

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

On the Biomedical Elite: Inequality and Stasis in Scientific Knowledge Production

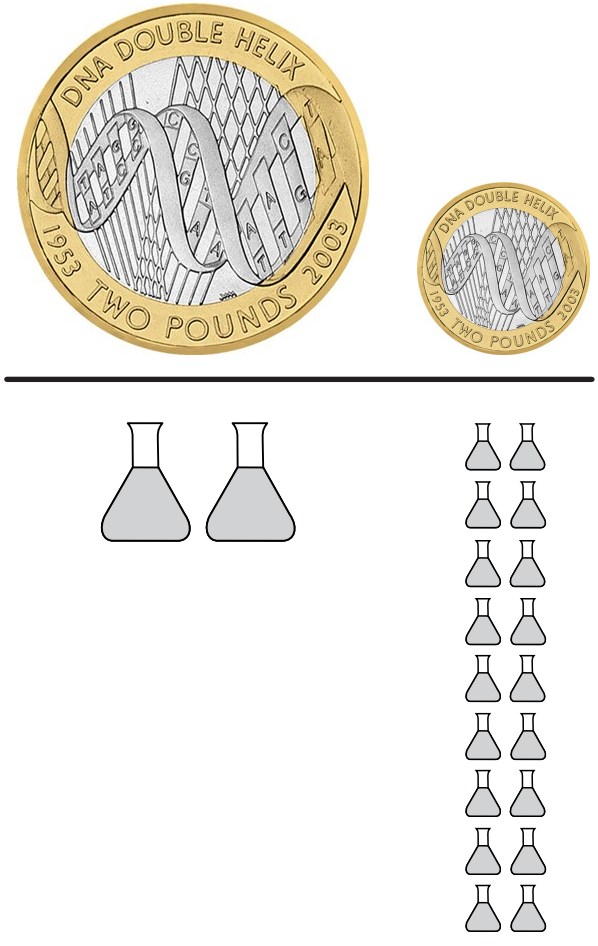
The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation	Katz, Yarden and Ulrich Matter. 2017. On the Biomedical Elite: Inequality and Stasis in Scientific Knowledge Production. Berkman Klein Center for Internet & Society Research Publication.
Accessed	September 29, 2017 2:37:32 AM EDT
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:33373356
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)

On the Biomedical Elite: Inequality and Stasis in Scientific Knowledge Production

Yarden Katz
Ulrich Matter



On the Biomedical Elite: Inequality and Stasis in Scientific Knowledge Production

Yarden Katz[1,2] and Ulrich Matter[2,3]

[1]Dept. of Systems Biology, Harvard Medical School, Boston, MA, USA
[2]Berkman Klein Center for Internet & Society, Harvard University, Cambridge, MA, USA
[3]Faculty of Business and Economics, University of Basel, Switzerland

July 11, 2017

Abstract

Researchers and research institutes are increasingly being evaluated using metrics (from bibliometrics to patent counts), which are core instruments of a longstanding effort to quantify scientific productivity and worth. Here, we examine the relationship between commonly used metrics and funding levels for investigators funded by the National Institutes of Health, the largest public funder of biomedical research in the United States, in the years 1985-2015. We find that funding inequality has been rising since 1985, with a small segment of investigators and institutes getting an increasing proportion of funds, and that investigators who start in the top funding ranks tend to stay there (which results in stasis, or lack of mobility). Furthermore, funding levels are a strong quantitative predictor of the interrelated set of metrics frequently used by economists and policy makers to evaluate scientific research. Our results suggest that the widespread system of metrics favors a minority of elite, highly funded researchers and institutes. Current attempts to "optimize" science are inextricably linked to the concentration of funds in the biomedical research system and are likely to further reduce diversity in the research community.

Introduction

The scientific enterprise in the United States and Europe, which is largely subsidized by government agencies, is molded by society's prevailing ideologies (Keenan, 1966; Dickson, 1984; Harding, 1991; Leslie, 1993). In recent decades, and in the United States in particular, there has been a dramatic economic and political shift toward neoliberal policies that promote the use of "markets" (competition under a price system) over alternative mechanisms (such as democratic or hierarchical decision making) to allocate resources. Advocates of these policies stress their allocative efficiency, an argument that is based on theoretical frameworks commonly used within economics and that requires having an objective measure for quantifying output. The proclaimed advantages of market solutions have been repeatedly challenged, however, by a rapidly growing empirical literature that uncovers their negative consequences for large parts of society. Neoliberal policies have produced large income and wealth disparities (Duménil and Lévy, 2011; Piketty, 2014) and a decrease in social mobility (Chetty et al., 2017). Neoliberal logic and its treatment of knowledge as a commodity has been reflected in academic research and universities more broadly (Mirowski, 2011; Brown, 2015). An elaborate system of metrics for quantitatively ranking and evaluating researchers (e.g., using publication and citation counts) is now commonplace, and can be seen as part of the project of creating a "market place" of scientific research (Mirowski, 2011)—an idea deeply rooted within economic thought (Stigler, 1964). This system operates in a research environment that, like the broader economy, is increasingly competitive (positions and funding opportunities are scarce) which compels decision makers to focus

Yarden Katz (yarden@hms.harvard.edu) is a fellow at the Berkman Klein Center for Internet & Society and the Department of Systems Biology at Harvard Medical School. Ulrich Matter (umatter@protonmail.com) is a Visiting Researcher at the Berkman Klein Center for Internet & Society at Harvard University and a WWZ Research Fellow at the Faculty of Business and Economics, University of Basel.

more on the quantification of output. What are the consequences of the neoliberal view of science that emerges in such competitive settings, and how is it linked to the funding sources that scientists rely on?

In the scientific arena, intense competition can produce a highly uneven distribution of resources needed to perform research. Increased competition for research funds can also shape outcomes for researchers in the academic labor market, where the ability to successfully raise funds is prized. If the allocation of resources is increasingly governed by metrics, then diversity—both within the directions of scientific inquiry and in the identity of researchers who pursue those—may suffer, depending on what segments of the research community are favored by the metrics. Thus, the virtual "market place" for scientific research, symbolized by metrics, may affect tangible outcomes insofar as the metrics are adopted by funding agencies, policy makers and scientists.

In this work, we use data on projects funded by the National Institutes of Health (NIH), the primary funder of biomedical research in the United States, to study the relationship between the allocation of NIH funds and the system of metrics in a competitive research ecosystem. Though the NIH funding system allocates funds through a centralized bureaucratic process, it relies on a peer-review system that, mirroring the U.S. economy, has gotten more competitive (and has seen lowered success rates for grant applications in recent years (Zerhouni, 2006)). We find that the NIH-funded biomedical research system shares two other salient features with the broader American economy. First, the distribution of research funds, like that of incomes, has been getting more unequal since the 1980s. Second, the biomedical funding system is marked by stasis, or lack of mobility, where initially well-funded investigators tend to persist in the top funding ranks. We show that the metrics used for evaluation of science strongly correlate with funding levels and that their use, taken in light of the rising concentration of funds and stasis present in the system, therefore privileges a minority of elite researchers and institutes. Our results suggest that the reliance on "market solutions" via metrics may work to further concentrate resources within biomedical research, potentially at the expense of the diversity of the research community.

Background

Pierre Bourdieu noted how some areas of intellectual and cultural production, such as the art world, outwardly "disavow" economic considerations pertaining to markets (Bourdieu, 1993). In contemporary biomedical science, by contrast, the language of business and finance is used explicitly by policy makers, economists and some scientists themselves when discussing how the scientific enterprise should be organized or how individual scientists should make decisions (Stidham et al., 2012; van Dijk et al., 2014; Kaiser, 2017a). Under the established view, public funding for scientific research is an "investment" (where the research projects being funded make up a "portfolio") and the "return on investment" can be quantified using metrics that will guide future funding and policy agendas, as well as hiring and tenure decisions. Various metrics—such as publication, citation and patent counts—are widely used by economists and policy analysts as proxies for scientific productivity, impact, innovation or quality (Azoulay et al., 2011; Boyack and Jordan, 2011; Li and Agha, 2015; Sampat and Williams, 2015) (see Mirowski (2011) and Tyfield (2013) for a critical perspective on these mainstream economic approaches to science). While the practice of counting publications in order to rank scholars dates as far back as the late 19th century (Csiszar, 2017), today's systematic practice of counting publications and citations is typically traced to the work of Eugene Garfield (Garfield, 1955, 1972), who viewed citation networks in part as a way to assess the "productivity" of researchers (or research fields) and journal "importance." The digital transformation in publishing has made it possible to implement bibliometric analyses on the grand scale envisioned by Garfield with far greater ease. Social network scientists have recently used such large-scale analyses to attempt to quantify and predict individual scientists' "impact" (Wang et al., 2013; Sinatra et al., 2016). The increased popularity and broad application of university rankings (Taylor and Braddock, 2007)—that are in turn a function of scientists' "impact"—contributes to this trend.

Apart from using bibliometrics to identify "high impact" work, there has also been a concerted effort to identify commercializable research. In recent decades, there has been a sharp rise in patenting of basic research, particularly within biomedicine (Mirowski, 2011). Landmark court cases and new legislation in the early 1980s expanded the range of biomedical research that can be patented (to include, for example, genetically modified organisms and gene products) (Heller and Eisenberg, 1998; Halbert, 2005; Katz, 2015). Since this period, universities have dramatically expanded administrative bodies that oversee faculty research

(Ginsberg, 2011), such as technology transfer offices that aim to patent potentially lucrative academic research. Patentability became another metric for evaluating scientific research and success in the patent office is often taken as an indicator of whether scientific research is yielding practical results for society (Li et al., 2017).

The metrics-based approach has been internalized by funding agencies. For instance, the U.S. National Science Foundation uses metrics regarding publications, patents and citations to quantify national scientific productivity (National Science Board, 2014) and the National Institutes of Health track publications by funded investigators that appear in journals with high "impact factors" (National Institutes of Health, 2016c). Publishing platforms and research-related social networking sites have also integrated metrics (Hammarfelt et al., 2016) (which are compiled and heavily promoted by companies like Google, who have the computing power and access necessary to pool publications at very large scale). The appeal of these metrics to funding agencies and university administrators, the ease with which they are calculated and broadcast, and their influence on the career prospects of researchers (whether real or imagined), has meant that metrics of scientific "worth" play an important role in the cultural and social fabric of scientific research.

The web of metrics used to quantify science is intertwined with funding through a number of mutually reinforcing relationships. Investigators who produce more publications are more likely to receive grants, which in turn enable them to expand their laboratories and increase the number of publications produced. Well-funded researchers in larger laboratories are likely to pursue more avenues of research in parallel, thereby increasing the chances of publishing in high profile journals, which begets more citations (Brembs et al., 2013)—in part because of the journals' own self-citation biases (Chorus and Waltman, 2016). A strong track record of publications boosts the chances of getting patents, viewed as a sign of the "translational" value of research, which in turn appeals to funding agencies and attracts industry sponsors. The circularity becomes apparent in quantitative analyses of impact: (Sinatra et al., 2016), for example, use citations as a proxy for scientific impact and attempt to validate this measure with Nobel prizes as an "independent" indicator, although the two cannot be decoupled. Thus, the scientific research enterprise is prone to circular "rich-get-richer" effects (sometimes referred to as "The Matthew Effect" (Merton, 1968)).

These effects do not apply uniformly across researchers, but rather reflect the social prejudices and norms of the community. Gender differences provide a paradigmatic example: laboratory job applications that have a traditionally male name are scored more highly by faculty members than those with a traditionally female name (Moss-Racusin et al., 2012), women receive overall lower scores on fellowship applications compared with men (Wenneras and Wold, 1997) and are cited less frequently than men (Caplar et al., 2017) (while men self-cite more than women (King et al., 2016)), and elite biomedical academic laboratories—which have access to long-term large funding sources and are likely to publish in prestigious journals—hire men over women at a higher rate than non-elite laboratories (Sheltzer and Smith, 2014). In light of this background, rich-get-richer effects threaten the diversity of the research ecosystem by concentrating resources in the hands of a few and doing so in a way that potentially amplifies existing social biases and hierarchies. There is a large body of historical evidence that suggests that the composition of the scientific community, which is sculpted by the feedback loops we have mentioned, can substantially alter the structure, methodology and standards of evidence in scientific work. Here too feminist scholarship provides illuminating instances from evolutionary biology, genetics, physiology and reproductive biology (Harding, 1986; Hubbard, 1990; Harding, 1991; Fausto-Sterling, 2008). Recent work on 19th century American science also highlighted how the marginalized perspectives of black scholars offered radically different analyses of human genetics, evolutionary history, biological determinism and brain function, along with distinct ways of using statistical evidence to draw causal inferences (Rusert, 2017)—all further substantiating the notion that the content of scientific inquiry is shaped by the perspectives of those who participate in it.

Here, we quantitatively analyze the allocation of funding in the biomedical research system in order to explore how rich-get-richer effects in funding may work to privilege different researchers and research institutes. We begin with a longitudinal analysis of the rich-get-richer phenomenon in public research funding and then relate funding levels of investigators and institutes to a set of commonly used metrics.

Analysis

Rising inequality in NIH funding for investigators and organizations

To examine the evolution of funding patterns through time and their relation to various metrics, we obtained data about NIH-funded projects from 1985-2015 from the NIH ExPORTER database (see Methods for detailed description of data acquisition and preprocessing). These data contain information about grants awarded to nearly 90,000 Principal Investigators (PIs) across ∼2,700 organizations, including the publications and patents that have resulted from these projects (Figure S1). We have also obtained citation counts for all ∼993,000 million publications in this data set (see Methods). We have compiled all of these data into a single publicly available relational database with standardized names for institutes and organizations to enable consistent longitudinal analyses.[1]

For the purposes of our analysis, investigators can be thought of as participants in an "economy," where overall "income" is the total NIH funding awarded that year, and each investigator (and by extension, the investigator's institute) obtains a fraction of this income. Through the interrelated ways mentioned above, this funding is used to produce publications, patents and other quantifiable research "outputs" that feed into metrics for evaluation and ranking. We can then ask: (1) how egalitarian the distribution of NIH funding is, both at the level of individual investigators as well as the level of institutes where these investigators work, and (2) how funding levels relate, quantitatively, to the web of metrics for evaluation, ranking, and prestige that are prevalent in the biomedical world. We address these questions by drawing on tools from the economics of inequality that are commonly used to analyze income and wealth disparities (Atkinson, 1970; Sen, 1973; Piketty, 2014; Duménil and Lévy, 2011; Cobham et al., 2016).

We examined inequality at two levels: (1) PI-level (how equal is the allocation of funding to different PIs?) and, (2) at the organization-level (how equal is the allocation of funding to different organizations (e.g., universities)?). We only included in the analysis NIH grants that are geared for research, excluding grants for clinical drug testing, educational curriculum development, or individual awards for students or postdoctoral staff. We examined funding levels based on the complete set of research grants ("All grants"), as well as separately using a subset of these grants that belong to the NIH R-series ("R grants")—see Methods for complete list of grant types used. R grants are generally awarded to individual laboratories for pursuit of their research (such as R01 or R21 grants) and often provide the core funding of the average biomedical research lab (particularly in universities). These grants are distinct from large cross-institute grants (such as U01 grants), which would be included in "All grants" but not "R grants." Therefore, the set of all grants provide a measure of the overall research funding landscape, while R grants are a proxy for the individual laboratory-focused state of funding. Since the NIH data do not distinguish cases when a PI received no funding because of grant applications being denied from cases when the PI did not apply for any grants, we only considered PIs who received some funding when computing funding summary statistics for the funding distribution at any given year. Given that NIH success rates can be as low as 10% or less (National Institutes of Health, 2016a), this important caveat means that our analyses likely underestimate disparities in funding among PIs (and consequently, among organizations) by not accounting for PIs who submitted grant applications but were denied funding.

We calculated the *cumulative funding* for each PI and organization using these data for the period 1985-2015. The cumulative funding for a PI/organization at a given year is the funding available to that entity in that year, carrying over relevant portions of funding awarded in previous years (relevance is determined from the budget duration field of awarded grants and funds from multi-PI grants were assumed to be divided equally among PIs; see Methods for details). All remaining references to funding refer to cumulative funding. As has been previously observed (Hand, 2012), we found that the distribution of NIH funding at a given year is highly unequal, both at the level of PIs and organizations for both sets of grants we considered (Figure 1A and Figure 1B, respectively; note logarithmic scale). In the year 2010, for example, the top 10% of PIs received ∼40% of total research grant funding, while the bottom 40% received only ∼11% (Figure 1C and Table 1). PIs in the top 10% that year had median funding of $1.45M, compared to the bottom 40% who had median funding of $140,000 (Table 1). The disparity among organizations was even larger: the top 10% of organizations in 2010 received ∼80% of total funding (with median funding of $108 M), while the

[1] Database is available on Github.

bottom 40% received less than 1% (with median funding of $220,000) (Figure 1D and Table 1).[2]

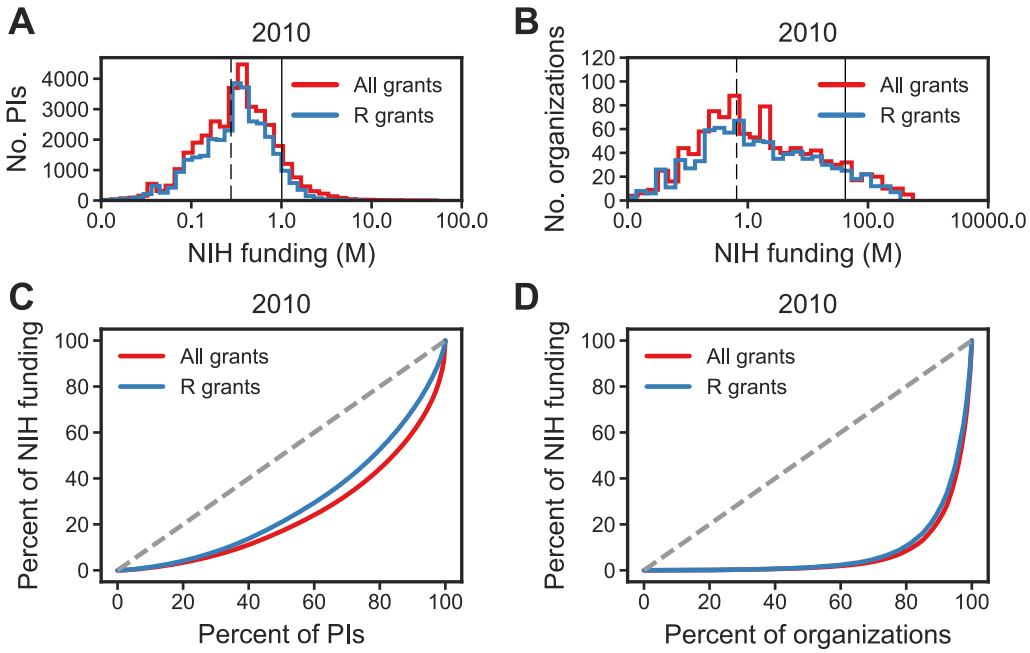


Figure 1: NIH funding inequality in a single year, 2010. (A) Distribution of NIH funding per PI for all in red and R-series grants in blue (see text for distinction), x-axis on $\log_{10}$ scale (in millions of dollars, with unlogged axis labels). Dotted vertical line marks end of bottom 40% of PIs (based on all grants); solid vertical marks beginning of top 10% of PIs (based on all grants). (B) Distribution of NIH funding per organization (as in A; note $\log_{10}$ x-axis scale). (C) Lorenz curves of funding per PI for all grants (red) and R-series grants (blue), where x-axis is cumulative bottom percentile of PIs and y-axis is percentage of total NIH funding belonging to the bottom cumulative percentile. (D) Lorenz curves for funding per organization for all grants (red) and R grants (blue) (as in C).

Top 10%									
2005					**2010**				
Median	Share (%)	No. of PIs/Orgs	Level	Grant type	Median	Share (%)	No. of PIs/Orgs	Level	Grant type
$1.25 M	38.9	3,185	PI	All	$1.45 M	40.0	3,529	PI	All
$1.03 M	29.7	2,898	PI	R	$1.13 M	30.3	3,168	PI	R
$83.99 M	76.8	100	Org	All	$108.00 M	78.1	102	Org	All
$78.90 M	74.1	91	Org	R	$90.71 M	73.8	89	Org	R
Bottom 40%									
2005					**2010**				
Median	Share (%)	No. of PIs/Orgs	Level	Grant type	Median	Share (%)	No. of PIs/Orgs	Level	Grant type
$0.13 M	11.7	12,737	PI	All	$0.14 M	11.1	14,112	PI	All
$0.13 M	14.3	11,588	PI	R	$0.15 M	13.8	12,668	PI	R
$0.17 M	0.6	399	Org	All	$0.22 M	0.6	406	Org	All
$0.15 M	0.6	361	Org	R	$0.22 M	0.7	353	Org	R

Table 1: NIH funding statistics for PIs and organizations in the years 2005 and 2010.

[2]In this data, grants are associated with organization identifiers that distinguish between institutes that are in fact part of the same umbrella institute. For instance, in our data, Harvard Medical School and the Harvard School of Public Health are both treated as distinct from Harvard University's main campus, although these are all part of the same institute. Similarly, Harvard Medical School is treated separately from the Dana-Farber Cancer Institute in Boston and other research hospitals affiliated with the medical school although these can share staff, resources and grants. This granular representation of organizations probably underestimates the extent to which funds are concentrated at the organization level.

In order to assess inequality trends through time, we computed the Gini index for the years 1985-2016 across organizations and PIs (Figure 2A and Figure 2B, respectively). The Gini index is computed from the Lorenz curve of a distribution (such as those shown in Figure 1C, D) and ranges from 0 to 1 where the higher the Gini value, the more unequal the distribution of funding: 0 corresponds to a perfectly egalitarian distribution of funds, while 1 indicates that a single entity obtains all funds. We found that Gini values have been rising since 1985, both at the PI level and organization level (Figure 2A and Figure 2B, respectively), with a sharper increase when considering all research grants compared with only R grants. Since 2010, inequality appears to be decreasing at both the PI and organization level (the source of this decrease is likely explained in part by changes in NIH policy which are discussed in the Appendix).

While the Gini index continues to be popular in analyses of wealth and income inequality, several disadvantages of the index have been pointed out (Damgaard and Weiner, 2000; Piketty, 2014; Cobham et al., 2016), notably that the Gini is overly sensitive to changes in the middle range of the distribution but insufficiently sensitive to changes near its tails (Cobham et al., 2016). The Palma index of inequality (Cobham et al., 2016), which is simply the ratio of the share of income going to the top 10% to that of the bottom 40%, has been proposed as an alternative. We calculated the Palma index across PIs and organizations and found a similar pattern of rising inequality to that obtained with Gini (Figure 2C and Figure 2D, respectively), where PIs in the top 10% had $\sim$2-2.5 fold more funding than the bottom 40% in 1985-1995, but $\sim$3-3.5 fold more funding by the 2000s. Similarly, organizations in the top 10% went from having 50-60 fold more funding than the bottom 40% in 1985-1995 to over 100-fold in the 2000s (though the disparity appears to have decreased in recent years).[3] Since the Palma ratio is more interpretable than the Gini index and highly correlated with Gini values in our data (Figure S2), we use the Palma to summarize inequality in the remaining analyses.

A distribution can be made less equal in numerous ways that shrink and grow the share of income obtained by different percentiles (or "classes") of the population. To better understand the way in which funding inequality has changed since 1985, we looked directly at the share of funding awarded to the top 10% and the bottom 40% across years (Figure 3). At both the PI and organization level, we found that the share of funding going to the top 10% has been increasing, and the share going to the bottom 40% decreasing, but these shifts appear to take place at different rates (Figure 3). When considering all research grants, the share of funding going to the top has been increasing faster than the share to the bottom has been decreasing (Figure 3A), and yet faster when analyzed by organization (Figure 3B). When considering only R-series grants, the rate of funding increase to the top has been slower than the rate of funding decrease to the bottom for PIs (Figure 3C), but faster at the organization-level (Figure 3D).

Funding inequality, for both individual PI laboratories and organizations, is connected to the size of these entities. Highly funded PIs can expand laboratory personnel which in turn requires larger budgets. Likewise, organizations with more research PIs will receive more overhead funds (and have a larger research and administrative community that can enable PIs to be more effective at fund raising), and as a result, these organizations can expand and hire more research staff. Such feedback loops mean that the size of a laboratory or organization can work to both determine and help increase its funding. To account for these factors, ideally one would use information about the number of personnel in laboratories and organizations through time and their success rates in applying to grants, but these data are not available and difficult to reliably estimate. To mitigate the contribution of organizations that may, due to their size, only rarely apply for funding on our assessment of inequality, we examined inequality for the subset of organizations that have received NIH funding for at least a decade (Figure S4; note that as in the rest of our analyses, the calculation of inequality statistics at any given year is made using only organizations that had some funding that year). We observed a similar qualitative pattern to that shown in Figure 3, namely that since the 1980s, inequality has generally been on the rise but appears to have decreased in recent years (see Appendix for more information).

[3] Inequality as analyzed here is a composite of multiple NIH subdivisions that have budgets of varying sizes and award grants to distinct research communities (National Institutes of Health, 2016b). We further analyzed PI funding inequality across time separately for each NIH subdivision that is present in all years of our observation period (see Figure S3 for top 12 most highly funded institutes). As expected, inequality trends were different in distinct subdivisions, as measured by Palma in 1985-2015 (Figure S3). The National Inst. of Cancer, for instance, had a faster increase in inequality compared to the National Inst. of Arthritis & Musculoskeletal & Skin Diseases, while the National Eye Inst. maintained a nearly perfectly constant level of inequality. The majority (62.5% for all grants; 75% for R grants) of NIH subdivisions shows rising inequality when we calculate the slope of the best linear fit for the Palma index for each NIH subdivision. Among the top 12 most highly funded subdivisions 9 (11 for R grants only) show a trend towards higher inequality.

Stasis and mobility across funding ranks

Our results indicate that the top percentile of the NIH-funded biomedical field has been getting richer with funding, while the bottom has been getting a smaller share of the overall funding. We next asked how the rising funding inequality manifests in the funding trajectories of individual investigators: for an individual PI, what are the prospects for mobility across the funding ranks? In the year when a PI first receives NIH funding, the PI may be in one of three ranks, depending on the received funding amount: the top 10%, the bottom 40% or the "middle" (between top 60th and top 10th percentiles). Within a given time period, the PI may move up or down these ranks, stay in place, or fail to receive funding altogether (potentially by exiting the funding system and not applying for NIH funding). To examine how the initial funding rank affects PI funding mobility, we divided the years 1985-2015 into sliding eight-year windows and for each year t, selected only the PIs who received their first NIH funding in t and examined their funding trajectory in $[t, t+7]$. We then calculated a *stasis score* for each PI in $[t, t+7]$, defined as the fraction of years the PI spent in the funding rank held at t. For instance, a PI who entered into the bottom 40% at year t and remained there for an additional year before moving up to be in the top 60% percentile would have a stasis score of $2/8$ (0.25), since 2 out of the 8 years were spent in the rank the PI was in at time t (bottom 40%).

For each eight-year window, we calculated the average stasis score across PIs who first entered the NIH funding system in the top 10% at the beginning of the time window (solid lines in Figure 4A; 2008 marks the value for final eight-year window). For comparison, we simulated a simple null model where the allocation of funding per year t is independent of previous years and depends only on the distribution of grant sizes and durations (i.e., the number of years in which the PI will receive the grant) awarded that year. Under the null model, a population of 1,000 PIs is seeded during the first year, 1985, and awarded grants

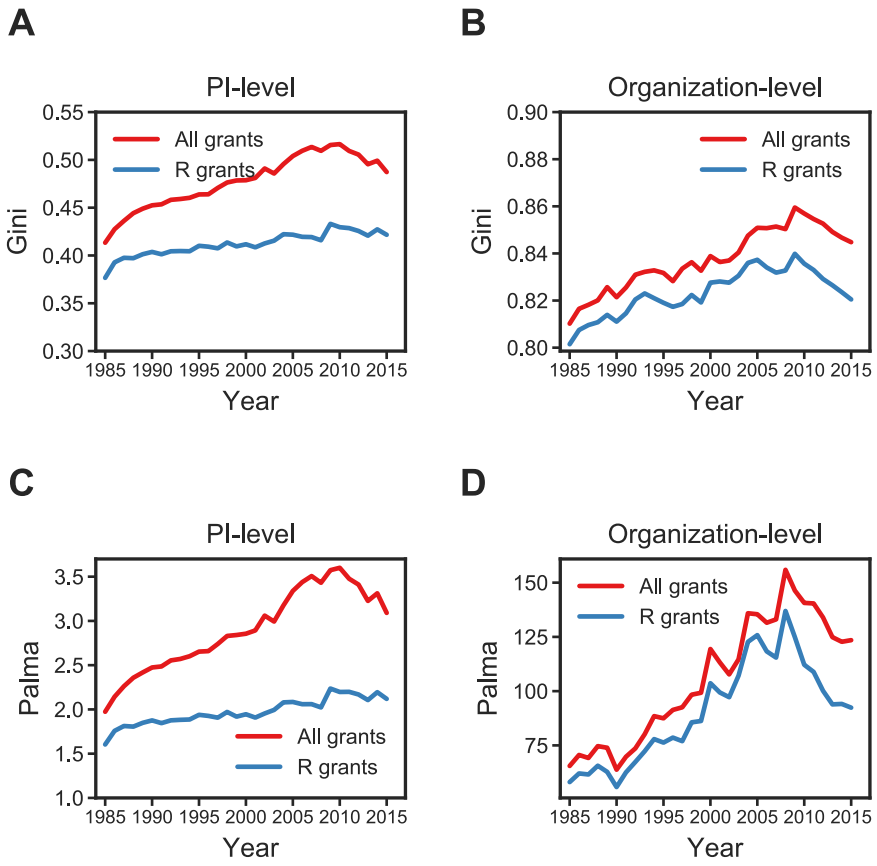


Figure 2: Inequality trends in NIH funding, 1985-2015. (A) Gini values for funding at PI-level. (B) Gini values for funding at organization-level. (C) Palma ratios for funding at PI-level. (D) Palma ratios for funding at organization-level.



Figure 3: Funding trajectories for top and bottom ranked PIs and organizations, 1985-2015 (red indicates calculation using all grants, blue using only R grants). Slope (m) calculated from best fit line. (A) Top: percentage of total NIH funds (all grants) granted to PIs in top 10% of funding distribution. Bottom: percentage of total NIH funds (all grants) granted to PIs in bottom 40%. (B) Top: percentage of total NIH funds (all grants) granted to organizations in top 10%. Bottom: percentage of total NIH funds (all grants) granted to organizations in bottom 40%. (C) Top: percentage of total NIH funds (only R grants) granted to PIs in top 10%. Bottom: percentage of total NIH funds (only R grants) granted to PIs in bottom 40%. (D) Top: percentage of total NIH funds (only R grants) granted to organizations in top 10%. Bottom: percentage of total NIH funds (only R grants) granted to organizations in bottom 40%.

sampled independently from the empirical joint distribution of grant size and duration in 1985. In the following year, the population are randomly assigned grants by the same mechanism, carrying over additional funding obtained the previous year (i.e., in cases where grants with duration greater than 1 year were assigned), and so on, until the first sliding window of 8 years is complete. The same procedure is then repeated for the next sliding windows. Following exactly the same procedure as outlined above, the null stasis score is then taken to be the average across the simulated population of PIs.

The average stasis score observed in the data for PIs who started in the top 10% was $\sim$0.58 (solid lines, Figure 4A), substantially greater than the $\sim$0.27 expected under the null model (dotted lines, Figure 4A). PIs who start in the bottom 40% generally tend to spend more time in the bottom than expected under the null model, but this effect is much smaller (Figure 4B). This analysis suggests that PIs who start in the top percentiles of funding tend to stay in the top.

A quantitative relationship between funding levels, publication metrics and journal prestige

Our results so far point to a concentration of research funds that has generally become more pronounced through the years, in a funding system where investigators who start out being well-funded tend to remain well-funded. Given the central role that money plays in sustaining biomedical research laboratories and enabling their expansion, we next asked how the unequal distribution of

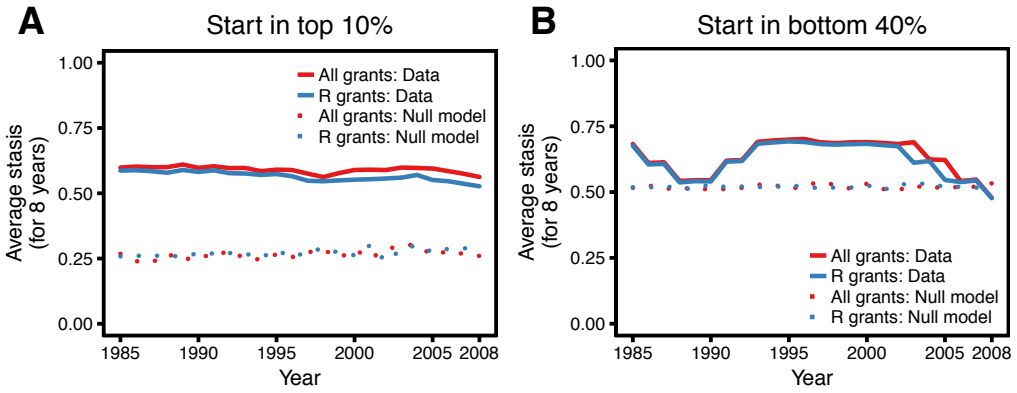


Figure 4: Stasis and mobility in the NIH funding system, 1985-2015. (A) Average PI stasis score calculated for eight-year sliding windows, starting in 1985, for PIs whose first funding placed them in top 10%. Red curves for all grants, blue for R grants. Dotted red and blue curves correspond to average stasis scores under a null model that controls for the size and duration of all grants and R grants, respectively (see text for details). (B) Same as (A), except considering PIs whose first funding placed them in bottom 40%. Red curves for all grants, blue for R grants; dotted red and blue curves correspond to null models for all grants and R grants, respectively.

funding relates to scientific production as measured by the set of metrics that we discussed earlier.

We first explored the relationship between the funding available to PIs and the number of publications these PIs (co-)author. We compiled the list of all publications associated with R grants in the years 2005-2010[4] using NIH data and then extracted citation counts for these publications from an external database (see Methods). We first calculated the average funding percentile for each PI over the period 2005-2010 using only funding from R grants. This average percentile provides us with a ranking of how well-funded each PI is relative to all PIs who received *some* funding in the period 2005-2010. We then selected a subset of those PIs that: (1) had funding during *every* year in 2005-2010, and (2) co-authored at least one publication (associated with R grants) in this period. We focused on PIs who had consistent funding throughout the entire six year period in order to exclude those who have exited the funding system (e.g., through retirement). (This conservative choice, along with our requirement of at least one publication in this time period, means that our selection of PIs is likely skewed towards those investigators that are generally successful fund raisers.)

We next asked how the funding rank of PIs (calculated as described above) compared to the total number of publications they (co-)authored. This revealed a quantitative relationship: the more highly funded PIs have a substantially larger number of publications (Figure 5A, note logarithmic y-axis). Similarly, and as expected, more highly funded PIs accumulate a larger number of citations (Figure 5B, note logarithmic y-axis). Both analyses point to a strong stratification at the top end of funding: e.g., PIs between the 80th and the 90th percentile have nearly 2-fold fewer publications than PIs in the top 10% (a similar stratification was observed for citation counts, Figure 5B). At the level of organizations, we observed a similar result. As for PIs, we calculated the average funding percentile of organizations who received some R grant funding in 2005-2010. We then selected only those organizations that were consistently funded throughout the entire period and compared their average funding percentile to the number of publications and citations that these organizations produce (Figure 5C and Figure 5D, respectively, where each point denotes an organization; note logarithmic scale). Similarly to the case of PIs, we found that more highly funded organizations produce substantially more publications and citations and that this increase is more pronounced in the top funding ranks (Figure 5C,D).

In addition to publication and citation counts, journal prestige can play an important role in the social fabric of academic research, depending on the subfield, and so much so within biomedicine that the most prestigious journals have come under fire from some biologists in recent years (Schekman, 2013). In basic biomedical research, the so-called "glamour" journals—*Nature*,

[4]We chose 2005-2010 since this six year period is: (i) long enough to contain many grants, (ii) recent enough that it is plausibly representative of current trends, yet (iii) distant enough in the past so that publications associated with grants funded in this period have been entered in the NIH databases and their citations have had time to accumulate.

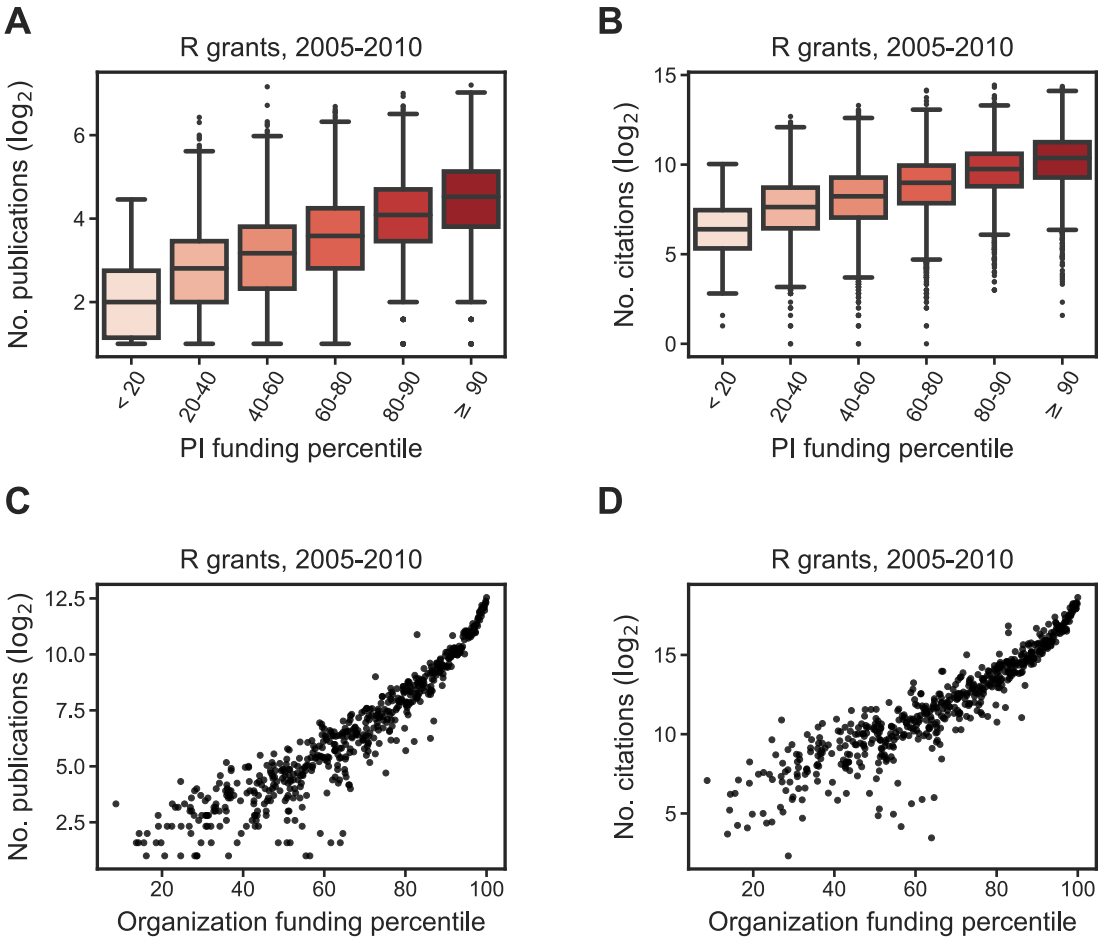


Figure 5: Roles of PI and organization funding rank in number of publications and citations. (A) Relationship between PI funding percentile (averaged from R grants in the years 2005-2010) and number of publications produced (co)authored by these PIs (including only publications linked to R grants; note $\log_2$ scale on y-axis). (B) Same as (A), except y-axis is number of citations (plus 1 to account for zero citations, on $\log_2$ scale). Citation counts were obtained from Elsevier database as described in Methods. (C) Relationship between organization funding percentile (averaged from R grants in the years 2005-2010) and number of publications produced by PIs in these organizations ($\log_2$ scale); each point corresponds to an organization. (D) Same as (C), except y-axis is number of citations (plus 1 to account for zero citations, on $\log_2$).

Science and *Cell*—are considered the most prestigious "high impact" venues, while in the more clinically-oriented communities publications in the *The Lancet* or *New England Journal of Medicine (NEJM)* are similarly prized. We compared the relationship between the "impact factors" of a set of 30 biomedical journals, often used as an indicator of prestige, and the median funding level of PIs published by the journals (based on publications associated with R grants in the years 2005-2010[5]). We found that journals with higher impact factors published papers by PIs with higher median funding levels, with the elite journals (such as *Nature*, *Science* and *Cell*) being most extreme in this regard (Figure 6A). The full distribution of PI funding percentiles (computed based on R grants) for a set of 10 general interest biology and genetics journals is shown in Figure 6B (ordered by impact factor). As an example, the median funding percentile of PIs published in a "high impact" journal such as *Nature* (impact factor of ∼38) is ∼80, while a "low impact" journal such as *PLoS One* had a median funding percentile of ∼70 for its published PIs. These results suggest that more prestigious journals tend to publish PIs who are less diverse in terms of funding.

[5]We use publications and funding ranks based only on R grants here in order to specifically exclude very large cross-center grants. Such grants often result in publications in high impact factor journals, which would potentially inflate the relationship between PI funding and journal prestige.

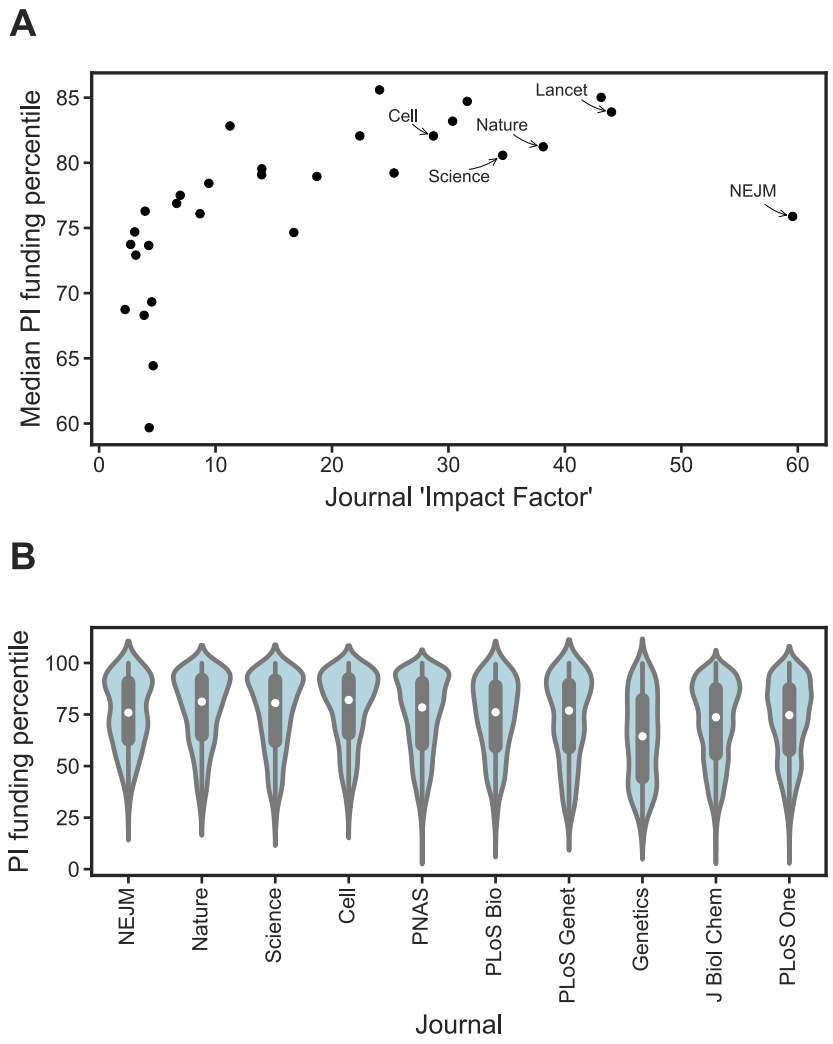


Figure 6: Prestigious journals publish more highly funded PIs. (A) Comparison of a journal's "Impact Factor" (JIF; obtained from Thomson-Reuters) and the median funding percentile of PIs who authored papers in the journal (averaged from R grants in the years 2005-2010). (B) Distribution of PI funding percentiles for ten popular scientific journals, sorted by JIF (highest on left, lowest on right).

Minority of highly funded institutes own majority of NIH-funded patents

NIH-funded research can also be patented. Patents are often viewed by economists as measures of productivity and innovation (Li et al., 2017) and are also taken into consideration by the NIH when evaluating grant applications. Following legislative changes and landmark court cases regarding patentability of biotechnology in the early 1980s (Katz, 2015; Heller and Eisenberg, 1998), the number of patents stemming from NIH-funded research has increased substantially (Figure 7A). We asked whether patents, like publications and citations, are also strongly concentrated in a minority of institutes. Indeed, we found that the vast majority ($\sim$80%) of NIH-funded patents are owned by just the top 10% of patent producing organizations, as shown in Figure 7B. As an example, the Massachusetts Institute of Technology and Stanford University own a combined $\sim$7% of all NIH-funded patents filed between 1971 and 2010. As with number of publications and citation counts, here too we observed a strong quantitative relationship between funding rank and number of patents owned (Figure 7C; note logarithmic y-axis), showing that the more well-funded organizations produce substantially more patents. Similarly to the case of other metrics we examined, the relationship between funding rank and the metric of interest was non-linear with stratification at the top tail (Figure 7C). These results are consistent with a model where only a minority of elite institutes have the legal and financial means to successfully obtain patents.

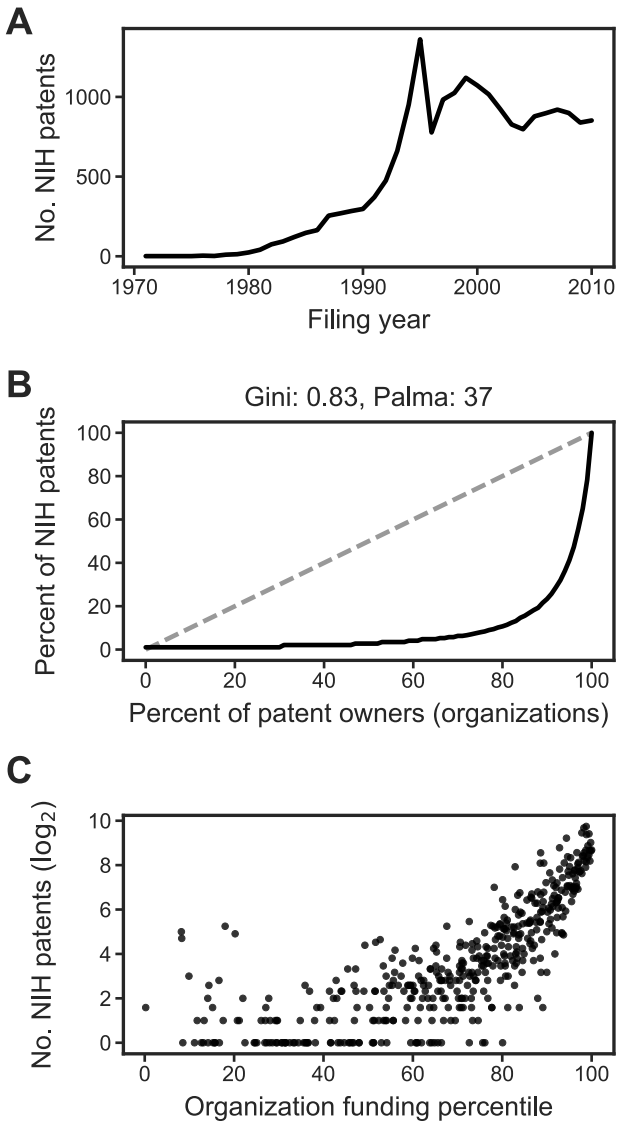


Figure 7: Inequality in NIH-funded patent production. (A) Number of NIH-funded patents by filing year. (B) Lorenz curve for NIH-funded patents (x-axis is bottom cumulative percentile of patent owning organizations, y-axis is percent of total NIH-funded patents). Dashed grey line marks perfect equality. (C) Comparison of the funding percentiles of organizations (average of percentiles computed from R grants in the years 2005-2010) to the number of NIH-funded patents owned by organizations ($\log_2$ scale). Each point corresponds to an organization.

Discussion

As Thorstein Veblen observed a century ago (Veblen, 1918), the social values internalized by institutes (universities, in Veblen's case) can sometimes only be seen in retrospect, when their imprint on institutional policy becomes apparent. Our longitudinal analysis here shows that a hallmark of neoliberal economies—a widening disparity in wealth across the population (Duménil and Lévy, 2011; Piketty, 2014)—is also reflected in the realm of scientific knowledge production. We found that funding levels are increasingly concentrated in a minority of institutes and investigators, and are strongly linked to the ranking and evaluation metrics promulgated by economists and policy analysts. A second trend observed in the neoliberal U.S. economy—that of decreased social mobility (Chetty et al., 2017)—also has a correlate in our analysis: investigators who enter the system in the top funding ranks tend to stay there, resulting in a kind of "stasis." As in the real economy, a low level of mobility in the context of NIH funding may be a contributing factor to increasing inequality.

An important limitation of our analysis is that we have only considered investigators who successfully applied for NIH funding (excluding those whose funding applications were denied). In light of the intensifying competition for NIH grants in recent years, our analysis probably substantially underestimates the overall inequality in the funding system. We have also not accounted for any private sources of funding that investigators may have (from industry partners or non-profit organizations such as the Howard Hughes Medical Institute), which are also likely to be concentrated in hands of elite investigators. While we have not directly addressed the underlying causal mechanisms of rising inequality, the analysis presented here can serve as a basis for policy debates concerning U.S. biomedical research. Since different fields operate by distinct institutional mechanisms, our results cannot be directly extrapolated to other disciplines (or to biomedical funding in other countries), but attempts to quantify knowledge production via metrics are widespread, and it's likely that the basic relationship between funding levels and success by the metrics that we have found is not idiosyncratic to biomedicine in the United States. For example, in economics (and related social sciences), metrics and rankings to quantify the "worth" of research papers, journals, individual academics and university departments are ubiquitous and considered much more central to an early-career economist's trajectory today than several decades ago (Frey, 2009). This becomes apparent in discussions within economics surrounding the development of more sophisticated bibliometrics (in the United States) (Perry and Reny, 2016) and rankings (in Germanic-speaking Europe) (Gygli et al., 2017).

It is tempting to try to disentangle the complex causal processes that drive the strong associations we have observed between funding and commonly used metrics. One might hypothesize, for instance, that investigators at elite institutes are simply more likely to submit their work to prestigious journals, and that this might partly explain the correlation between funding and publications. However, the underlying factors driving these correlations aren't taken into account in the neoliberal view of science. Part of the ideological appeal of metrics-based "objectivity" lies precisely in its offer of a supposedly universal standard of measurement that transcends the distinct social, cultural and methodological practices of different investigators and scientific communities. Though there are attempts to normalize metrics by coarse field-specific values (e.g., the Relative Citation Ratio used by NIH (Kaiser, 2017a)), the metrics-based view is still predicated on the idea that the local practices found within science—its "disunity" not just across but within scientific fields (Galison and Stump, 1996)—can be divorced from evaluation and ranking.

Our intention here is not to tease apart the set of metrics from funding or from one another, but rather to highlight their interrelation, which is at work when individuals and institutes are judged by metrics taken at face value. The picture that emerges from our results is very much in line with what we would expect from an economic perspective, given the incentives set in the current research system. If having more "impact" according to the defined metrics is rewarded, universities and funding bodies will (at least on average) try to maximize these metrics. If university presidents are perceived to be successful if their university is highly ranked, they will pursue a personnel strategy that leads to higher rankings, etc. However, this logic, and the circularity it is mired in, holds independently of whether or not the knowledge and scientific insights produced in the process are actually considered beneficial for the scientific endeavour or for society at large. This is why judging that the system "works well" from the observation that PIs who receive more funds score more highly by the metrics is fallacious reasoning.[6]

The metrics-based view of science found in mainstream economics is bound to leave its imprint on the culture and composition of the biomedical research community. There are already a number of forces that work to reproduce the elite ranks and limit diversity: for instance, elite biomedical research laboratories tend to employ fewer women (Sheltzer and Smith, 2014) and top ranked institutes tend to hire their own (Clauset et al., 2015) (in addition to the well-known intergenerational reproduction of academic faculties famously documented in (Bourdieu, 1988)). The use of metrics, as our results show, can be another force contributing to privilege an elite minority. Moreover, the increased competition vis-à-vis the quantification of output via simple metrics is likely to affect self-selection into biomedical research. That is, the field may become more attractive to young researchers

[6]Our argument here can be related to a form of "performativity" where the act of adopting a representation alters the state of the world so as to better conform to that representation. A notable example is the argument by MacKenzie and Millo that once option traders started using the Black-Scholes formula, market dynamics were altered so as to fit the assumptions of the mathematical model underlying the formula; the model wasn't necessarily a faithful representation of pre-existing market dynamics, but its adoption drove market changes that the model is suited to explain (MacKenzie and Millo, 2003). A crucial difference between this phenomenon and the potential "performativity of metrics"' in our context is that the adoption of metrics by institutes would *not* simply make it so that the scientific value or knowledge gained (what the metrics are purportedly a proxy for) will in fact be proportional to, say, short-term gains in citation counts. A weaker notion of performativity trivially holds at the institutional level: if funding agencies, universities and various committees decide to reward scientists based on how quickly their work accumulates citations then those individuals and their projects will serve as exemplars for subsequent funding and promotion decisions (meaning that by accepting a given set of metrics as a proxy for scientific value, communities will alter their composition to reflect this belief).

who see themselves as suited to a highly competitive environment where having a career is perceived to be contingent on one's ability to maximize clearly defined metrics such as the number of "prestigious" publications or citation counts.[7] In response to criticisms of the highly skewed distribution of funding in biomedical research, the NIH recently proposed a cap on the funds a single investigator can receive (National Institutes of Health, 2017a), which sparked strong backlash from some elite biomedical scientists (Weisman, 2017) and was promptly withdrawn (Kaiser, 2017b). The selective pressure induced by the maintenance of this concentration of wealth (exerted in part by a small and influential segment of biomedicine that profits from the status quo), coupled with the emphasis on metrics, is likely to exacerbate the lack of diversity within biomedical science. An interesting extension of our work would be to characterize the investigators in different funding ranks (and by scores achieved via metrics) to better understand their identities, subfields and scientific trajectories. This may shed light on the broader question of how differential access to funding in a highly competitive system constrains and guides avenues of scientific inquiry.

Acknowledgments

We thank Ariella Azoulay, Nazim Bouatta, Vincent Butty, Jenny Chen, Bruno S. Frey, Lukas Rieppel, Alois Stutzer, Lauren Surface, Caleb Weinreb and the Berkman Klein FH community for helpful discussions and comments. We also thank the NIH RePORT team for their careful work in maintaining the ExPORTER data tables and for answering several questions about the data.

Methods

Data sources and preprocessing

1. NIH grants: data about funding, publications and patents for NIH-supported projects were downloaded from NIH ExPORTER in April-May 2017.
2. Patents data: filing dates for NIH patents were obtained from Google patents database using custom code.
3. Citations data: citations for NIH-supported papers were obtained using the Elsevier Developer's API.
4. Journal rankings data: 'Impact Factors' for journals were obtained from the Thomson-Reuters InCites Journal Citation Reports.

NIH grants selection: We focused on research grants (as opposed to grants for development of curricula, or grants aimed at businesses for targeted clinical development and testing) whose primary recipients are principal investigators at research institutions. We excluded grants considered as individual "training grants" at the undergraduate, graduate or postdoctoral levels, or grants that are often granted to research core facilities (such as P20 or P30 grants). Two grant sets were used:

1. *All grants*: Includes NIH research grants of type:
    - R01, R03, R21, R33, R35, R37, R55, R56
    - RC1, RC2, RC4
    - RF1
    - RL1, RL2
    - DP1
    - K05, K06, K08

[7] Adverse or unintended effects of extremely competitive markets are becoming increasingly recognized in a growing body of empirical evidence in labour economics and management science (Niederle and Vesterlund, 2007, 2011).

- P50
- U01, U24, U54

2. *R grants*: The subset of R, RC, RL and RF grants.

Calculation of cumulative funding: In each year, we considered the "cumulative" funding available to every PI (identifier by their unique PI ID as it appears in NIH records). To calculate the cumulative funding for a PI at a given year t, we considered all grants that carry over from years prior to t, as well as the yearly budget available to the PI from any new grants awarded in year t (using the `Total_Cost` column in the NIH data when available and `Total_Cost_Sub_Project` otherwise). For grants with multi-year budgets, we assumed that the funds are split equally across the budget years—e.g., a PI who receives a $3M grant in the year 2000 with a budget lasting three years would be treated as having $1M in each of the three years (2000, 2001 and 2002). Grants with negative budget duration were discarded. In the case of multi-PI grants, we assumed the grant money was split equally among all PIs. Cumulative funding for organizations was calculated in the same way as for PIs, except grant funding was aggregated using a unique identifier assigned to each organization (which is the nine-digit DUNS ID for the organization when available and the organization's name otherwise; `org_id` in our database).

A publication was associated with the PI if the manuscript was supported by a grant awarded to the PI and if the PI was an author on the publication (i.e., if the PI's last name appeared in the publication's author list as it appears in PubMed).

Data availability: All data has been entered into a SQLite database (available on Github).

Supplementary Figures

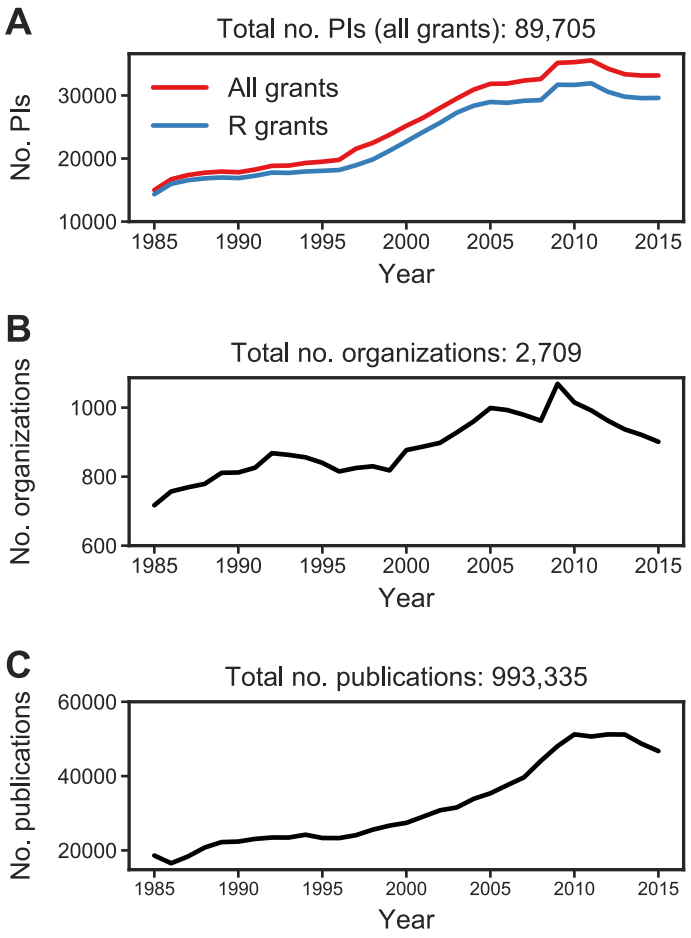


Figure S1: (A) Number of PIs with non-zero (cumulative) NIH funding by year. (B) Number of organizations funded by NIH by year. (B) Number of publications associated with NIH funded projects by year.

Appendix

Changes in NIH grant composition for organizations across time: As shown in Figure 2 and Figure 3, inequality appears to have been decreasing from its high levels in recent years. To explore the factors that may be contributing to this trend, we analyzed the composition of grants awarded to the top 10% versus the bottom 40% in 1985-2015 (Figure S5). For each year, we calculated the percentage of total funding (awarded to top or bottom class) that was contributed by each of the grant types listed in Methods and plotted these percentages for only the subset of grant types that contributed substantially to the total funding of both classes (Figure S5). In 2009-2010, the top 10% started receiving a smaller percentage of the total funding (Figure S5A) and while the composition of grants awarded to the top has changed between the period before 2009-2010 and the period after (Figure S5B), when the bottom started to receive an increasing share of funding (Figure S5C), this change was smaller compared to that observed in the composition of grants awarded to the bottom (Figure S5D). For example, although both the top and bottom received funding from NIH stimulus grants in 2009-2010 (part of the American Recovery and Reinvestment Act), such as RC1 grants, the bottom has had a larger relative share of its funding coming from these grants (Figure S5D) compared to the top. In addition, both larger center grants (U01, U24, U54) and smaller laboratory research grants (R03, R21, R33) started contributing more substantially to

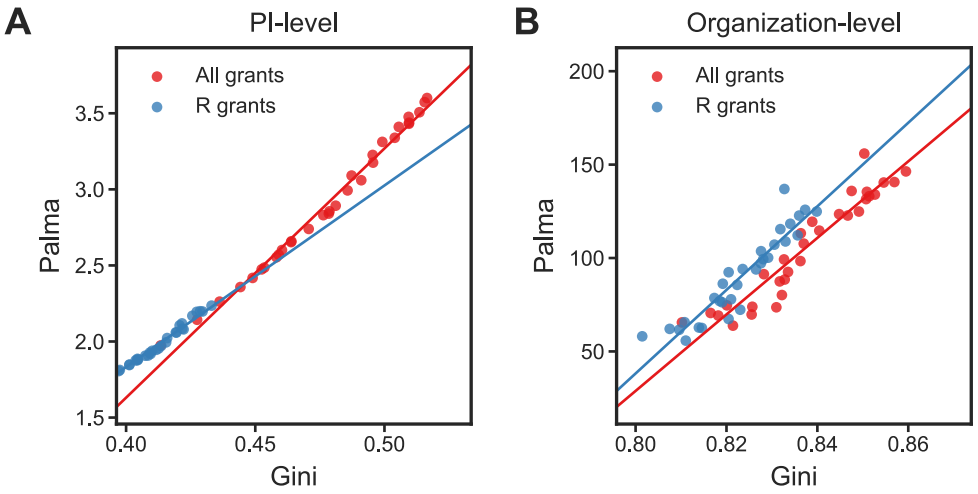


Figure S2: (A) Comparison between Gini and Palma values for funding at PI-level, for all grants (red) and R grants (blue). Each point represents a year. (B) Same as (A) for funding at organization-level.

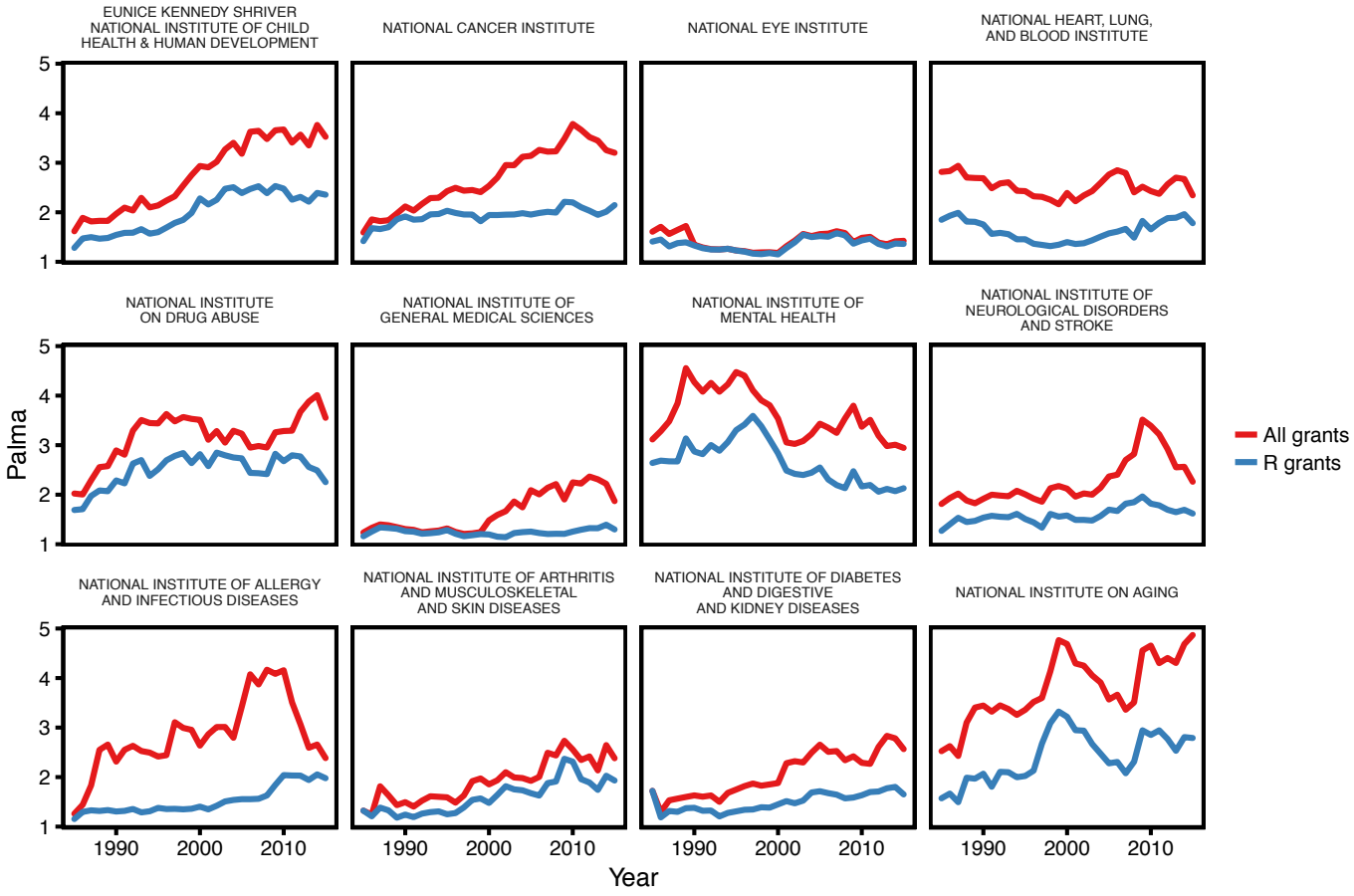


Figure S3: Inequality of funding at PI-level (as measured by Palma) for top 12 most highly funded NIH subdivisions.

the pool of funding awarded to the bottom around 2009-2010. These changes in grant composition may be partly explained by NIH policy initiatives that are intended to increase diversity (National Institutes of Health, 2017b).

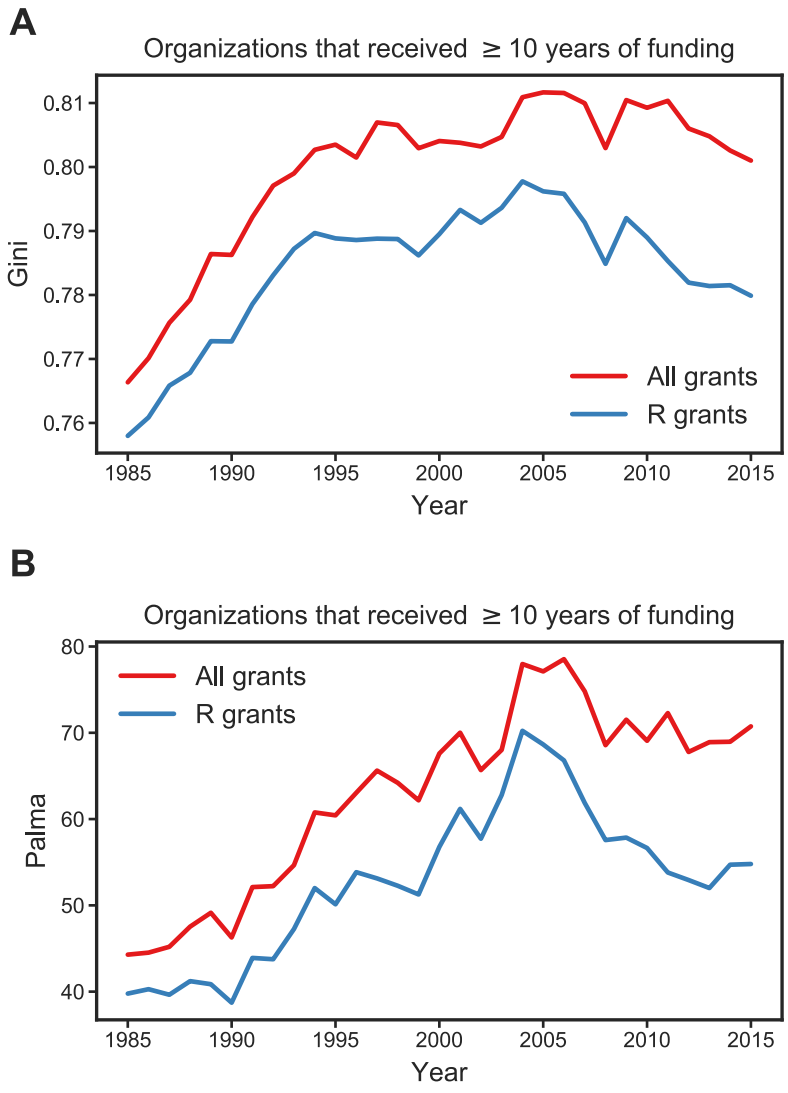


Figure S4: Inequality of funding for organizations with longstanding NIH funds. (A) Gini across years for organizations that received at least 10 years of funding in the period 1985-2015. (B) Same as (A), except showing Palma values.

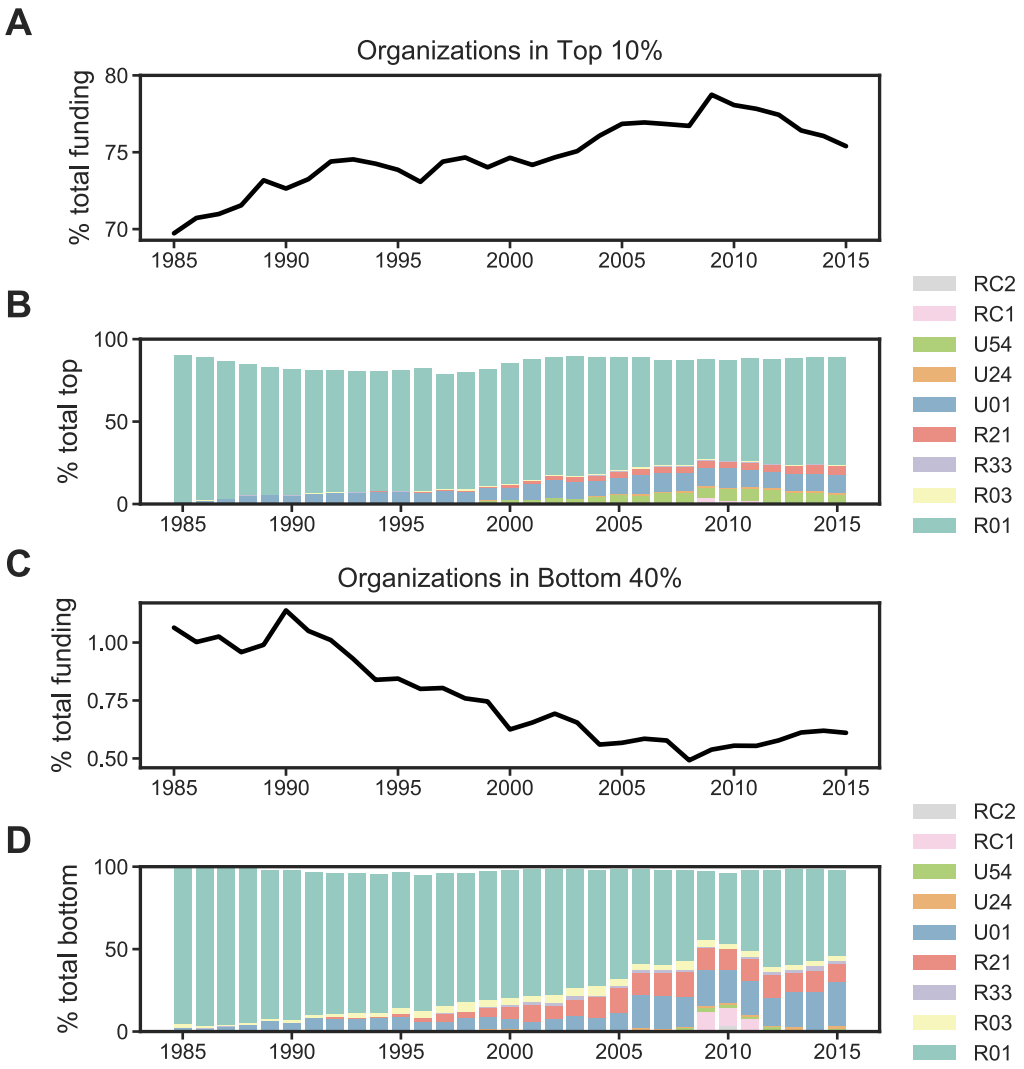


Figure S5: Changing composition of grant types awarded to top 10% and bottom 40% of organizations in 1985-2015. (A) Percent of total funding given to top 10%. (B) Percent of total funding awarded to the bottom 40% (including all grants listed in Methods) by grant type. Note that only a subset of the grants listed in Methods that contribute substantially to the total are shown here, which is why the sum of percentages per year does not exactly sum to 100%. (C) Percent of total funding given to bottom 40%. (D) Percent of total funding awarded to the bottom 40% by grant type (calculated in the same way as (B)).

References

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3):244 – 263.

Azoulay, P., Zivin, J. G., and Manso, G. (2011). Incentives and Creativity: Evidence from the Howard Hughes Medical Investigator Program. *The RAND Journal of Economics*, 42.

Bourdieu, P. (1988). *Homo Academicus*. Stanford University Press.

Bourdieu, P. (1993). *The field of cultural production: Essays on art and literature*. Columbia University Press.

Boyack, K. W. and Jordan, P. (2011). Metrics associated with NIH funding: a high-level view. *J Am Med Inform Assoc*, 18(4):423–431.

Brembs, B., Button, K., and Munafò, M. (2013). Deep impact: unintended consequences of journal rank. *Frontiers in human Neuroscience*, 7.

Brown, W. (2015). *Undoing the demos: Neoliberalism's stealth revolution*. MIT Press.

Caplar, N., Tacchella, S., and Birrer, S. (2017). Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nature Astronomy*, 1:141.

Chetty, R., Grusky, D., Hell, M., Hendren, N., Manduca, R., and Narang, J. (2017). The fading American dream: Trends in absolute income mobility since 1940. *Science*, 356(6336):398–406.

Chorus, C. and Waltman, L. (2016). A large-scale analysis of impact factor biased journal self-citations. *PLoS One*, 11(8):e0161021.

Clauset, A., Arbesman, S., and Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Sci Adv*, 1(1):e1400005.

Cobham, A., Schlögl, L., and Sumner, A. (2016). Inequality and the Tails: the Palma Proposition and Ratio. *Global Policy*, 7(1):25–36.

Csiszar, A. (2017). How lives became lists and scientific papers became data: cataloguing authorship during the nineteenth century. *Br J Hist Sci*, 50(1):23–60.

Damgaard, C. and Weiner, J. (2000). Describing inequality in plant size or fecundity. *Ecology*, 81(4):1139–1142.

Dickson, D. (1984). *The New Politics of Science*. University of Chicago Press.

Duménil, G. and Lévy, D. (2011). *The Crisis of Neoliberalism*. Harvard University Press.

Fausto-Sterling, A. (2008). *Myths Of Gender: Biological Theories About Women And Men, Revised Edition*. Basic Books.

Frey, B. S. (2009). Economists in the PITS? *International Review of Economics*, 56(4):335–346.

Galison, P. L. and Stump, D. J. (1996). *The disunity of science: Boundaries, contexts, and power*. Stanford University Press.

Garfield, E. (1955). Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas. *Science*, 122(3159):108–111.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178(4060):471–479.

Ginsberg, B. (2011). *The Fall of the Faculty*. Oxford University Press.

Gygli, S., Haucap, J., Ruffner, J., Sturm, J.-E., and Südekum, J. (2017). Handelsblatt-ranking 2.0: Wissenschaftlicher, flexibler, transparenter. http://www.oekonomenstimme.org/artikel/2017/06/handelsblatt-ranking-20-wissenschaftlicher-flexibler-transparenter. Accessed: 2017-06-20.

Halbert, D. (2005). *Resisting Intellectual Property*. RIPE Series in Global Political Economy. Taylor & Francis.

Hammarfelt, B., de Rijcke, S., and Rushforth, A. D. (2016). Quantified academic selves: the gamification of research through social networking services. *Information Research*, 21(2):21–2.

Hand, E. (2012). Extra scrutiny for 'grandee grantees'. *Nature*, 482(7386):450–451.

Harding, S. (1986). *The Science Question in Feminism*. Cornell paperbacks: Philosophy of science/women's studies. Cornell University Press.

Harding, S. (1991). *Whose Science? Whose Knowledge?: Thinking from Women's Lives*. Cornell paperbacks. Cornell University Press.

Heller, M. A. and Eisenberg, R. S. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science*, 280(5364):698–701.

Hubbard, R. (1990). *The politics of women's biology*. Rutgers University Press.

Kaiser, J. (2017a). Data check: Critics challenge nih finding that bigger labs aren't necessarily better. *Science*, 356(6342):997–997.

Kaiser, J. (2017b). Updated: NIH abandons controversial plan to cap grants to big labs, creates new fund for younger scientists. *Science*.

Katz, Y. (2015). Who Owns Molecular Biology? *Boston Review*.

Keenan, B. (1966). *Science and the University*. Columbia University Press.

King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J., and West, J. D. (2016). Men set their own cites high: Gender and self-citation across fields and over time. *arXiv preprint arXiv:1607.00376*.

Leslie, S. (1993). *The Cold War and American Science: The Military-industrial-academic Complex at MIT and Stanford*. Columbia University Press.

Li, D. and Agha, L. (2015). Big names or big ideas: Do peer-review panels select the best science proposals? *Science*, 348(6233):434–438.

Li, D., Azoulay, P., and Sampat, B. N. (2017). The applied value of public investments in biomedical research. *Science*, 356(6333):78–81.

MacKenzie, D. and Millo, Y. (2003). Constructing a market, performing theory: The historical sociology of a financial derivatives exchange. *American Journal of Sociology*, 109(1):107–145.

Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810):56–63.

Mirowski, P. (2011). *Science-Mart: Privatizing American science*. Harvard University Press.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. U.S.A.*, 109(41):16474–16479.

National Institutes of Health (2016a). Awards of R01 and R21 grants in FY 2016. https://gsspubssl.nci.nih.gov/blog/articles?funding_patterns/2016. Accessed: 2017-06-06.

National Institutes of Health (2016b). HHS FY2016 Budget in Brief. https://www.hhs.gov/about/budget/budget-in-brief/nih/index.html#budget. Accessed: 2017-06-04.

National Institutes of Health (2016c). High Impact Journals: Superfund Research Program. https://tools.niehs.nih.gov/srp/publications/highimpactjournals.cfm. Accessed: 2017-06-08.

National Institutes of Health (2017a). New NIH Approach to Grant Funding Aimed at Optimizing Stewardship of Taxpayer Dollars. https://www.nih.gov/about-nih/who-we-are/nih-director/statements/new-nih-approach-grant-funding-aimed-optimizing-stewardship-taxpayer-dollars. Accessed: 2017-06-25.

National Institutes of Health (2017b). Research Supplements to Promote Diversity in Health-Related Research. https://grants.nih.gov/grants/guide/pa-files/PA-08-190.html. Accessed: 2017-07-04.

National Science Board (2014). *Science and Engineering Indicators 2014*. National Science Foundation, Arlington, VA.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *Quarterly Journal of Economics*, 122(3):1067–1101.

Niederle, M. and Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, 3(1):601–630.

Perry, M. and Reny, P. J. (2016). How to count citations if you must. *American Economic Review*, 106(9):2722–41.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Rusert, B. (2017). *Fugitive Science: Empiricism and Freedom in Early African American Culture*. NYU Press.

Sampat, B. and Williams, H. L. (2015). How do patents affect follow-on innovation? Evidence from the human genome. Technical report, National Bureau of Economic Research.

Schekman, R. (2013). How journals like Nature, Cell and Science are damaging science. *The Guardian*, 9:12–23.

Sen, A. (1973). *On Economic Inequality*. Clarendon paperbacks. Clarendon Press.

Sheltzer, J. M. and Smith, J. C. (2014). Elite male faculty in the life sciences employ fewer women. *Proc. Natl. Acad. Sci. U.S.A.*, 111(28):10107–10112.

Sinatra, R., Wang, D., Deville, P., Song, C., and Barabási, A.-L. (2016). Quantifying the evolution of individual scientific impact. *Science*, 354(6312):aaf5239.

Stidham, R. W., Sauder, K., and Higgins, P. D. (2012). Using bibliometrics to advance your academic career. *Gastroenterology*, 143(3):520–523.

Stigler, G. J. (1964). The intellectual and the marketplace. *Society*, 1(5):11–12.

Taylor, P. and Braddock, R. (2007). International university ranking systems and the idea of university excellence. *Journal of Higher Education Policy and Management*, 29(3):245–260.

Tyfield, D. (2013). *The Economics of Science: A Critical Realist Overview: Volume 1: Illustrations and Philosophical Preliminaries*. Ontological Explorations. Taylor & Francis.

van Dijk, D., Manor, O., and Carey, L. B. (2014). Publication metrics and success on the academic job market. *Curr. Biol.*, 24(11):R516–517.

Veblen, T. (1918). *The Higher Learning in America: A Memorandum on the Conduct of Universities by Business Men*. B. W. Huebsch.

Wang, D., Song, C., and Barabási, A.-L. (2013). Quantifying long-term scientific impact. *Science*, 342(6154):127–132.

Weisman, R. (2017). Scientists worry about plan to cap individual labs' federal funding. *Boston Globe*.

Wenneras, C. and Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, 387(6631):341–343.

Zerhouni, E. A. (2006). Research funding. NIH in the post-doubling era: realities and strategies. *Science*, 314(5802):1088–1090.