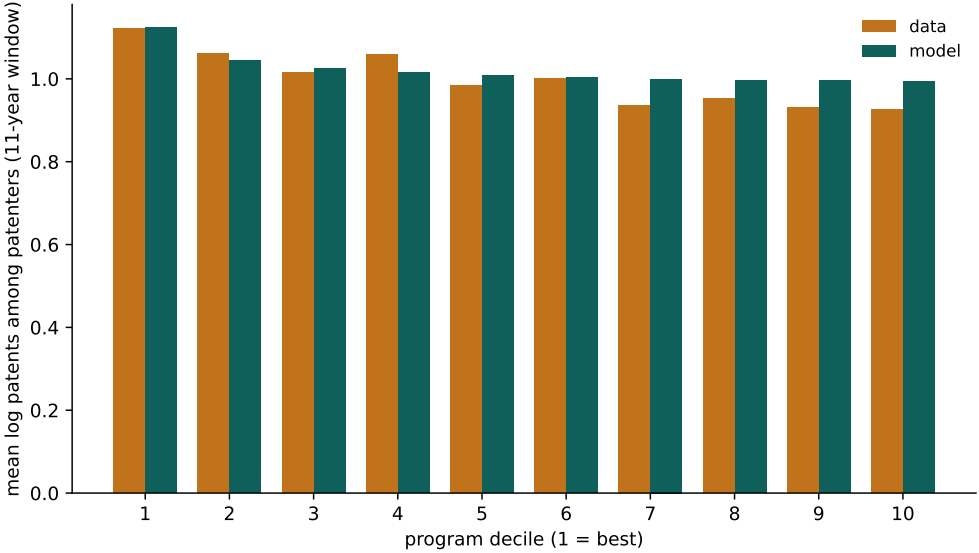

Back

# Patent Share

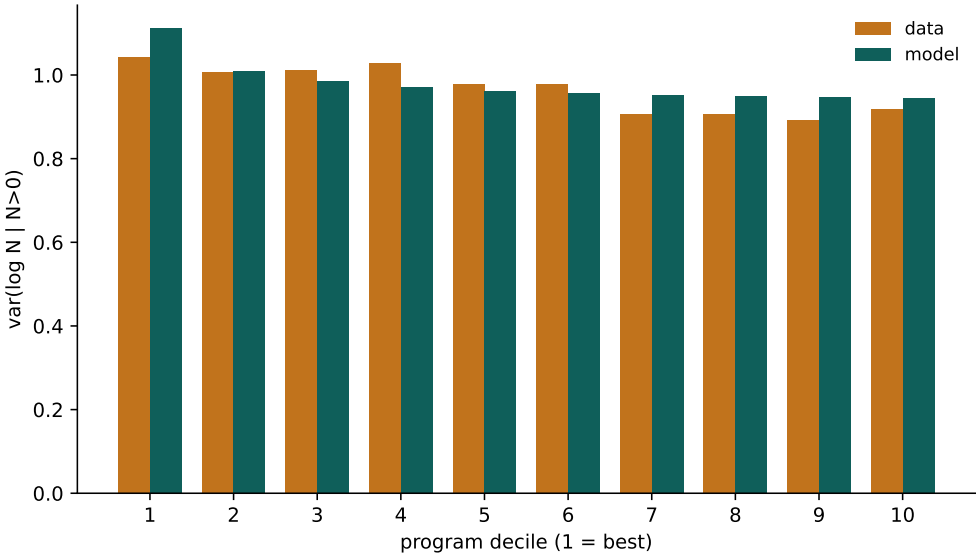



Back

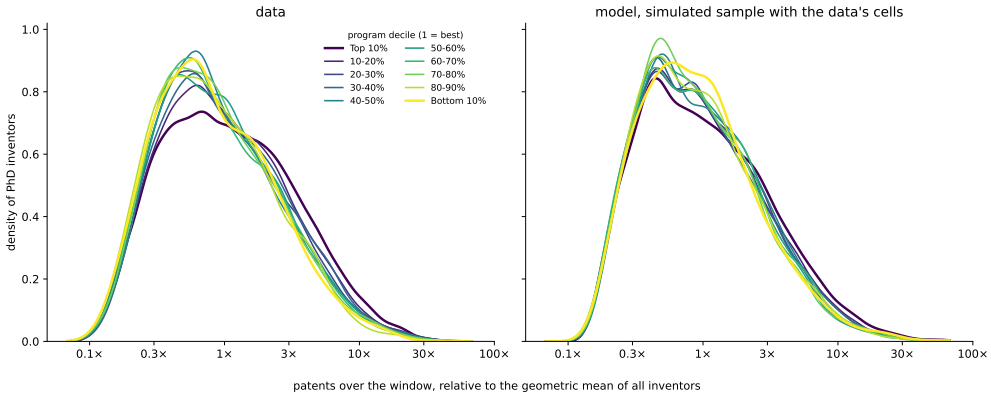
# Means


data
model
mean log patents among patenters (11-year window)
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8
9
10
program decile (1 = best)

# Variances

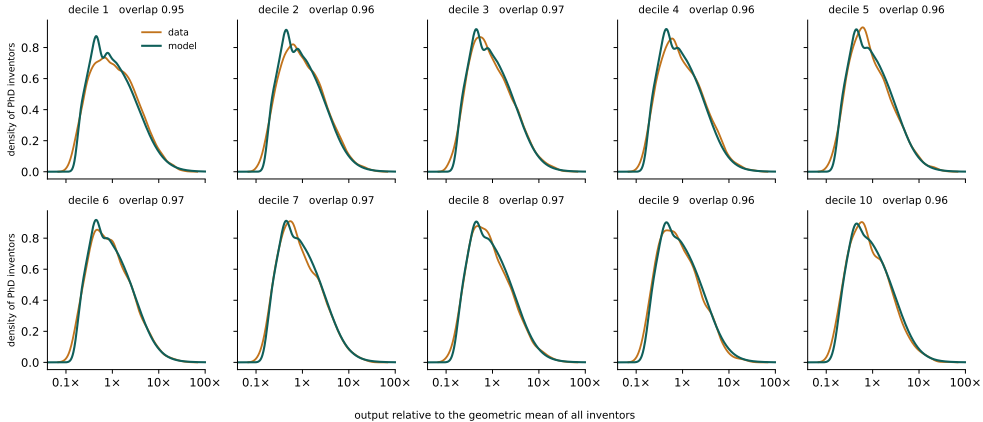

data
model
var(log N | N>0)
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8
9
10
program decile (1 = best)

# Cell Rate Adjusted



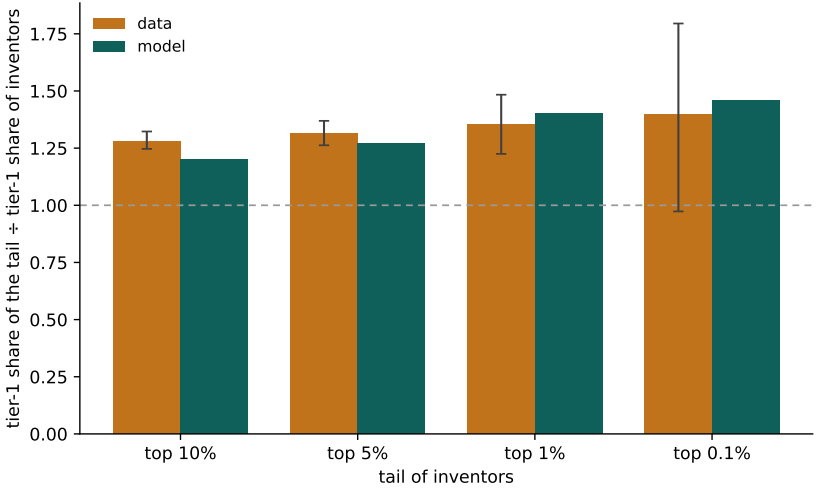

Back

# By Decile



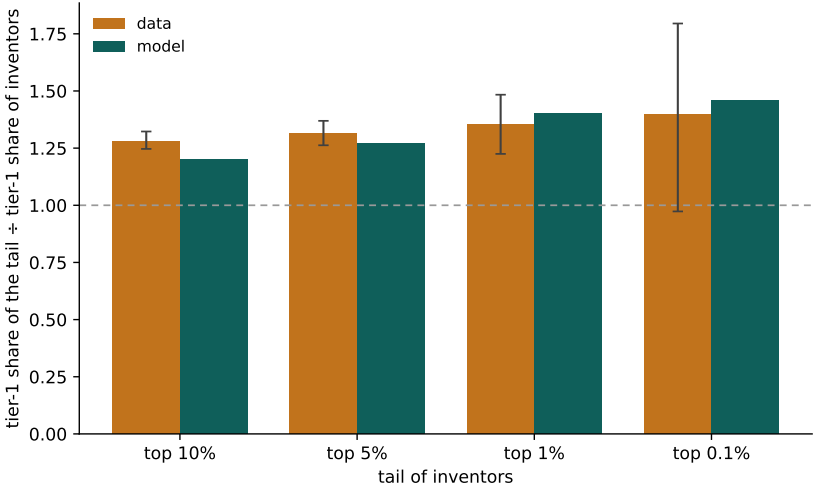

Back

# Identifying the Tail



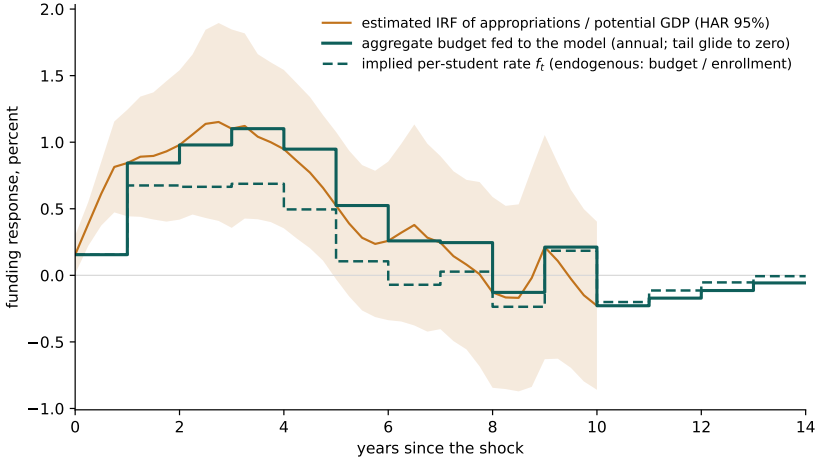

Tail Index by Decile
Back

# Capture Curve




Back

# Funding Input



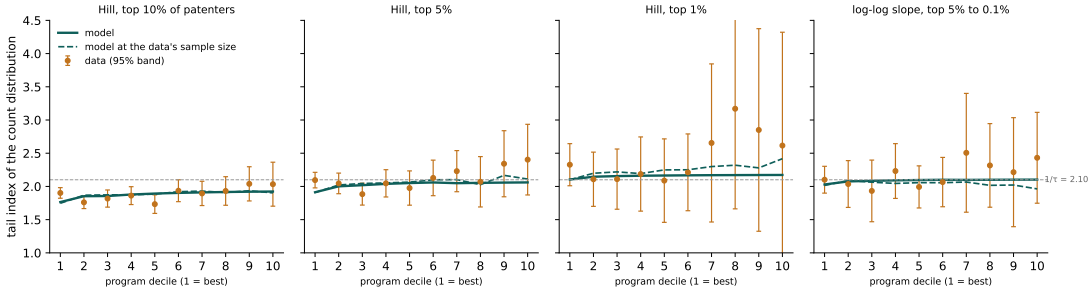

Back

# Untargeted Moments

| | decile 1 (best) | | decile 10 | | abs. % error, 10 deciles | |
|---|---|---|---|---|---|---|
| Untargeted moment | Model | Data | Model | Data | mean | max |
| share of patenters with exactly one patent | 0.323 | 0.332 | 0.354 | 0.409 | 6.2% | 13.3% |
| share of patenters with exactly two patents | 0.163 | 0.166 | 0.175 | 0.163 | 2.8% | 9.4% |
| share of patents held by the top 1% of PhDs | 0.341 | 0.362 | 0.400 | 0.466 | 9.4% | 14.1% |
| kurtosis of ln $N$ among patenters | 3.00 | 2.56 | 3.09 | 2.78 | 8.8% | 17.4% |
| skewness of ln $N$ among patenters | 0.73 | 0.47 | 0.79 | 0.66 | 23.1% | 54.9% |

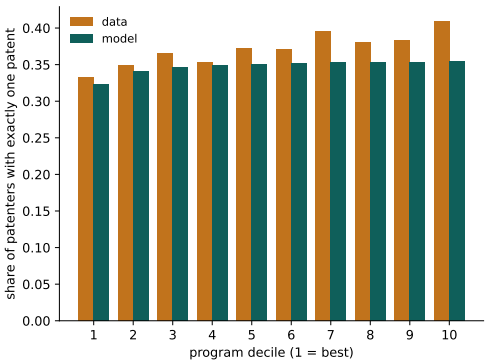
▸ Ones and Twos ▸ Top 1% Share ▸ Log-Count Shape ▸ Tail Index by Decile ▸ Tail Composition ▸ Danish Expansion ▸ Rank-Rank Slope
◂ Back

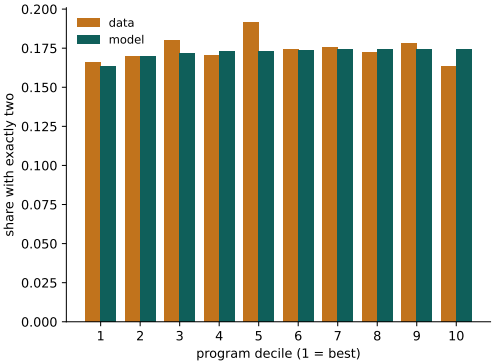
# Tail Index: Model vs. Data




◂ Identifying the Tail ◂ Untargeted Moments

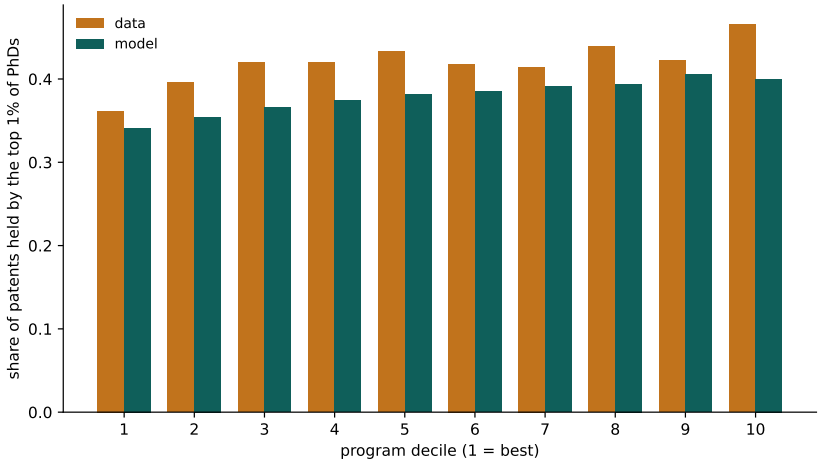
# Ones and Twos



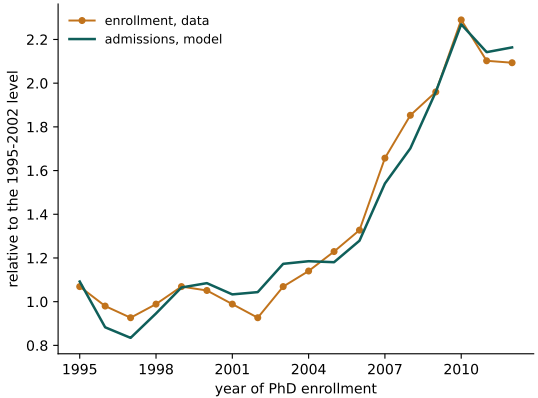

Back

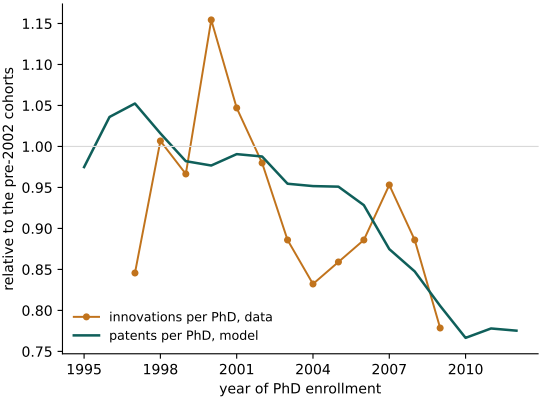
# Top 1% Share




Back

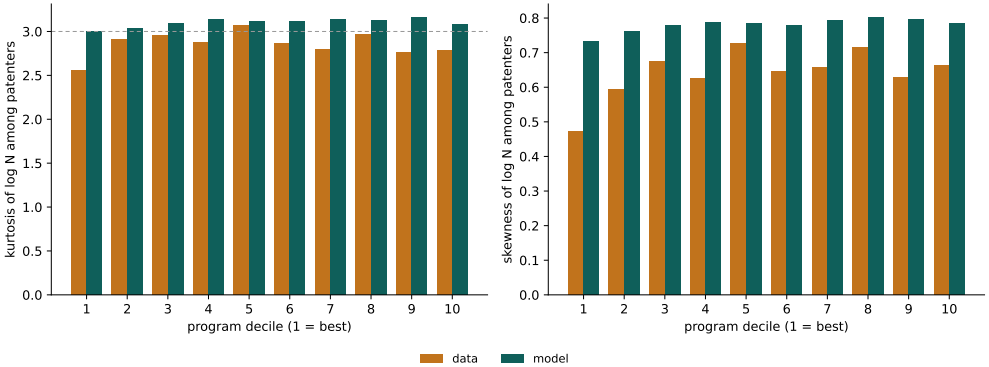
# The Danish PhD Expansion: Data and Model



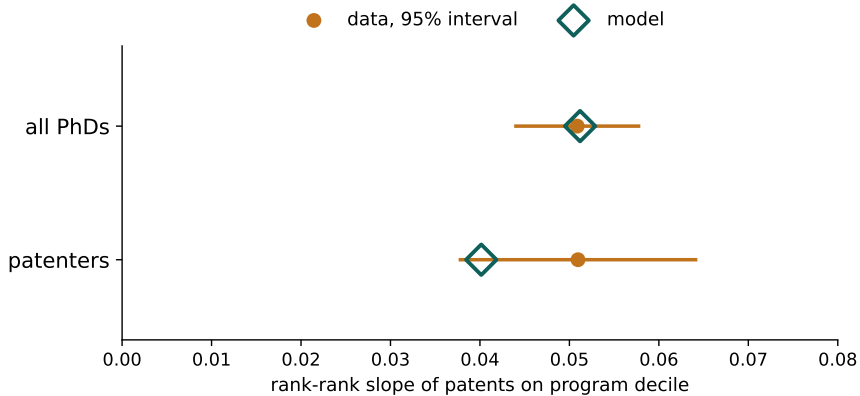

Back

# Log-Count Shape



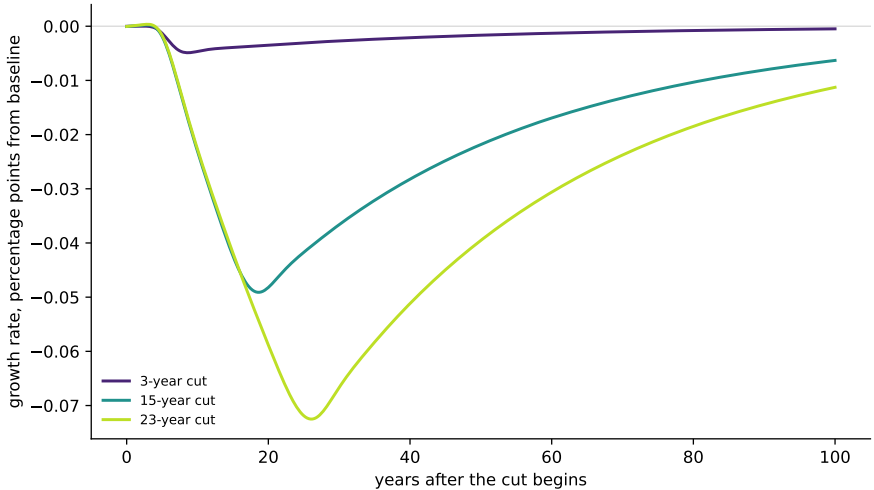

Back

# Rank-Rank Slope: Model vs. Data




Back

# The Growth Effect of Funding Cuts



# References I


Aghion, Philippe, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen (2017) "The social origins of inventors," technical report, National Bureau of Economic Research.

Aghion, Philippe and Peter Howitt (1992) "A MODEL OF GROWTH THROUGH CREATIVE DESTRUCTION," *Econometrica*, 60 (2), 323–351.

Akcigit, Ufuk, Jeremy Pearce, and Marta Prato (2025) "Tapping into talent: Coupling education and innovation policies for economic growth," *Review of Economic Studies*, 92 (2), 696–736.

Azoulay, Pierre, Joshua S Graff Zivin, Danielle Li, and Bhaven N Sampat (2019) "Public R&D investments and private-sector patenting: evidence from NIH funding rules," *The Review of economic studies*, 86 (1), 117–152.

Babina, Tania, Alex Xi He, Sabrina T Howell, Elisabeth Ruth Perlman, and Joseph Staudt (2023) "Cutting the innovation engine: how federal funding shocks affect university patenting, entrepreneurship, and publications," *The Quarterly Journal of Economics*, 138 (2), 895–954.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen (2019) "Who becomes an inventor in America? The importance of exposure to innovation," *The Quarterly Journal of Economics*, 134 (2), 647–713.

# References II


Campbell, Sheila, Jaeger Nelson, Eli A Schrag, Heidi L Williams, and Caleb K Wroblewski (2026) "Estimating the Economic Effects of Federally Funded R&D," technical report, National Bureau of Economic Research.

Chade, Hector, Gregory Lewis, and Lones Smith (2014) "Student portfolios and the college admissions problem," *Review of Economic Studies*, 81 (3), 971–1002.

Dyevre, Arnaud (2023) "Public R&D spillovers and productivity growth," technical report, Working paper. Eaton, J., & Kortum, S.(1999). International technology ....

Fieldhouse, Andrew and Karel Mertens (2025) "The returns to government R&D: Evidence from US appropriations shocks."

de Fraisse, Antoine Hubert, Peter Maxted, Kunal Sangani, and David Sraer (2026) "Pre-Trends Ahead of Monetary Policy Surprises: Recovering the Response of Corporate Investment."

Hendricks, Lutz and Oksana Leukhina (2018) "The return to college: selection and dropout risk," *International Economic Review*, 59 (3), 1077–1102.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow (2019) "The allocation of talent and us economic growth," *Econometrica*, 87 (5), 1439–1474.

# References III


Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy (2021) "Measuring technological innovation over the long run," *American Economic Review: Insights*, 3 (3), 303–320.

Manski, Charles F (1989) "Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon," *Economics of Education review*, 8 (4), 305–312.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny (1991) "The allocation of talent: Implications for growth," *The quarterly journal of economics*, 106 (2), 503–530.

Myers, Kyle R and Lauren Lanahan (2022) "Estimating spillovers from publicly funded R&D: Evidence from the US Department of Energy," *American Economic Review*, 112 (7), 2393–2423.

Pearce, Jeremy G (2020) *Idea Production and Team Structure*, Ph.D. dissertation, The University of Chicago.

Romer, Paul M (1990) "Endogenous technological change," *Journal of political Economy*, 98 (5, Part 2), S71–S102.

Terviö, Marko (2009) "Superstars and mediocrities: Market failure in the discovery of talent," *The Review of Economic Studies*, 76 (2), 829–850.

# References IV

Tham, Wei Yang, Joseph Staudt, Elisabeth Ruth Perlman, and Stephanie D Cheng (2024) "Scientific talent leaks out of funding gaps," *arXiv preprint arXiv:2402.07235*.

Wu, Alice (2026) "Employer Learning, Sorting, and Productivity: Evidence from Computer Scientists," technical report, ROCKWOOL Foundation Berlin (RFBerlin).