Booms, Busts, and the PhD Pipeline: The Career Impacts of Biomedical Science Funding Volatility

Max Mayca* Gerald Marschke†

March 2, 2026

Abstract

Federal funding is the primary source of support for basic science, sustaining the laboratories where research is conducted and new scientists are trained, but it is often subject to significant variability. This study examines the effects of two types of funding shocks on research outcomes: system-wide shocks, such as the NIH doubling, and idiosyncratic shocks, such as when individual investigators unexpectedly secure multiple grants simultaneously. Using combined data at the investigator level on publications, theses, and grants, we find that while Principal Investigators (PIs) produce more papers during funding surges, these papers are of lower average quality and produced in labs that expand and reorganize their teams. PhD students who trained under PIs exposed to system-wide funding surges face weaker job-market outcomes and lower subsequent research productivity, consistent with graduating into a labor market temporarily flooded with new scientists. Idiosyncratic funding shocks, which do not create such market-level supply pressures, nonetheless also reduce the research productivity of affected mentees after graduation. Together, these patterns suggest that rapid expansions in funding can dilute effective supervision and strain mentoring capacity, lowering the quality of both scientific output and the next generation of researchers.

*University at Albany, SUNY.

†University at Albany, SUNY and NBER. Email: marschke@albany.edu. We thank Neil Smalheiser and Vetle Torvik for the use of their Author-ity data. We gratefully acknowledge support from National Institute on Aging, Office of Behavioral and Social Sciences Research, and National Science Foundation (NSF) SciSIP through P01 AG039347 (PI, Bruce Weinberg), and NSF 1661278 and 1918445 (PI, Marschke). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the National Institutes of Health. All errors are the responsibility of the authors.

I Introduction

The National Institutes of Health (NIH) is the primary source of support for basic biomedical science, sustaining the laboratories where research is conducted and new scientists are trained, but it is often subject to substantial variability. Over the past half-century, NIH appropriations in real terms have followed a boom–bust path with sizable swings: an unusually rapid doubling in 1998–2003, followed by a decade of erosion in purchasing power, a temporary ARRA spike in 2009–2010, and sequestration-era cuts that left real funding in the 2010s well below its early-2000s peak. At the lab level, these aggregate changes—together with uncertainty in grant adjudication—can translate into meaningful year-to-year variation in funding received by individual principal investigators (PIs).

How a sudden funding increase translates into scientific output depends on how quickly laboratories can scale and on which labor inputs adjust. Existing evidence documents that funding interruptions and shortfalls can reduce research output and disrupt scientific careers (e.g., Babina et al., 2023; Tham et al., 2024). This paper examines how *positive*, short-lived funding surges affect training conditions and the subsequent careers of the cohorts trained during those expansions.

The mechanisms are especially salient given an asymmetry in labor supply elasticities. In the short to medium run, the supply of independent PIs is inelastic, whereas the supply of PhD students and postdocs is relatively elastic.[1] Sudden funding increases therefore tend to scale up existing labs by hiring more junior researchers rather than by creating many new independent laboratories. This tendency may be amplified if incremental NIH funding is disproportionately awarded to established investigators; for example, Larson et al. (2012) documents that the average age of scientists receiving grants increased from 1998 to 2003. Together, these features imply that funding surges are likely to raise team size and publication volume, while creating countervailing pressures on training and research quality insofar as mentoring and supervisory capacity adjusts slowly.

These concerns appear prominently in policy-oriented analyses of the NIH doubling and its aftermath (Freeman and Van Reenen, 2009; Teitelbaum, 2008). A central channel is the training pipeline: expansions disproportionately increase the number of trainees (especially postdocs) relative to the number of independent positions, so when funding growth ends, large cohorts face a contraction in demand for research labor and elevated career risk. A related feature of the NIH funding environment is that low paylines and high uncertainty encourage investigators to submit multiple proposals as a form of insurance rather than calibrating submissions to a single target scale of operations (Freeman and Van Reenen, 2008, 2009). Even holding aggregate funding fixed, this behavior implies that some investigators will experience short-run *windfalls*—more proposals funded in a brief window than they anticipated ex ante.[2]

[1] Goolsbee (1998) argues that increases in R&D funding primarily raised wages for highly trained researchers rather than substantially increasing their numbers, consistent with inelastic short-run supply.

[2] Illustrative qualitative accounts describe strategies such as submitting "many grants a year" and "taking as many

In this paper we study two kinds of positive shocks. The first is a system-wide surge—the NIH doubling (and the temporary ARRA expansion)—which shifts the aggregate funding environment and can generate cohort-wide pipeline effects. The second is idiosyncratic PI-level funding windfalls that occur even in the absence of aggregate budget shocks. Our empirical strategy treats these PI-level windfalls as shocks to lab scale that are plausibly distinct from the steady, anticipated component of a PI's funding stream: low success rates, combined with multiple submissions per PI, occasionally generate large, discrete jumps in support for individual labs. We then ask how these windfalls reverberate through the careers of the next generation of scientists trained in those labs.

These concerns about funding volatility map naturally to a dynamic training problem. When funding surges, labs scale up primarily by hiring more PhD students and postdocs rather than creating new independent labs, since the supply of PIs adjusts slowly. This creates a trade-off: increased production today comes at the cost of diluted mentoring per trainee. Moreover, when funding growth ends, large cohorts trained during expansions face a labor market with constrained demand for new PIs. We formalize this intuition in a simple overlapping-generations model (Appendix B) in which research output depends on both current mentor intensity and the quality of training received by previous generations. The model shows that temporary funding surges can push the system off its optimal growth path, reducing both current research quality and the long-run productivity of scientists trained during the expansion.

Related work shows that funding increases can have sizable ripple effects on research teams—including changes in staffing—with downstream implications for scientific output (Sattari et al., 2022). We complement this evidence by focusing on human-capital formation *within* labs: how short-run windfalls affect training intensity, training quality, and trainees' subsequent careers.

To do so, we link NIH administrative grant records to biomedical publication histories and U.S. dissertation-based advisor–advisee relationships using large-scale person identifiers. We measure research output and idea content from MEDLINE-based publication records and use citation links to construct forward citations in a fixed window. We also construct advisor-year measures of doctoral training intensity and student-level early-career outcomes in the years following graduation. Using staggered difference-in-differences and panel event-study designs, we show that funding surges increase measured production on impact (publications and trainees), but reduce average quality on some margins as labs scale. Consistent with a training-channel interpretation, trainees who pass through labs during surge periods exhibit weaker early-career outcomes, including lower bibliometric measures and reduced success in securing early-career NIH funding.

shots as we can" to increase the probability that at least one proposal is funded (Webel, 2025; Drugmonkey, 2025; Dobson, 2025).

II NIH through the doubling and ARRA eras

From 1998 to 2003, Congress doubled the NIH's budget from $14 billion to $27 billion, significantly increasing both the quantity and size of research grants. Although this surge in funding was expected to improve both the quantity and quality of research output, doubts have been raised about its effectiveness. For example, Freeman and Van Reenen (2008) argue that the period of doubling did not only fail to deliver the anticipated boost in research production, but also had adverse impacts in subsequent years. As funding levels were not sustained, budget cuts and interrupted projects became more common, hampering research continuity and long-term progress.

Another positive shock occurred in 2009 with the American Recovery and Reinvestment Act (ARRA). Although not as extreme as the doubling, it serves as a secondary shock for our preliminary study. However, the details of the two shocks differ greatly. Indeed, ARRA was the Obama government-led fiscal stimulus in response to the Great Recession. The ARRA stimulus was probably understood by PIs and research universities to be short-lived. The doubling, however, was not explicitly communicated as temporary when it was implemented. In fact, many universities and research institutions made long-term commitments and expansions based on the assumption that higher funding levels would be maintained or continue to grow at normal rates after 2003.

Using data from the Graduate Student Survey (GSS), the Integrated Postsecondary Education Data System (IPEDS), and NIH RePORTER data, we can make a preliminary assessment on the effect of budget increases on research inputs such as number of PhD students and number of PIs. We separate out the top 20 schools as schools which receive the most funding. We take the top 20 schools because the distribution of scientific activity is highly skewed with a relatively small number of schools responsible for most of the research grant money and most of the research output. Moreover, faculty researchers in the top 20 universities are intensely engaged in mentoring PhD students and postdocs while simultaneously engaged in research.

Figure 1 panel (a) shows NIH expenditures per year derived from NIH RePORTER data. It shows a steady increase during the doubling period and the sharp, temporary increase under the ARRA stimulus. Notably, the data suggest that top 20 schools did not gain a disproportionate advantage during the doubling—in fact, their share of NIH funding declined slightly during this period. This pattern suggests that the increased funding was spread relatively broadly across institutions rather than concentrated in the most well-funded universities. However, after the doubling ended, the funding distribution shifted, and by the time of ARRA, the top 20 schools had regained and even increased their share of NIH funding. This reversal suggests that while the doubling initially supported a more decentralized allocation of resources, subsequent funding cycles reinforced the concentration of research funding among well-established institutions.

The data suggest that the expansion of biomedical research during these funding surges was achieved primarily by increasing the number of PhD students, rather than by hiring more postdocs. Figure 1 panel (b) shows the number of full time PhD students in the life sciences increases shortly

after the doubling begins, particularly in top 20 schools. This increase lags the budget increase slightly, consistent with the time required to recruit new students. We can also observe a spike in doctoral students during Covid probably caused by potential graduates delaying their graduation because of the state of the job market. In contrast, Panel (c) shows that the number of postdocs experienced no or only a small change in the trend during the doubling, suggesting that PIs met their increased staffing requirements more with additional graduate students rather than postdoctoral researchers. However, during the ARRA period, postdoc appointments show a pronounced spike, consistent with the expectation that short-term funding increases are more likely to result in hiring postdocs or research staff under short-term contracts rather than committing to multi-year PhD training. PhD students not only require a longer commitment but are often ineligible for intensive research in a lab until after coursework.

Figure 1 Panel (d) shows a delayed increase in the number of distinct PIs receiving NIH funding during both the doubling and ARRA periods. The rise in distinct PIs, coupled with a decline in the proportion of PIs from top 20 programs, suggest that funding increases were not disproportionately allocated to well-established, well-funded researchers and institutions. This pattern runs counter to the expectation of a Matthew effect in funding allocation (Bol et al. (2018)). However, during and after ARRA period, the share of funding claimed by top 20 programs rose sharply, surpassing pre-doubling levels.

Figure 1 Panel (e) shows that the ratio of student, postdoc, or graduation per PI. The overall students-per-PI ratio shows a subdued but notable shift, with the decline in the ratio of doctoral students per NIH-funded PI reversed midway through the doubling, rising modestly through 2008. This suggests that while the increase in new PIs helped distribute the mentoring burden more broadly, established PIs still took on more students as the period progressed. At top universities, the pattern is similar but more pronounced. Panel (f) shows that at the top 20 institutions, the number of students per PI also increased midway through the doubling, reversing the prior downward trend. However, these institutions exhibit a distinct response in their reliance on postdocs: top schools consistently have more postdocs per PI, and this trend accelerates sharply during the ARRA period.

The descriptive trends suggest that funding increases were accompanied by adjustments in the research workforce, with evidence that the mentoring burden on PIs increased. During the doubling, the number of PhD students per PI rose, reversing a prior decline, while postdoc numbers remained relatively stable. Under ARRA, postdoc appointments increased sharply, particularly at top universities, likely reflecting a preference for hiring more immediately productive researchers in response to short-term funding boosts.

These figures have limitations. Graduate student counts are not necessarily tied to NIH-funded PIs, as they come from university-level sources rather than NIH grant records, making direct comparisons between PIs and student or postdoc numbers imprecise. They imprecisely show how expansions play out in the lab size and composition. This pattern is consistent with evidence from Sattari et al., who find that additional funding tends to expand research teams by adding non-

PhD student and -postdoc persons, such as professional research staff undergraduates more than postdocs and graduate students.

To better understand how funding shocks shaped research training and careers, we turn to PI-level data linking grants, publications, and student outcomes. Increased mentoring demands may have weakened training, leading to worse outcomes for doctoral students. Alternatively, expanded funding may have improved training by providing access to better equipment, larger-scale projects, or more research experience. However, it could also have shifted students toward marginal projects with lower long-term returns, reshaping the scientific workforce in unexpected ways.

Our multivariate analysis directly tests these possibilities by examining how funding shocks at the PI level influence research output, student outcomes, and career trajectories. We now turn to this analysis.

III Theoretical framework

Appendix B introduces a simple overlapping-generations framework that we use to motivate the empirical work. Young researchers (PhD students and postdocs) collaborate with old researchers (PIs) to produce papers. The key variable is the mentor ratio: the amount of PI time available per trainee. This ratio affects both immediate research quality and the long-run productivity of the next generation, since today's trainees become tomorrow's PIs.

A sudden funding increase allows existing PIs to rapidly expand their labs by hiring more trainees. This expansion lowers the mentor ratio—spreading PI attention more thinly across projects and students. Research output increases (more papers), but average quality may decline due to reduced supervision per project. More critically, trainees receive lower-quality mentoring, affecting their subsequent research careers.

The resulting optimal dynamics exhibit saddle-path stability (see equation (B.5)). There exists a unique stable trajectory converging to a steady-state mentor ratio. If the initial ratio is sufficiently low, the optimal allocation progressively shifts resources toward young researchers, reducing mentoring intensity and eventually leading to a vanishing stock of senior researchers. Conversely, if the ratio is sufficiently high, resources shift toward senior researchers until feasibility constraints bind.

A key constraint limits the number of senior researchers to the previous period's stock of young researchers. This imposes a feasibility boundary on the system. Under a constant budget, the economy converges to a steady-state mentor ratio provided it lies within the feasible region. Otherwise, the constraint binds and determines a new steady state.

Budget dynamics play a crucial role. Rapid budget expansions shrink the feasible region because hiring must initially expand through young researchers. This temporarily lowers the mentor ratio, increasing quantity at the expense of quality. If the expansion is sufficiently large, the system may

cross the stable path. When the budget later stabilizes or declines, the economy may move toward a low-mentoring equilibrium characterized by declining senior researcher capacity.

The model therefore highlights a central policy implication: gradual funding increases allow the system to adjust the mentor ratio smoothly, preserving long-run research quality. By contrast, large temporary funding spikes risk destabilizing the research workforce and inducing persistent quality losses.

IV Data

We combine three main data sources, as shown in Figure 2. First, we use dissertation records to identify biomedical PhD graduates and advisor–advisee links. Second, we use NIH grant records to measure both aggregate funding conditions and investigator-level award receipt. Third, we use publication, citation, and text-based measures to follow scientists' research output, impact, and novelty over time. Our analyses use annual panels: a PI-year panel for publication outcomes, an advisor-year panel for mentoring outcomes, and an advisee-level file (summarized in fixed post-PhD windows and attached to the advisor-year of graduation).

A central challenge is linking individuals consistently across data sources. We address this by anchoring all person-level information to the Author-ity disambiguation of PubMed/MEDLINE developed by Torvik, Smalheiser, and coauthors (Torvik et al., 2005; Torvik and Smalheiser, 2009). Author-ity assigns publications to *author clusters* (probable authors) and provides a stable Author-ity identifier (AUID) for each cluster. The Author-ity release we use is based on a PubMed snapshot taken in July 2009, and thus provides publication histories through that date. We then use the companion *authorlink* (crosswalk) files to map external administrative records—notably NIH investigators and dissertation authors/advisors—into this common AUID space using probabilistic matches (retaining links above the recommended match-probability threshold).[3]

NIH ExPORTER is the NIH's public administrative database of funded grant applications and awards, providing project identifiers, activity codes, institutes/centers, dates, and annual costs, along with named investigators. We restrict attention to the NIH's core investigator-initiated research-project mechanisms—so-called *R01-equivalent* awards (DP1, DP2, DP5, R01, R37, R56, RF1, RL1, U01, and R35)—and omit mechanisms primarily designed for training, fellowships, or career development, which differ in purpose and funding structure. Using the NIH authorlink crosswalk, we map ExPORTER investigators to AUIDs via probabilistic matches (the public authorlink release retains links with match probability exceeding 0.5) and then collapse grant support across projects to the PI-year level. This yields annual investigator-level measures of award receipt and

[3]Author-ity 2009 and the corresponding NIH/NSF/USPTO authorlink files are publicly available via the Illinois Data Bank; see `https://databank.illinois.edu/datasets/IDB-4222651` and `https://databank.illinois.edu/datasets/IDB-4370459`. Benchmark evaluations report high accuracy for Author-ity at scale (e.g., recall 98.8% and precision about 98%), implying that a small share of publications are misattributed due to residual splitting or lumping errors.

award amounts that can be merged directly to AUID-year publication outcomes constructed from Author-ity/MEDLINE.

We identify doctoral training relationships using ProQuest Dissertations & Theses (PQDT), a comprehensive bibliographic repository of dissertations and theses submitted by universities; university library documentation commonly describes PQDT as indexing roughly 90% of North American doctoral dissertations accepted each year. The PQDT data have This dataset, spanning 1867 to 2013, includes 1.2 million dissertations and theses. Using the ProQuest–Author-ity linkage file, we map dissertation authors and primary advisors to AUIDs and retain dissertation metadata (degree type, institution, field, and year of completion). We restrict to doctoral dissertations. Appendix 1 describes our cleaning and timing rules to determine PhD year and for dealing with multiple doctoral theses (Appendix 1).

These data enter our analyses in two ways. First, for advisor-year analyses we construct doctoral training intensity as the number of graduating doctoral dissertations in year t for which advisor i is listed as the primary advisor. Second, for student-outcome analyses we link each advisee AUID to publication and (when relevant) NIH funding records and summarize post-PhD outcomes in a fixed window following dissertation completion (e.g., outcomes in years g through $g + 5$), then attach these outcomes to the advisor-year corresponding to graduation year g.

We measure research output and text-based novelty using Author-ity/MEDLINE publication histories (publications through July 2009). We measure citation-based impact using Web of Science/Science Citation Index Expanded (SCIE) citation links for the MEDLINE-matched set of articles; to ensure comparability across cohorts we construct forward citations within a fixed window (five years after publication). We also incorporate the text- and diversity-based metrics from Yu et al. (2023); because these measures are not available for all authors/papers in our extracted files, missing values reflect lack of coverage rather than zeros.

We merge NIH funding measures, ProQuest mentoring measures, and publication outcomes by AUID and year, restrict the analysis window to 1985–2008, and reshape to an annual panel, filling missing years with zeros for production measures and trimming inactive years at the beginning and end of observed careers (Appendix 1).

After collapsing at the individual-year level, we can then merge this dataset with our two previous ones obtaining a final dataset of 21.7 million observations. Moreover, by combining the data we have from ProQuest and Author-ity we can obtain publishing information on advisees per advisors. We then only keep observation within the common range (1985 to 2008) and drop individuals who have never had advisors or publications (we keep individuals who have never received grants) while expanding for missing years within an individual's career and account for inactivity at the beginning and at the end of each advisor's career. Additionally, we drop individuals with production (advisee, publication or grant) on fewer than 10 years during our sample. The resulting analytic panel contains 1,135,640 person-year observations.

The descriptive statistics are detailed in Table 1. All the observations are at the advisor-year level and a short description of the variables can be found in Table 2. Forward citations are measured five years after publications to avoid any truncation problems (the dataset containing publication information, Author-ity, goes to 2013). Fractioned variables are the ratio of the variable of choice with the team size working on the project. Student quality observations are measured at the advisor level from the time of graduation to five years afterward, e.g., a student graduating in 2000 will appear on his advisor's observations for the year 2000 and will count all the papers published by said student between 2000 and 2005. Mean team size is the average of authors per paper per year for each advisor, while lab size represents the total amount of collaborator per year for each advisor (a distinction is that lab size do abstraction of multiple collaboration with the same author throughout a year). As we can see, the number of observations greatly decreases for our study of student's quality which could lead to a low-level of predictive power for our analysis.

We also repeat the same process keeping only the 20 universities which received the most funding. The top 20 universities are: Johns Hopkins University, University of California San Francisco, University of Michigan, University of Pittsburgh, University of Pennsylvania, Columbia University, Yale University, Washington University, Stanford University, University of California San Diego, University of Washington, Duke University, New York University, University of North Carolina Chapel Hill, University of California Los Angeles, Emory University, University of Minnesota, University of Wisconsin Madison, Northwestern University, Harvard University. ProQuest has information about the institution of the advisor of a thesis. As such, we can identify advisors who works for such establishments.

V Model

We investigate how unexpected increases in PI funding affect (1) PI research output and labor organization, including the number of publications, team size, and collaboration structure, and (2) the development and career outcomes of PhD students trained under PIs experiencing funding shocks. These two dimensions are interrelated: how PIs scale up their lab operations in response to additional funding affects the training environment for PhD students.

In response to a positive shock, we expect PIs to increase research output by publishing more papers, expand their research teams by hiring more PhD students and postdocs, and collaborate with more co-authors. If these additional publications involve newly recruited PhDs, postdocs, or more senior collaborators—or if new PhDs and postdocs are not yet fully integrated into the research process—publication quality may decline

There is a timing issue: PIs cannot immediately bring on new research personnel, as funding must first be secured, and then time is required to recruit and integrate them into the lab. For example, PhD students must complete coursework before becoming fully engaged in research, and postdoc hiring requires identifying and vetting qualified candidates. In the meantime, the PI's

research output may change in ways that affect quality. This could happen because the PI is stretched across more papers, or because the marginal research projects funded under the shock are of lower quality. We measure these effects using total publications per PI and total citation-weighted publications (overall productivity), number of co-authors per PI (team or lab size), and co-authors per paper (how collaboration is structured). This allows us to confirm that the funding shock expands research activity while also examining whether it reshapes how research is structured and produced in the lab, before turning to its effects on student careers.

When funding increases, a PI must expand their lab, often by hiring more graduate students or postdoctoral researchers. This creates a potential mentorship trade-off. More PhD students and postdocs mean mentorship is spread thinner, reducing per-student investment and leading to lower-quality training. If PIs need to scale up quickly, they may lower admission standards, bringing in weaker students, which could reduce research output. If the PI expands into lower-value projects, students may work on less impactful research, affecting their long-term success. We test these hypotheses by examining the number of PhD students graduating under the PI, student publication and citation records and post-graduation funding success (early-career productivity).

To test those hypotheses, we estimate the effect of a positive funding shock on research outcomes using a difference-in-difference model with PI and year fixed effects.

$$y_{i,t+n_j} = \beta \text{Shock}_{i,t} + \gamma_i + \delta_t + \epsilon_{i,t} \tag{1}$$

where i indexes the PI, the period (yearly or in ranges of years) is denoted by t, and y represents the research output of the PI or the results of the PhD student. Fixed effects γ and δ account for the PI-specific ability and broader funding trends, respectively. Since it will take time from the moment a laboratory receives funding to when the output is produced, we include an effect n_j which lags the dependent variable, with $j \in \{1, 2\}$, $n_1 = 1$, and $n_2 = 2$, differentiating between short-term and slightly delayed effects of funding shocks. Variables related to PI production (e.g., publications, citations) are expected to respond more immediately.[4] Student productivity outcomes take longer to manifest because newly recruited PhD students typically take several years to complete their research and graduate. As a result, we expect the effects of a funding shock on student career outcomes to emerge only after a lag of a few years. We run this analysis for both the overall sample and within the top 20 universities only.

The key independent variable, *Shock*, is an indicator of whether the PI receives an unexpected, substantial increase in funding. We use a binary indicator instead of the level of funding to capture cases where the PI must rapidly scale up their lab in a way that is not easily adjustable. If we used funding levels as a continuous variable, the model would capture gradual adaptations, which

[4]When a PI experiences a funding shock, there is typically a delay before new PhD students are recruited and trained. During this transition period, research output may still increase as the PI may shift their publishing strategy to increase publication counts, for example, by splitting the same research findings into several smaller papers of lesser quality. This would allow them to increase their publication rate even before new students are in place and productive, consistent with incentives created by funding agencies.

PIs can often manage smoothly. Instead, we focus on cases where the scale of the PI's operation must expand quickly, forcing them to hire additional personnel and restructure their research efforts—exactly the conditions under which the constraints we describe above should bind. We now need to define what a shock is.

1. NIH Shock:

NIH's Institute Centers (IC) all saw a significant increase in funding around the same period, although some ICs were shocked slightly earlier and others later. Instead of sorting between ICs receiving different levels of raises, we separate the control group as individuals who do not receive any sort of funding or receive most of the funding before getting shocked (the current threshold is 75%) and the treatment group as researchers who receive a significant amount of their funding while the IC is getting shocked.

2. Individual Shock:

PIs need sustained grant funding to run their labs, but because NIH proposals are often rejected, they submit more proposals than necessary. Occasionally, multiple proposals are unexpectedly funded, creating a positive funding shock that forces rapid expansion. At the individual level, a researcher is considered shocked if the moving average of his funding increases by more than 80% from one year to the next. Therefore, the control group is made up of individuals who have never had a big funding burst but more likely received their funding more regularly.

The time distribution of the first period of shock greatly differs between our two definitions as shown in Figure 3. Since IC shocks mostly happen during the 1998-2003 period, initial period of shocks will be condensed through the same period. For individual shocks, it seems to follow a normal distribution centered around 2000, since individual shocks will also arise based on career experience even without budget increase from the funding agency.

We also want to test whether an individual being shocked affects the output of the rest of its department, granted that they have not been shocked yet. We define a department as a combination of field and university. For example, a researcher publishing from Boston University in the field of Biology is considered to belong in that specific department. Due to that broad definition, one individual can belong into more than one department. Then, a department is considered shocked if at least one individual in the department has been shocked, following the previous definitions of shocks. We only use our IC definition of shock for this analysis. For an individual i who has not been shocked yet and a department k, we have:

$$y_{i,t+n_j,k} = \beta \text{Shock}_{k,t} + \gamma_i + \delta_t + \epsilon_{i,t,k} \tag{2}$$

To better understand the dynamics before and after such shocks, we also conduct an event

study analysis:

$$y_{i,t+n_j} = \sum_{\tau=-5}^{5} \beta_\tau \mathbb{I}\{t - T_i = \tau\} + \gamma_i + \delta_t + \epsilon_{i,t} \tag{3}$$

The indicator $\mathbb{I}$ serves as a dummy around the first period an individual i is being shocked T_i. This allows us to observe the behavior of the dependent variables around the shock, 5 years prior and 5 years later. With an event analysis a relevant lag selection has a lesser importance, but helps move the window to obtain a frame that is interesting to observe.

VI Results

The findings from equation (1) are presented in Table 3. Notably, although the two shock definitions yield the same directional effects across all variables, their statistical significance differs. Consistent with previous literature, a positive funding shock appears to correlate with increased production. Specifically, advisors experience a rise in publication rate—approximately 0.09 additional publications per year—and an increase in the graduation rate of their advisees by about 0.02. With a higher budget, it is plausible that advisors are better equipped to undertake additional projects concurrently. However, the dynamics of team composition also change: the average team size per paper decreases by 0.29 individuals, which suggests that principal investigators (PIs) may be redistributing their available workforce across a greater number of projects, rather than merely expanding the size of individual projects. Concurrently, the overall lab size increases by roughly 1.22 individuals, indicating a broader search for diverse collaborative opportunities among researchers.

A critical concern associated with this production increase is the apparent decline in the quality of output. Both forward citations per paper and fractioned citations per paper drop—by 1.71 and 0.67 citations respectively. This decline suggests that while increased funding enables PIs to produce more publications, the attention each project receives may be diluted. One possible interpretation is that the elevated production pace leaves less time for refining each project, or alternatively, projects might be subdivided into smaller, less impactful units. Moreover, the reduction in the average fractioned publications per paper reinforces the idea that, despite larger lab teams, the per-project quality declines. Similarly, the use of new concepts appears to be declining after the doubling, to the benefit of older concepts. This is an indication of less innovative research. However, the doubling appears to increase the diversification of the citations, with references more narrow and forward citations spreading out to other fields.

Turning to Table 4, the effects on students are also noteworthy. Although students publish more and graduate at higher rates during periods of increased funding, the average impact of their publications—as measured by forward citations—drops by approximately 2.46 citations per paper. Same as their advisors, students formed during the doubling tend to use fewer new concepts and cite less diverse research. This downward trend in quality is further supported by other indicators:

both the percentage of students publishing and the percentage receiving funding decline by 4% and 0.9% respectively. These findings resonate with Tham et al. (2024), who show that funding gaps push researchers—especially early-career personnel—out of U.S. science and depress their earnings, and with Babina et al. (2023) who find that negative funding shocks reduce high-quality, basic publications. Our contribution is to show that even positive funding shocks, which expand labs and temporarily increase output, can have adverse consequences for the next generation of scientists trained in those labs.

One plausible explanation for the poorer student outcomes is that the increased workload imposed on advisors leaves less time for individualized mentorship. Alternatively, the decline in student outcomes reflects lower admission standards as PIs scrape closer to the bottom of the applicant barrel when expanding (Akcigit et al., 2024). These mechanisms have different implications. If student quality is the binding constraint, we should observe departmental spillovers: when a colleague in the same department receives a funding shock and expands, they should poach higher-quality applicants, leaving unshocked faculty with a worse pool. If instead diluted mentoring is the key mechanism, only direct exposure to a shocked advisor should matter.

Table 5 summarizes the results of this departmental analysis that results from equation (2). In general, the effects are statistically insignificant, which suggests that having another researcher from the same department receive a significant increase in funding does not lead to having students of worse quality. This outcome points to a scenario where the performance loss is mainly attributable to the dispersion of advisors' attention over more tasks, rather than an inherent drop in the quality of the students recruited due to competition for a scarce pool of applicants with high-ability. However, the limited size of the control group in this analysis—primarily due to the prevalence of funding increases in most departments—complicates the interpretation, particularly as departments loosely connected to the biomedical field (e.g., psychology) remain less affected, possibly introducing endogeneity issues. Expanding this analysis to include more departments within the same field could yield further insights.

A similar pattern emerges when we focus on the top 20 universities, as shown in Tables 6 and 7. Although the overall trends mirror those observed in the full sample, statistical significance diminishes—likely a result of the smaller sample size and the increased data requirements of our methodological approach.

Given the variability in lag structures depending on the dependent variable under consideration—and the absence of a universally accepted method to determine an optimal lag—it is standard practice to complement regression analyses with a panel event study. Such a study examines the effect on dependent variables over a continuous range of time surrounding the event, thereby providing a more nuanced view of whether the observed effects are attributable to a specific event or merely represent a trend.

Figure 4 and 5 depicts the event study results, from equation (3). For advisees, a moderate pre-event trend is evident; however, following the funding shock, there is a marked increase in advisee

count followed by a gradual reversion to baseline levels. A parallel pattern is observed for advisors' publication counts. In contrast, forward citations per PI paper initially show an inconsistent pattern but transition to a pronounced downward trend post-event, with no observable rebound. Mean team size and lab size display distinct trends—downward and upward respectively—with both metrics beginning to converge towards pre-shock values approximately five years after the funding event.

For student-specific metrics, the analysis is more challenging due to the smaller sample size and the inherent variability in graduation timelines, which complicates the estimation of an appropriate lag. Despite these limitations, a delayed downward trend is observable across the student variables, implying that our lag selection may be underestimating the time required for the full effect to manifest.

While we obtained statistically significant results across the board that fits our initial hypothesis, it is important to note that we face non negligible pretrend issues. Those issues suggest that our control and treatment group are fundamentally different researchers with different career paths and that using a staggered DiD can result into biased coefficients. The pretrend issue is stronger when it comes to the IC shock, suggesting that individuals getting shocked by different ICs do not follow the same dynamics. We are currently working to resolve this issue by adjusting our control and treatment groups and potentially adding weights or adjusting for the difference in slope.

VII Conclusion

The conventional wisdom that "the more, the better" does not always apply to research funding. Between 1998 and 2003, the National Institutes of Health (NIH) underwent a significant funding increase, an expansion that many anticipated would translate into improved research outcomes in the biomedical sciences—namely, higher-quality publications and stronger mentorship for emerging researchers. However, our analysis challenges this assumption.

By combining three large national dataset- ProQuest, Author-ity, and ExPorter - we assessed the impact of NIH funding increases through a staggered difference-in-differences model. This method allowed us to evaluate a broad spectrum of outcomes, including principal investigators' (PIs) publication productivity and citation impact, changes in team composition, and the training and early career trajectories of their mentees. Unlike prior studies that have predominantly concentrated on the direct effects of funding surges on PI output, our research uniquely examines how these fiscal shocks influence the formation and subsequent development of PhD students who are being trained during periods of expansive funding.

Our results reveal that, although an increase in funding corresponds with higher publication counts and a greater number of graduations, these gains come at a cost. Specifically, the publications produced in this environment tend to attract fewer citations, suggesting a compromise in research impact or visibility. Moreover, the early career prospects for new researchers suffer. PhD

graduates who trained under PIs benefitting from system-wide funding surges tend to face a more competitive and saturated job market, which subsequently dampens their research productivity. Even when funding shocks occur on an idiosyncratic basis—without generating broad market-wide pressures—the negative effect on post-graduation productivity persists, potentially due to factors such as reduced mentorship quality.

These findings have broader implications for the well-documented aging of the biomedical workforce and declining opportunities for early-career investigators (Larson et al., 2012; Teitelbaum, 2008). Our model suggests that funding boom-bust cycles contribute to these trends through a dynamic mechanism: temporary expansions create large trainee cohorts who cannot be fully absorbed into PI positions when funding stabilizes, while PIs hired during the expansion remain in the system. This bottleneck persists for decades, as illustrated by the continued challenges facing scientists trained during the doubling era. While other factors—including elimination of mandatory retirement and potential peer review capture—likely contribute to workforce aging, our analysis shows that funding volatility itself generates lasting distortions in the age structure and career opportunities in science.

Our results are subject to an important limitation: they rely on comparisons across researchers who may already be on different career trajectories prior to experiencing a funding shock. Aggregating such individuals can introduce bias if treatment timing is correlated with underlying momentum in research output, potentially affecting our estimates and interpretation. We are currently addressing this concern by refining the composition of treatment and comparison groups and by incorporating adjustments that account for differential pre-shock trends.

In summary, our findings indicate that funding volatility—whether experienced across the system or at the level of individual PIs—can have pronounced adverse effects on the training and career development of early-career researchers. Our results complement recent evidence that funding interruptions and negative shocks can permanently alter who stays in science and what kind of science is produced (Babina et al.; Tham et al.), by showing that large, temporary increases in funding can also degrade training quality and early-career outcomes. Our results underscore the potential benefits of adopting measures that favor more stable and predictable funding trajectories. Reducing extreme fluctuations in financial resources could help mitigate disruptions, thereby supporting both the quality of research outputs and the professional growth of the next generation of scientists.

References

Akcigit, Ufuk, Jeremy Pearce, and Marta Prato, "Tapping into Talent: Coupling Education and Innovation Policies for Economic Growth," *The Review of Economic Studies*, 2024.

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

Bol, Thijs, Mathijs De Vaan, and Arnout van de Rijt, "The Matthew effect in science funding," *Proceedings of the National Academy of Sciences*, 2018, *115* (19), 4887–4890.

Dobson, Sarah, "How to Write Multiple Grants Simultaneously," Sarah Dobson & Co. Blog 2025. Posted May 15, 2025. Available at `https://www.sarahdobson.co/sarah-dobson-tips/how-to-write-multiple-grants-simultaneously`.

Drugmonkey, "NIH Will Only Accept Six Applications per PI per Calendar Year," Drugmonkey Blog 2025. Posted July 17, 2025. Available at `https://drugmonkey.wordpress.com/2025/07/17/nih-will-only-accept-six-applications-per-pi-per-calendar-year/`.

Freeman, Richard B and John Van Reenen, "Be careful what you wish for: A cautionary tale about budget doubling," *Issues in Science and Technology*, 2008, *25* (1), 27.

Freeman, Richard B. and John Van Reenen, "What If Congress Doubled R&D Spending on the Physical Sciences?," in Josh Lerner and Scott Stern, eds., *Innovation Policy and the Economy, Volume 9*, Chicago: University of Chicago Press, 2009, pp. 1–38.

Goolsbee, Austan, "Does government R&D policy mainly benefit scientists and engineers?," 1998.

Larson, Richard C, Navid Ghaffarzadegan, and Mauricio Gomez Diaz, "Magnified effects of changes in NIH research funding levels," *Service Science*, 2012, *4* (4), 382–395.

Sattari, Reza, Jung Bae, Enrico Berkes, and Bruce A Weinberg, "The ripple effects of funding on researchers and output," *Science Advances*, 2022, *8* (16), eabb7348.

Teitelbaum, Michael S, "Structural disequilibria in biomedical research," *Science*, 2008, *321* (5889), 644–645.

Tham, Wei Yang, Joseph Staudt, Elisabeth Ruth Perlman, and Stephanie D. Cheng, "Scientific Talent Leaks Out of Funding Gaps," CES Working Paper 24-08, U.S. Census Bureau, Center for Economic Studies, Washington, DC February 2024.

Torvik, Vetle I. and Neil R. Smalheiser, "Author Name Disambiguation in MEDLINE," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2009, *3* (3), 11.

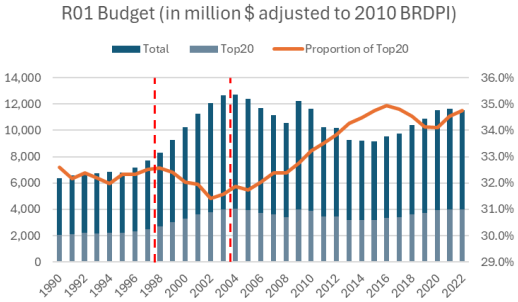
_ , **Marc Weeber, Don R. Swanson, and Neil R. Smalheiser**, "A Probablistic Similarity Metric for Medline Records: A Model for Author Name Disambiguation," *Journal of the American Society for Information Science and Technology*, 2005, *56* (2), 140–158.

Webel, Allison, "What the New NIH Grant Submission Cap Means for Investigators," Notes on Innovation and Research, University of Washington School of Nursing 2025. Blog post, accessed November 2025. Available at `https://nursing.uw.edu/what-the-new-nih-grant-submission-cap-means-for-investigators/`.

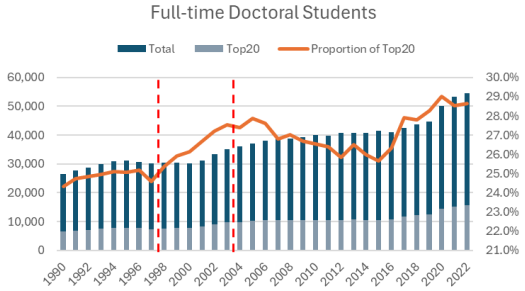
Yu, Huifeng, Gerald Marschke, Matthew B Ross, Joseph Staudt, and Bruce A Weinberg, "Publish or perish: Selective attrition as a unifying explanation for patterns in innovation over the career," *Journal of Human Resources*, 2023, *58* (4), 1307–1346.

Figure 1: Budget and Labor Supply

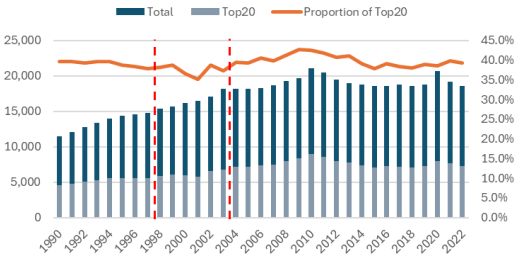
(a)



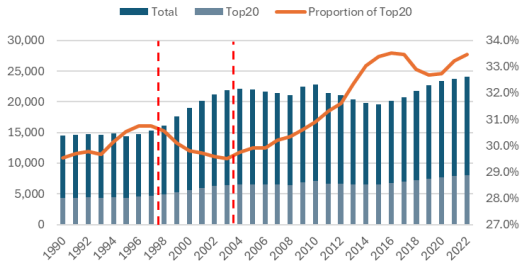
(b)



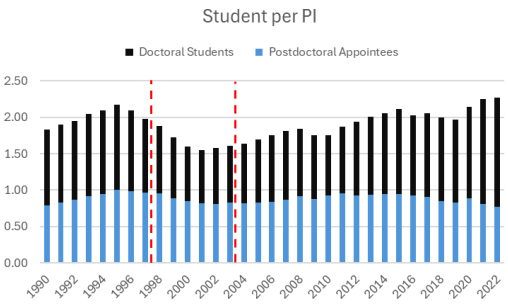
(c)



(d)



(e)



(f)

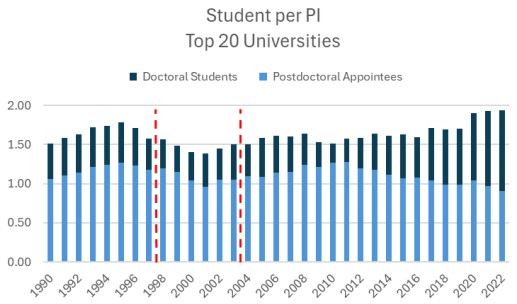


Figure 2: Data Combination Flow Chart

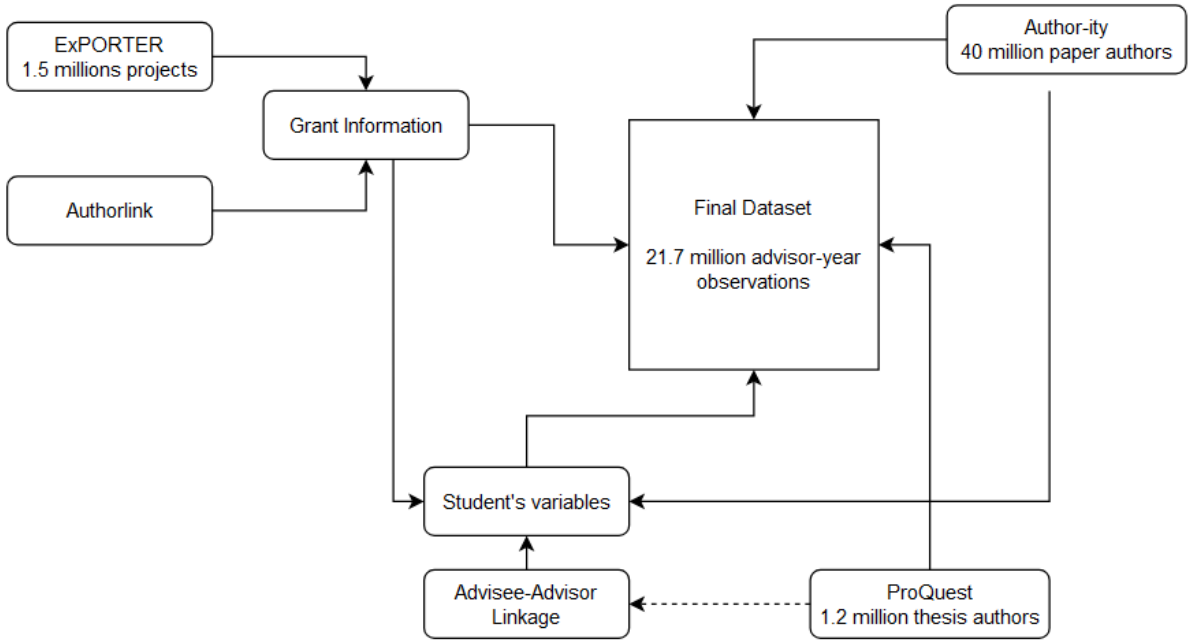
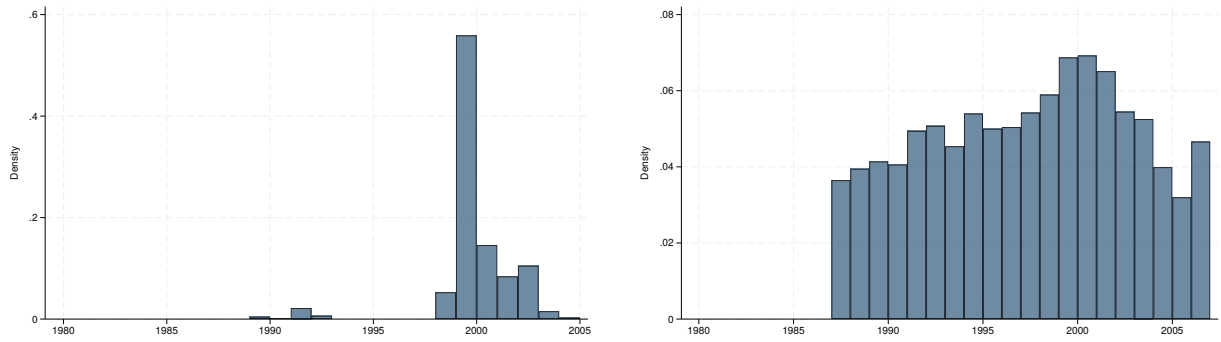


Figure 3: Distribution of shocks



(a) NIH Shocks

(b) Individual Shocks

Table 1: Descriptive Statistics

Panel A: PI-level Variables	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
Advisee Count	1,135,640	0.288	0.645	0	25
Publication Count	1,135,640	2.018	2.993	0	142
Fractioned Publications	1,135,640	0.567	0.835	0	21.9996
Fractioned Publications per Paper	707,798	0.300	0.162	0.001	1
Forward Citations	707,798	83.140	180.440	1	11,199
Forward Citations per Paper	707,798	21.161	35.761	1	5,157
Fractioned Forward Citations	707,798	20.516	45.886	0.002	3,518.617
Fractioned Forward Citations per Paper	707,798	5.535	9.231	0.002	820.167
Maximum Forward Citations	707,798	40.594	88.604	1	10,967
New Concepts	615,941	1.405	3.580	0	113
Average New Concepts	615,941	0.363	0.797	0	18
Old Concepts	615,941	56.893	69.427	0	1,721
Average Old Concepts	615,941	16.074	10.199	0	98
Diversity of References	601,074	951.286	32.482	0	987.578
Diversity of Citations	611,441	951.840	26.037	0	989.196
Mean Team Size	707,798	5.861	22.584	1	743
Lab Size	707,798	12.985	30.346	1	1,340
Panel B: Student-level Variables					
Publication Count	245,759	3.031	5.762	0	295
Publications per Student	245,759	2.388	4.463	0	284
Fractioned Publications	245,759	0.757	1.360	0	56.002
Fractioned Publications per Student	245,759	0.599	1.068	0	56.002
Forward Citations	139,037	134.063	283.986	1	10,369
Forward Citations per Paper per Student	139,037	17.604	26.874	0.050	1,789
Fractioned Forward Citations	245,759	15.966	42.871	0	2,536.402
Fractioned Forward Citations per Student	245,759	2.241	4.502	0	397.5
Maximum Citations	139,037	48.744	104.777	1	7,261
New Concepts	127,631	2.156	5.283	0	146
Average New Concepts	127,631	0.336	0.667	0	16
Old Concepts	127,631	96.902	139.682	0	5,732
Average Old Concepts	127,631	17.064	10.318	0	94
Diversity of References	123,939	951.265	34.645	0	986.305
Diversity of Citations	126,284	949.814	28.030	0	986.361
Amount of Publishing Students	245,759	0.679	0.708	0	8
Fraction of Publishing Students	245,759	0.530	0.482	0	1
Amount of Funding from Students	245,759	6,697.535	127,454.1	0	13,600,000
Fraction of Students Receiving Funding	245,759	0.006	0.070	0	1

Table 2: Variables

Index	**Variable Name**	**Description**	**Formula**
A	Advisee Count	Total number of students graduating per PI each year, measured by counting the number of doctoral dissertations (Masters excluded) for which the PI is listed as advisor.	
B	Publication Count	Total number of publications per PI per year for which the PI appears as any author.	
C	Fractioned Publications	Total amount of fractioned publications per PI per year. A fraction publication is obtained by dividing 1 by the number of authors for each paper the PI has appeared as an author. (Example: A PI who publishes two papers, one with 4 authors and another with 2 authors, will have $\frac{1}{4} + \frac{1}{2} = 0.75$).	
D	Fractioned Publications Per Paper	The average fractioned publications per paper for each PI and each year.	C/B
E	Forward Citations	Total amount of forward citations per PI per year. To avoid recency bias, forward citations are registered 5 years after publication.	
F	Forward Citations Per Paper	The average forward citations per paper published for each PI and each year.	E/B
G	Fractioned Citations	Total amount of fractioned citations per PI per year. A fraction citation is obtained by dividing the number of forward citations (5 years after publication) by the number of authors for each paper the PI has appeared as an author. (Example: A paper with 10 citations and 4 authors contributes $\frac{10}{4} = 2.5$).	
H	Fractioned Citations Per Paper	The average fractioned forward citations per paper published for each PI and each year.	G/B
I	New Concepts	The number of new concepts used by a given paper using text analysis. This variable is then aggregated for each advisor-year.	
K	Average New Concepts	The number of new concepts averaged across the number of paper published for each advisor-year.	I/B

J	Old Concepts	The number of old concepts used by a given paper using text analysis. This variable is then aggregated for each advisor-year.	
L	Average Old Concepts	The number of old concepts averaged across the number of paper published for each advisor-year.	J/B
M	Diversity of References	A metric to measure how diverse the papers being cited are for a given paper. The metric is then averaged across publications for each advisor-year.	
N	Diversity of Citations	A metric to measure how diverse the forward citations are for a given paper. The metric is then averaged across publications for each advisor-year.	
O	Mean Team Size	The average of the number of authors over the number of publications per PI for a given year.	
P	Lab Size	Total number of unique collaborators per PI per year measured by using common authorship of publications.	
Q	Amount of Publishing Students	The number of students who are counted as graduating under a PI (per A) who also publishes within 5 years after graduation. Observations are made at the PI-year level (the year being the year of graduation, not publication).	
R	Fraction of Publishing Students	The fraction of publishing students per PI per year.	Q/A
S	Amount of Funding from Students	Total amount of funding (in adjusted \$) received by students graduating under a PI (per A) within 5 years of graduation. Observations are made at the PI-year level (the year being the year of graduation, not publication).	
T	Fraction of Students Receiving Funding	The fraction of students who receive funding within 5 years of graduation.	S/A
V	Student's Forward Citations Per Student Per Paper	The average of forward citations for students who publish within 5 years of graduation. Observations are at the PI-year level and averaged over publications and students.	

Table 3: Regression Results for PI variables

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Advisee Count	0.023***	1,135,640	0.065***	1,135,769
	(0.006)		(0.005)	
Publication Count	0.094***	1,135,640	0.487***	1,135,769
	(0.026)		(0.025)	
Fractioned Publications per Paper	-0.003**	707,798	-0.005***	707,892
	(0.001)		(0.002)	
Forward Citations per Paper	-1.712***	707,798	-2.808***	707,892
	(0.333)		(0.410)	
Fractioned Forward Citations per Paper	-0.673***	707,798	-0.893***	707,892
	(0.098)		(0.113)	
Maximum Forward Citations	0.198	707,798	1.465	707,798
	(0.972)		(0.916)	
New Concepts	-0.497***	615,941	-0.054	615,995
	(0.053)		(0.049)	
Average New Concepts	-0.099***	615,941	-0.112***	615,995
	(0.010)		(0.011)	
Old Concepts	4.208***	615,941	12.938***	615,995
	(0.798)		(0.752)	
Average Old Concepts	-0.152*	615,941	0.007	615,995
	(0.091)		(0.090)	
Diversity of References	-1.628***	601,074	-1.446***	601,128
	(0.178)		(0.170)	
Diversity of Citations	0.845***	611,441	-0.030	611,495
	(0.154)		(0.145)	

Continued on next page

Regression Results for PI variables (continued)

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Mean Team Size	-0.286***	707,798	-0.136***	707,892
	(0.053)		(0.035)	
Lab Size	1.220***	707,798	2.221***	707,892
	(0.175)		(0.158)	
Individuals in Control Group		49,134		45,999
Individuals in Treatment Group		12,914		16,058

Regression Results: Contains the ATT estimates of the stacked difference-in-difference for our dependent variables of choice regarding PIs. Column (1) uses our IC definition of shock, i.e., the control group contains individuals that either receive no funding or receive funding from IC when they are not being shocked. Column (2) uses our definition of shock at the individual level, among the productive researchers of 20 years or more of experience. A shock is defined as a growth of 80% from its moving average of 3 years. Only PIs with at least one publication during a year are counted for variables related to publications, explaining the decline in the number of observations. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Regression Results for Student variables

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Publications per Student	-0.695***	245,759	-0.316***	245,759
	(0.114)		(0.099)	
Fractioned Publications per Student	-0.190***	245,759	-0.119***	245,759
	(0.028)		(0.027)	
Forward Citations per Paper per Student	-2.464***	139,037	-0.712	139,037
	(0.773)		(0.736)	
Fractioned Forward Citations per Student	-0.795***	245,759	-0.372***	245,759
	(0.151)		(0.137)	
Maximum Citations	-6.837*	139,037	-2.884	139,037
	(3.962)		(5.275)	
New Concepts	-0.641***	127,631	-0.193	127,631
	(0.170)		(0.214)	
Average New Concepts	-0.058***	127,631	-0.043*	127,631
	(0.019)		(0.024)	
Old Concepts	-9.841**	127,631	1.765	127,631
	(4.622)		(4.341)	
Average Old Concepts	-0.422*	127,631	0.376	127,631
	(0.216)		(0.242)	
Diversity of References	-2.513***	123,939	-0.736	123,939
	(0.829)		(0.996)	
Diversity of Citations	0.682	126,284	0.559	126,284
	(0.604)		(0.452)	

Continued on next page

Regression Results for Student variables (continued)

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Amount of Publishing Students	-0.040**	245,759	-0.001	245,759
	(0.019)		(0.020)	
Fraction of Publishing Students	-0.040***	245,759	-0.030***	245,759
	(0.008)		(0.010)	
Amount of Funding from Students	-12520**	245,759	-9201*	245,759
	(5378)		(5501)	
Fraction of Students Receiving Funding	-0.009***	245,759	-0.007**	245,759
	(0.003)		(0.003)	
Individuals in Control Group		49,134		45,999
Individuals in Treatment Group		12,914		16,058

Regression Results: Contains the ATT estimates of the stacked difference-in-difference for our dependent variables of choice regarding students. A student's production is counted within 5 years of graduation and is accounted at the PI level the year of the student's graduation. Column (1) uses our IC definition of shock, i.e., the control group contains PIs that either receive no funding or receive funding from IC when they are not being shocked. Column (2) uses our definition of shock at the PI level, among the productive researchers of 20 years or more of experience. A shock is defined as a growth of 80% from its moving average of 3 years. Only PIs with at least one student during a year are counted for variables related to publications, explaining the decline in the number of observations. Only PIs with at least 1 student who published within 5 years of graduation are counted for the average of student's quality. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Departmental Shock Effects on Students

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Publications per Student	0.063	207,253	-0.015	195,983
	(0.080)		(0.024)	
Fractioned Publications per Student	-2.932*	207,253	-0.310**	195,983
	(1.595)		(0.153)	
Forward Citations per Paper per Student	-5.690	108,394	-0.240	100,330
	(4.022)		(0.244)	
Fractioned Forward Citations per Student	-0.019	207,253	14.157*	195,983
	(0.034)		(7.557)	
Maximum Citations	-0.745**	108,394	-4.925	100,330
	(0.371)		(4.104)	
New Concepts	0.570	98,388	0.032	90,693
	(1.763)		(0.019)	
Average New Concepts	0.012	98,388	1697.510	90,693
	(0.012)		(2781.843)	
Old Concepts	0.002	98,388	0.0016	90,693
	(0.002)		(0.002)	
Average Old Concepts	0.063	98,388	-0.015	90,693
	(0.080)		(0.024)	
Diversity of References	-2.932*	94,756	-0.310**	87,080
	(1.595)		(0.153)	
Diversity of Citations	-5.690	97,070	-0.240	89,382
	(4.022)		(0.244)	
Amount of Publishing Students	-0.005	207,253	0.032	195,983
	(0.013)		(0.019)	
Fraction of Publishing Students	-0.011	207,253	0.012	195,983
	(0.008)		(0.012)	
Amount of Funding from Students	-9351*	207,253	1698	195,983
	(5173)		(2789)	
Fraction of Students Receiving Funding	-0.004***	207,253	0.002	195,983
	(0.001)		(0.002)	

Continued on next page

Departmental Shock Effects on Students (continued)

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Individuals in Control Group		14,572		10,335
Individuals in Treatment Group		44,892		49,742

Table contains the ATT estimates of the stacked difference-in-difference for our dependent variables on student outcomes. A department is considered shocked if at least one individual during the year experienced a shock. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Regression Results for PI variables, Top 20 Universities

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Advisee Count	-0.005	213,501	0.084***	213,501
	(0.013)		(0.009)	
Publication Count	0.113**	213,501	0.680***	213,501
	(0.055)		(0.057)	
Fractioned Publications per Paper	-0.004	150,302	-0.002	150,302
	(0.003)		(0.003)	
Forward Citations per Paper	-1.987**	150,302	-2.917***	150,302
	(0.797)		(1.022)	
Fractioned Forward Citations per Paper	-0.613***	150,302	-0.717***	150,302
	(0.221)		(0.267)	
Maximum Forward Citations	-1.662	150,302	5.342**	150,302
	(2.180)		(2.258)	
New Concepts	-0.602***	133,214	0.198*	133,214
	(0.119)		(0.120)	
Average New Concepts	-0.093***	133,214	-0.106***	133,214
	(0.018)		(0.021)	
Old Concepts	5.267***	133,214	15.994***	133,214
	(1.728)		(1.682)	
Average Old Concepts	-0.109	133,214	-0.055	133,214
	(0.174)		(0.166)	
Diversity of References	-1.438***	130,862	-0.930***	130,862
	(0.346)		(0.281)	
Diversity of Citations	0.677**	132,491	0.461*	132,491
	(0.287)		(0.235)	

Continued on next page

Regression Results for PI variables, Top 20 Universities (continued)

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Mean Team Size	-0.473***	150,302	-0.329***	150,302
	(0.176)		(0.102)	
Lab Size	1.031**	150,302	2.342***	150,302
	(0.411)		(0.387)	
Individuals in Control Group		7,059		6,607
Individuals in Treatment Group		4,138		4,591

Same as Table 3, but only keeping PIs from universities in our list of most productive establishments. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Regression Results for Student variables, Top 20 Universities

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Publications per Student	-0.606***	49,321	-0.293***	49,321
	(0.137)		(0.106)	
Fractioned Publications per Student	-0.161***	49,321	-0.085***	49,321
	(0.036)		(0.032)	
Forward Citations per Paper per Student	-1.016***	31,447	-0.612	31,447
	(0.383)		(0.471)	
Fractioned Forward Citations per Student	-0.476***	49,321	-0.259**	49,321
	(0.104)		(0.112)	
Maximum Citations	-4.286	31,447	-1.918	31,447
	(2.794)		(3.025)	
New Concepts	-0.389***	29,049	-0.127	29,049
	(0.096)		(0.112)	
Average New Concepts	-0.048***	29,049	-0.031*	29,049
	(0.013)		(0.018)	
Old Concepts	-5.927**	29,049	0.894	29,049
	(2.763)		(2.511)	
Average Old Concepts	-0.311*	29,049	0.162	29,049
	(0.172)		(0.199)	
Diversity of References	-1.952***	28,317	-0.578	28,317
	(0.573)		(0.652)	
Diversity of Citations	0.454	28,783	0.283	28,783
	(0.398)		(0.327)	

Continued on next page

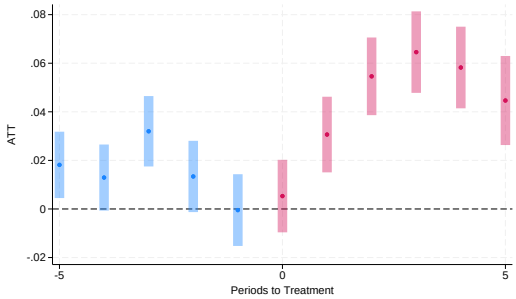
Regression Results for Student variables, Top 20 Universities (continued)

	(1) IC		(2) Individual	
Variables	Coeff. (S.E.)	N° obs.	Coeff. (S.E.)	N° obs.
Amount of Publishing Students	-0.086**	49,321	0.026	49,321
	(0.036)		(0.041)	
Fraction of Publishing Students	-0.042***	49,321	-0.003	49,321
	(0.015)		(0.020)	
Amount of Funding from Students	-14630	49,321	-15632	49,321
	(11614)		(14668)	
Fraction of Students Receiving Funding	-0.002	49,321	-0.004	49,321
	(0.005)		(0.006)	
Individuals in Control Group		7,059		6,607
Individuals in Treatment Group		4,138		4,591

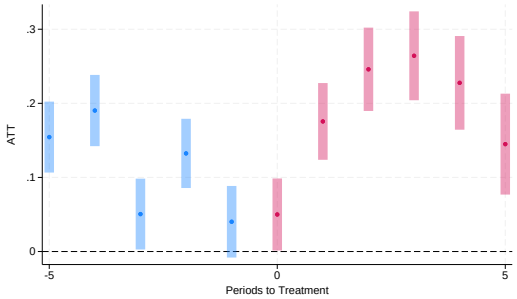
Same as Table 4, but only keeping PIs from universities in our list of most productive establishments. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 4: Panel Event Study

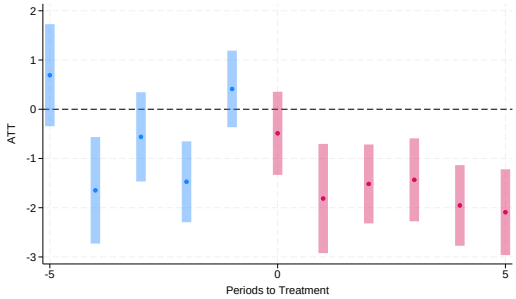
(a) Advisee



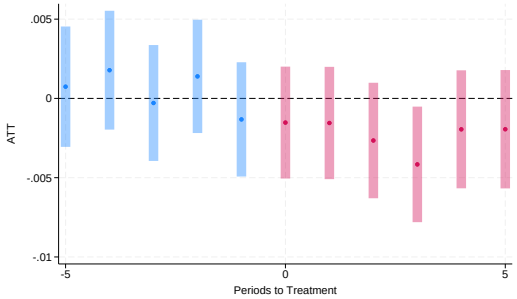
(b) Annual Publication Count



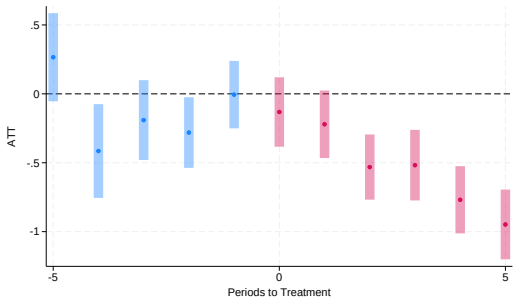
(c) Forward Citations per Paper



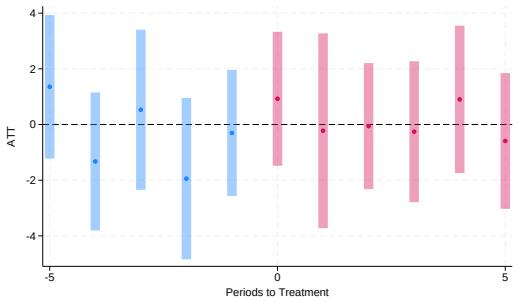
(d) Fractioned Publications per Paper



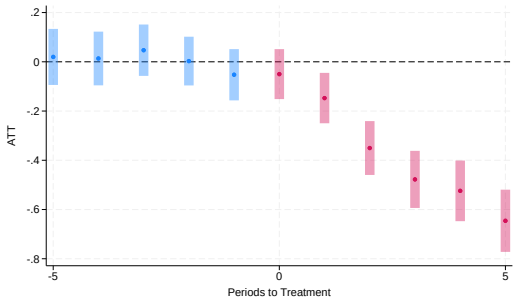
(e) Fractioned Citations per Paper



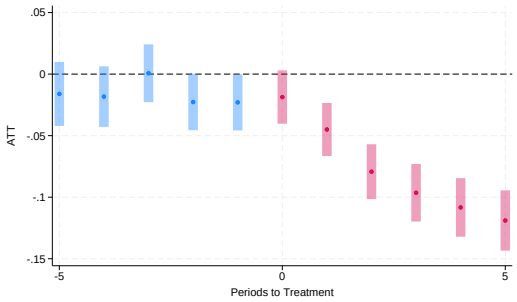
(f) Maximum Citations



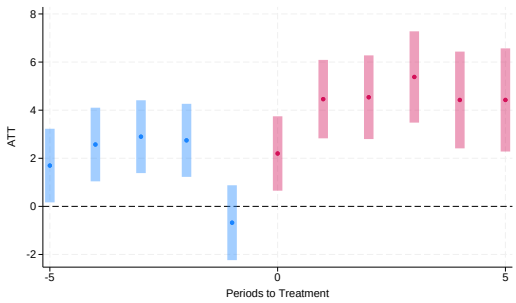
(g) New Concepts



(h) Average New Concepts



(i) Old Concepts



(j) Average Old Concepts

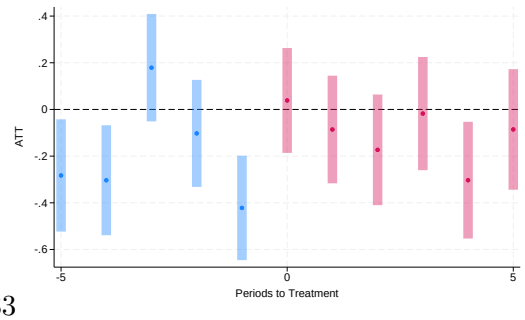
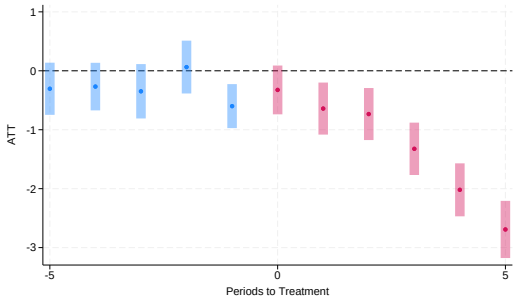
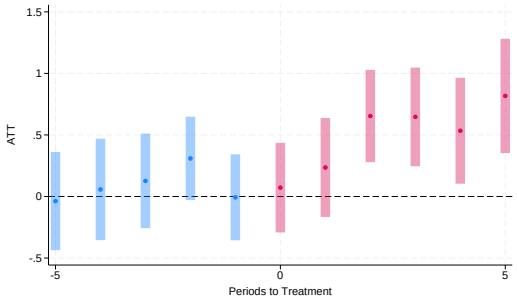


Figure 4: Panel Event Study (cont.)

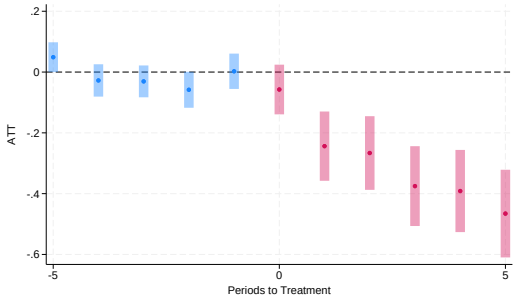
(k) Diversity of References



(l) Diversity of Citations



(m) Mean Team Size



(n) Lab Size

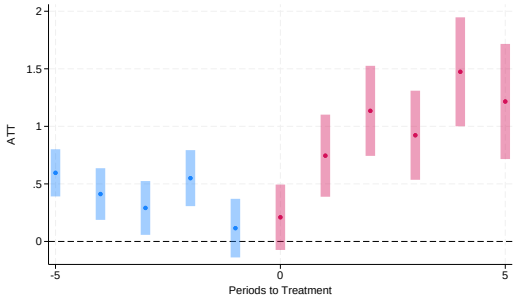
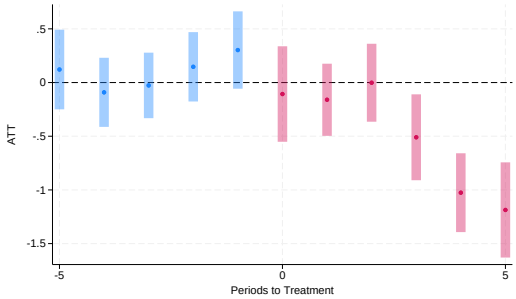
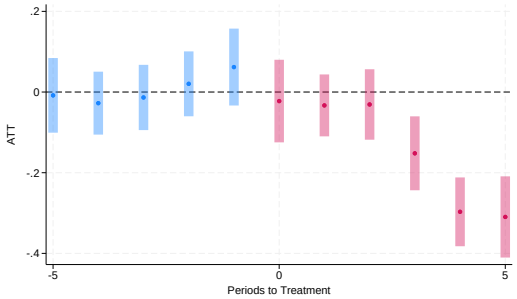


Figure 5: Panel Event Study (Students' Variables)

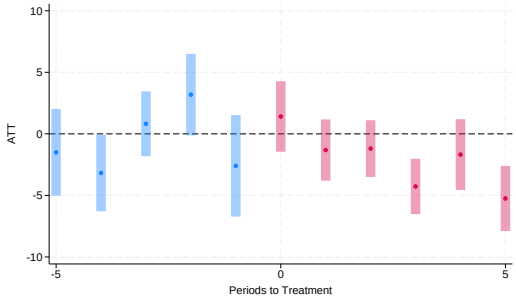
(a) Publication per Student



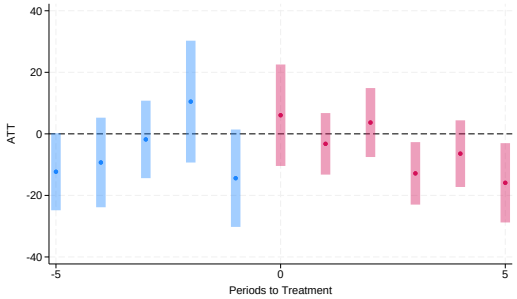
(b) Fractioned Publications per Student



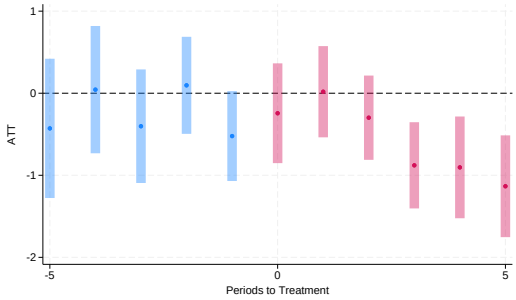
(c) Forward Citations per Paper per Student



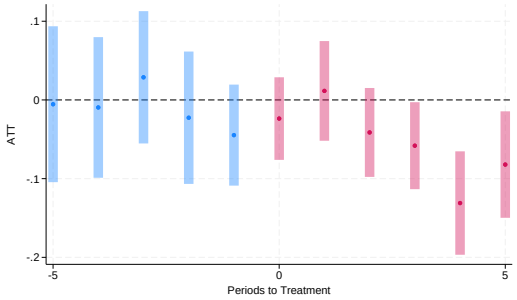
(d) Maximum Citations



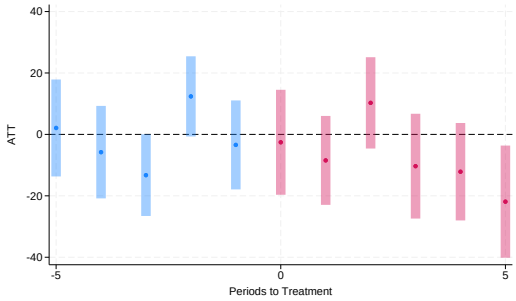
(e) New Concepts



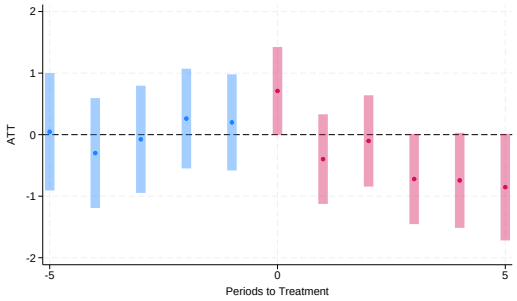
(f) Average New Concepts



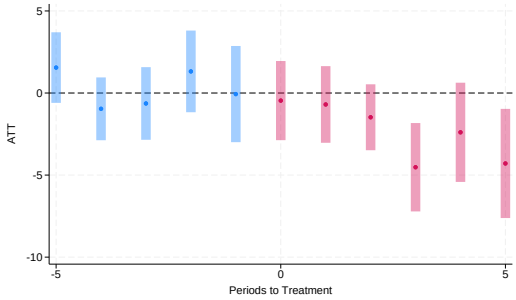
(g) Old Concepts



(h) Average Old Concepts



(i) Diversity of References



(j) Diversity of Citations

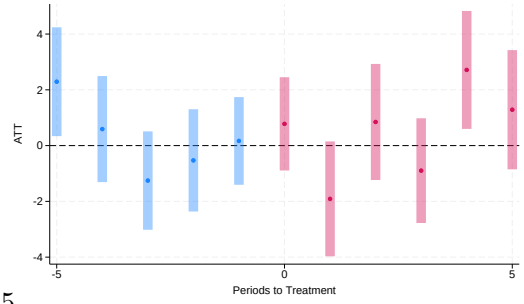
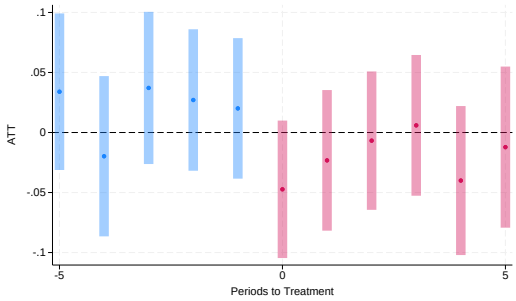
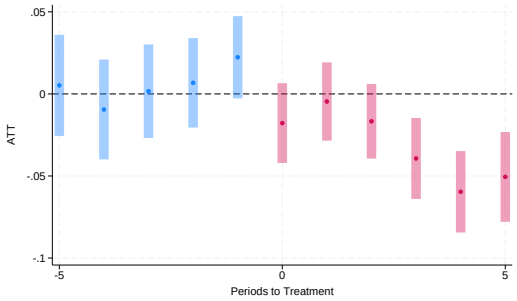


Figure 5: Panel Event Study (Students' Variables, cont.)

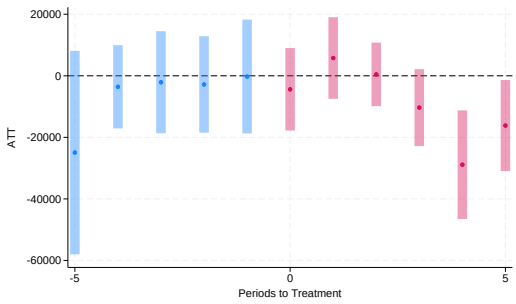
(k) Amount of Publishing Students



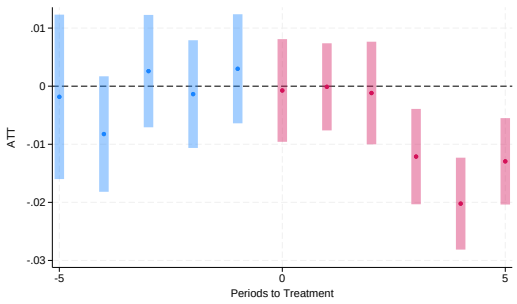
(l) Fraction of Publishing Students



(m) Amount of Funding from Students



(n) Fraction of Students Receiving Funding



A Dataset Construction

Throughout our data, the main individual identification is called "auid". The authorlink dataset is a convenient disambiguated dataset that combines project number (for grants), other project information, and auid. We first only keep R01-equivalent grants (R01, DP1, DP2, DP5, R37, R56, RF1, RL1, U01, R35), and then merge the file with an NIH ExPORTER file containing data about NIH grants and their PIs from 1985 to 2008, creating a link between auid and grant received.

Simultaneously, we use a disambiguated file from ProQuest which links auid with thesis information (advisor name, advisee name, degree, school, year of graduation). We drop master's degrees, and individuals with too many graduations (caused by disambiguation errors) and obtain a graduation date for individuals we have information on their doctorate thesis. For those we do not, we estimate a graduation date 16 years prior to their first advisee. The graduation year for individuals with more than one thesis is their second graduation, as long as this one is no more than 4 years after the first.

Additionally, we use an Author-ity file that contains data on all authorship and citations for papers published in the biomedical fields. We construct an indicator that only look at forward citations five years after publication to prevent recent papers from having less citations. We create a lab size variable that prevent the same author to be counted more than once as a collaborator for each author and each year. We also construct a mean size variable by averaging the total authors (here duplicates are counted) by the number of papers, for each year. We combine these data with the metrics obtained by Yu et al. (2023), namely their variables about the use of new concepts in a paper, old concept, and diversity of both references and citations. Yu's data does not contain information on all the authors that our dataset possesses and therefore lack of information will be treated as missing values as opposed to bull values.

Student's variables are obtained by first creating a linkage between advisor's auid and advisee's auid. Combining it with the three previous dataset we can obtain information on students such as grants received, paper published, and forward citations. For students with an ambiguous graduation date (more than one graduation), we take the average. This way we lose precision but keep observations.

We then combine all 4 datasets together, linking them through auid and years. Finally, we expand the dataset to obtain "0s" – for missing years therefore production is equal to 0 – discard observations due to inactivity both at the beginning and at the end of researchers' career. Then the dataset is merged with an indicator of shock, depending on the definition being used.

For the top 20 universities, we follow the same steps but only keeping observations from ProQuest where the university is in our top 20.

B Theoretical Framework

Researchers work (as researchers) a maximum of two periods. In each period, t, a new generation of young researchers (or students) begin their research career. The young researchers conduct research in period t and at t's end either continue on as an old researcher in $t+1$, or attrit to a non-research job. At the end of period $t+1$, all researchers retire. Because new researchers are "born" in each period, there is an overlap between old and young researchers. Research output is comprised of research papers, with each young researcher writing a single paper with one old researcher. The number of papers each old researcher writes is the ratio of young researchers to old researchers.

Science is funded by an external funding agency with the agency allocating its budget in each period between young and old researchers. If an individual receives 1 unit of funding, he works as a researcher. Otherwise, he takes a non-researcher job. The agency's budget in each period t is B_t. Because each researcher requires 1 unit of funding, in period t, there are $B_{y,t}$ young researchers and $B_{o,t}$ old ones, such that $B_{o,t} + B_{y,t} \leq B_t$. The young researchers brought in are useful to the scientific enterprise because they pair with old researchers to write papers, but also, they (potentially) become next period's old researchers who will lead research projects.

Each old scientist has a unit of effort to allocate to his papers. Thus, if there are $B_{y,t}$ young researchers and $B_{o,t}$ old researchers, there are $B_{y,t}$ papers (one for each young researcher) with each old researcher writing $B_{y,t}/B_{o,t}$ papers and devoting $\tau_t = B_{o,t}/B_{y,t}$ units of effort per paper, which determines the quality of articles. Effort is introduced here so that the funding agency has a reason to support old researchers.

Moreover, for each period t, each student's productivity is $g(\tau_t)$ and each professor's is $f(\tau_{t-1}, \tau_t)$ with $g(.)$ an increasing and concave function and $f(.,.)$ a negative semi-definite function. Indeed, the productivity of students depend on how much time their advisor can allocate for them, while the productivity of professors depends on the skills they earned while being students (hence τ_{t-1}) as well as how much time they can currently allocate to their research (τ_t).

Then, the total research output that can be produced becomes:

$$F(B_{o,t}, B_{y,t}) = B_{y,t}\big[f(\tau_{t-1}, \tau_t) + g(\tau_t)\big], \tag{B.1}$$

B.1 Unconstrained Optimization

Let $V(\tau_{t-1})$ denote the value function associated with an inherited mentor ratio τ_{t-1}. At each date, the young researcher (or planner) chooses τ_t to maximize the discounted sequence of aggregate outputs:

$$V(\tau_{t-1}) = \max_{\tau_t} \big[B_{y,t}\big(f(\tau_{t-1}, \tau_t) + g(\tau_t)\big) + \beta V(\tau_t)\big], \qquad 0 < \beta < 1. \tag{B.2}$$

The control variable is τ_t, and the state variable is τ_{t-1}.

Differentiating (B.2) with respect to τ_t gives the Euler condition:

$$B_t((f_2 + g')(1 + \tau_t) - f - g) + \beta(1 + \tau_t)^2 V'(\tau_t) = 0, \tag{B.3}$$

where f_1 denotes the partial derivative of f with respect to its first argument. The envelope condition is

$$V'(\tau_{t-1}) = f_1 \frac{B_t}{1 + \tau_t}. \tag{B.4}$$

Shifting one period forward and eliminating $V'(\tau_t)$ between (B.3) and (B.4) yields the Euler equation:

$$(f_2 + g')(1 + \tau_t) - f - g + \beta(1 + \tau_t)^2 f_{1,t+1} \frac{1}{1 + \tau_{t+1}} \frac{B_{t+1}}{B_t} = 0 \tag{B.5}$$

which links the mentor productivity ratios across three consecutive generations.

A steady state $\tau_t = \tau_{t-1} = \tau_{t+1} = \tau^\star$ satisfies

$$((f_2^\star + g'^\star)(1 + \tau^\star) - f^\star - g^\star) + \beta(1 + \tau^\star) f_1^\star = 0. \tag{B.6}$$

To analyze the dynamics we can write $x_t = \tau_{t-1}$ and $y_t = \tau_t$ so that:

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{\beta} & a \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix}, \quad a = \frac{(1 + y) f_{yy} + g'' + \beta(f_{xx}(1 + y) + 2 f_x)}{\beta(f_x - f_{xy}(1 + y))} \tag{B.7}$$

The characteristic polynomial of J is

$$\lambda^2 - a\lambda + \frac{1}{\beta} = 0.$$

Since $\lambda_1 \lambda_2 = \frac{1}{\beta} \geq 1$ under standard discounting, one eigenvalue lies inside and the other outside the unit circle, implying *saddle-path stability.*

The system admits one stable and one unstable direction. Along the stable direction, deviations from the steady state decay geometrically, while along the unstable direction they explode. If the initial condition does not lie exactly on the stable manifold, the trajectory will contain a component along the unstable eigenvector and will therefore diverge over time — that is, the variable τ_t "blows up." Hence, only points on the stable path converge to the steady state; all others eventually depart from it.

B.2 Constrained Optimization

However τ_t cannot be rising indefinitely with a constant budget. Indeed, there will come a point when the stock of students will be low enough that there will not be enough stock to hire the same amount of professors. Hence, the following constraint becomes binding:

$$B_{o,t} \leq B_{y,t-1} \tag{B.8}$$

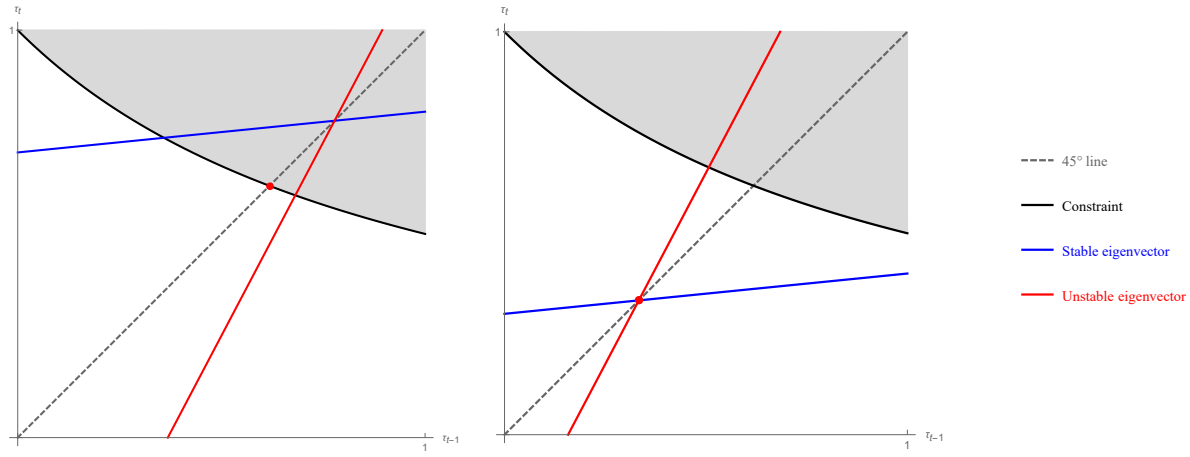
We can rewrite the constraint as:

$$\tau_t \leq \frac{B_{t-1}}{B_t} \frac{1}{1+\tau_{t-1}} \tag{B.9}$$

Constant Budget:

From inequality (B.9), we see that a constant budget creates a simple downward and concave curve that can be represented on the $\{\tau_t, \tau_{t-1}\}$ plane as shown on Figure B.1. There is two possibilities: the unconstrained steady state can either be in or out of the feasible area created by the constraint. If the steady state is unfeasible, then starting on the stable path (or above) will lead to the constraint. Once we are on the curve, we stay on it and the new steady state becomes the intersection between the 45° line and the curve. If the unconstrained steady state is feasible, then starting on the stable path still leads to the unconstrained steady state, and the constraint will not be binding.

Figure B.1: Dynamics of the Quality of Papers



(a) Unfeasible Equilibrium (b) Feasible Equilibrium

Varying Budget:

For an increasing budget, the ratio $\frac{B_{t-1}}{B_t}$ is decreasing leading to the constraint curve shifting to the left. Vice versa, for a decreasing budget the curve shifts to the right and when the budget stop growing it goes back in the middle. The two different cases illustrated in Figure B.1 become even more important. If the constraint was not binding in the first place (B.1.b) and we observe a relatively small budget increase then constraint non binding, therefore no change is observed. However, if the constraint is binding (B.1.a) ,or is not binding but observe a large positive budget shock, then reverting to a constant budget (or a decreased one) will create dynamics leading to a collapse of the system — where fewer and fewer professors are hired each year.

In regard of the discussed dynamics, we can conclude that incremental increases in budget — when the system has time to adapt to changes in professors-to-students ratio — should be favored against short spike in budget which can lead to a system collapse with only students remaining.