New facts and data about professors and their research

Kyle Myers[*†] Wei Yang Tham[*†] Jerry Thursby[†] Marie Thursby[†‡]
Nina Cohodes[†] Karim Lakhani[*†] Rachel Mural[§] Yilun Xu[†]

November 30, 2023

[*Draft — comments welcome*]

Abstract

We introduce a new survey of professors at roughly 150 of the most research-intensive institutions of higher education in the US. We document seven new features of how research-active professors are compensated, how they spend their time, and how they perceive their research pursuits: (1) there is more inequality in earnings within fields than there is across fields; (2) institutions, ranks, tasks, and sources of earnings can account for roughly 45% of the total variation in earnings; (3) there is significant variation across fields in the correlations between earnings and different kinds of research output, but these account for a small amount of earnings variation; (4) measuring professors' productivity in terms of output-per-year versus output-per-research-hour can yield substantial differences; (5) professors' beliefs about the riskiness of their research are best predicted by their fundraising intensity, their risk-aversion in their personal lives, and the degree to which their research involves generating new hypotheses; (6) older and younger professors have very different research outputs and time allocations, but their intended audiences are quite similar; (7) personal risk-taking is highly predictive of professors' orientation towards applied, commercially-relevant research.

[*]Harvard Business School
[†]D[3] Institute and the Laboratory for Innovation Science at Harvard
[‡]Georgia Institute of Technology
[§]Harvard Kennedy School
Corresponding author: kmyers@hbs.edu. This project received financial support from The Alfred P. Sloan Foundation and the Harvard Business School and excellent research assistance from Kate Powell. We are grateful for comments from Teresa Amabile, Peter Schiffer, and many others.

1 Introduction

Researchers are at the core of economic growth. The quantity and quality of their new ideas are what enable technological change (Jones 2022). In the U.S., the most common employer of Ph.D.-level researchers is academic institutions,[1] and a long body of evidence has documented the importance of academic science for industrial progress.[2] Amidst recent concerns of declining R&D productivity (Jones 2009; Bloom et al. 2020) and an increasing division of labor between industrial firms and professors in academia (Arora et al. 2020), it has become increasingly important to understand professors' roles in the scientific workforce.

Once dominated by philosophers (e.g., Popper 1934; Kuhn 1962), the "science of science" has become increasingly populated by quantitative analyses by economists, sociologists, physicists, and other scholars looking to turn their empirical tool-kits inwards (Azoulay et al. 2018; Fortunato et al. 2018). However, most of these empirical studies make use of a small number of data sources that are byproducts of conducting science, which were often not designed as intentional sources of data to study science. Inputs are usually measured using federal grant databases, which covers only a fraction of funding flows and leaves unobserved one of the most important input in the scientific production function, researchers' time. Outputs are usually measured using publication databases, which generates serious measurement error for any researcher whose output is not codified in print and makes across-field comparisons difficult to interpret given the range in publication norms. Many other important aspects of researchers' work and lives are left only partly visible at best: their professional position (e.g., tenure status, administrative duties); their incentives (e.g., sources of income); and the objectives of their research (e.g., their intended audience).[3]

In this paper, we document a new nationally representative survey of research-active professors at roughly 150 of the largest institutions of higher education in the US. The population and sample includes professors from all fields of science, broadly construed: humanities and related fields; medicine and health; natural, social, and mathematical sciences. We focus on professors who report a non-zero amount of their time being spent on research activities (95% of respondents). The survey instrument includes a number of

[1] According to the 2021 National Science Foundation Survey of Doctoral Recipients: 42% of Ph.D. scientists and engineers work at educational institutions; 37% work at private, for-profit businesses; 15% work at non-profit or government agencies; and the remainder are employed by other organization types or are self-employed (National Science Foundation 2021).

[2] See, for example, Cohen (2010) or Perkmann et al. (2021) for excellent reviews on this topic.

[3] Still, impressive efforts are still underway to improve the fidelity and interoperability of existing meta-science datasets (e.g., Marx and Fuegi 2020; Lin et al. 2023, UMETRICS at https://iris.isr.umich.edu/, and the Innovation Information Initiative at https://iii.pubpub.org/datasets).

novel elements related to professors' rank and tenure status, time use, funding, salaries and sources thereof, the nature of their research, and a battery of socio-demographic and household-related factors.

Our approach follows a long line of prior work that has used surveys of researchers to uncover their otherwise unobservable features, choices, preferences, or beliefs (e.g., Levin and Stephan 1991; Fox and Stephan 2001; Stern 2004; Sauermann and Cohen 2010; Curty et al. 2017; Levecque et al. 2017; Shortlidge and Eddy 2018; Ganguli and Gaulé 2019; Cohen et al. 2020; Philipps 2022). However, unlike many prior surveys of researchers, a guiding principle of our effort was "breadth over depth" such that many design choices reflect an objective of shedding new light on features of this market that have been largely ignored by empiricists. Our hope is that the summary statistics and correlations in this survey will spark more detailed, focused, and rigorous investigations into the causal effects underlying the patterns we see.

Before reporting the major findings of the survey, we describe the population, sampling methodology, recruitment protocol, and summary statistics. To test for representativeness, we use multiple dimensions of data that are observable for respondents and non-respondents. At the institutional level, we use data from the National Science Foundation's HERD survey (National Science Foundation 2023) to show that respondents come from institutions that receive relatively equal amounts of research funding from a variety of sources. At the individual level, we use data from publication and grant records to show a high degree of similarity in these metrics between respondents and non-respondents. The only key difference between our sample and the population is an under-response from professors at medical schools, which we discuss further below. We sometimes observe other statistically significant degrees of non-response bias, but the practical magnitudes are often relatively small and on the scale of 5 percent of means. Additionally, as one test of attention and honesty, we match a sub-sample of respondents to their publicly reported salaries and find a high degree of alignment.

We document seven key new findings about the research-active professor workforce:

> Finding 1: There is much more inequality in earnings within fields than across. This holds true at the household level because professors exhibit positive assortative matching.
>
> Finding 2: Differences in institutions, faculty ranks, tasks, and sources of earnings can account for roughly 45% of the variation in earnings across professors.
>
> Finding 3: There are significant differences across fields in the implied payoff to pro-

ducing observable research output (e.g., earnings per publication). However, research output and payoff differences can only account for 2–9% of the variation in earnings across professors.

Finding 4: Professors with higher gross output (i.e., annual publications) are not always more productive on a per-research-hour basis because of substantial variation in professors' time allocations. This is especially true for non-tenure-track professors.

Finding 5: Three of the strongest predictors of professors' beliefs about the riskiness of their research are: (1) the share of their time spent fundraising; (2) their risk-taking in their personal lives; and (3) their orientation towards generating new hypotheses with their research (as opposed to testing hypotheses).

Finding 6: Older professors have different intended research outputs than younger professors (i.e., focusing on books as opposed to journal articles), but their intended audiences are the same. Administrative duties exhibit a rise and fall over professors' careers, often with a discontinuous increase after receiving tenure, which can explain a large fraction of the decline in research hours post-tenure.

Finding 7: Professors' position on the basic–applied spectrum can be proxied with the intended output and audience of their research; more applied "Edison-like" professors, whose output is more likely to be tools and products and whose audience is more likely to be businesses and policymakers, report a higher willingness to take risks in their personal lives.

We do not report any causal effects in this paper, taking all equilibrium correlations as being representative of some combination of treatment and selection effects. Furthermore, the data currently exists only as a cross-section. Thus, variation across professors of different ages reflects both temporal dynamics as well as changes to the composition of this workforce.

In some cases, we report decompositions based on the R^2 and partial-R^2 statistics based on simple linear models.[4] Our goal with these exercises is to determine how much the variation in the focal outcome can be described by the covariates, which indicates the extent to which any treatment or selection effects are important along those dimensions. We also report the results of "observational regressions", a term which we will use to describe regressions of professors' features on a set of possibly endogenous covariates. Similarly, we

[4]The partial-R^2 for a given (independent) variable is the proportion of variation (in the dependent variable) explained by that variable that cannot be explained by the other (independent) variables.

sometimes employ ML-based covariate selection methods to identify variables with predictive power (e.g., Meinshausen and Bühlmann 2010). The broad patterns that emerge provide useful views of equilibrium relationships and motivate new hypotheses about the incentives facing academic researchers.

An anonymous version of a subset of the individual-level data is publicly available here: [data preparation still in progress]. We hope the results reported in this paper, and the public data, will spark further investigation into the academic research workforce. The rest of the paper is organized as follows: Section 2 describes the survey methodology and some summary statistics; Section 3 walks through our key new findings; Section 4 concludes with a discussion.

2 Methodology: Population, sampling, and survey

2.1 Population and sampling

Our target population is US professors who conduct research at major institutions of higher education. We identify this population by selecting the 158 largest institutions in the U.S. per their total R&D funding reported in the National Science Foundation's 2019 Higher Education R&D (HERD) survey (National Science Foundation 2023). We hired individuals to manually collect the emails of professors from these universities' websites.[5] We identified these individuals as people listed on institution's website with the word "professor" in their title, recording their title as well as information on each professor's name, program and/or department and/or college, and professorial rank. Our requirement of the word "professor" in the title was driven by the logistics of data collection.[6] Table 1 reports summary statistics for the institutions included in the population based on variables sourced from websites as well as the HERD survey. Figure B1 illustrates some joint distributions of fields, institutions, and ranks showing a significant amount of heterogeneity in the organizational structures.

Our sampling process was as follows. Based on the information gathered, we classified these emails into one of twenty fields of study and one of four ranks (assistant, associate, full or emeritus, and adjunct or other). We then sent an e-mail invitation to a randomly-

[5] See Appendix A for more.

[6] It is a simple, observable feature to rule identifying individuals in and out of population. Anything beyond this proved too complicated in our data collection process. An important question is to what extent we miss relevant individuals without this moniker. We cannot arrive at a conclusive estimate here since there is no clear definition of relevance with which to benchmark. As one potential benchmark, data from the US National Center of Education Statistics, indicates that only a few percentage points of these institutions' full-time instructional staff are not considered "faculty" (US Department of Education 2021).

selected half of the e-mails within each field-rank cell. These e-mails were distributed from October 2022 to March 2023.

The population consisted of 262,344 unique e-mails. We e-mailed a total of 130,785 individuals and 4,357 (3.33%) completed the survey.[7] We then restrict the sample to the 4,145 individuals (95.1% of respondents) who reported being a professor and spending a non-zero amount of time on research. Our final sample consists of professors from engineering, math and related sciences (699), the natural sciences (655; e.g., biology, chemistry, physics), social sciences (863; e.g., economics, political science, psychology, sociology), humanities and related science (783; e.g., art, history, education, linguistics), and health or medical sciences (1,145; e.g., schools of medicine or public health). These five aggregate groupings of fields were chosen partly based on the results of a principal component analysis to identify fields where professors responded to the survey similarly. Table B1 reports the results of a single-component PCA based on the entire survey, averaging at the field level. We then aggregated fields together by primarily relying on this score, with some minor adjustments to align with our understanding of these fields.[8] We sometimes use these five aggregate groupings of fields in our discussion and empirical analyses given the small sample sizes within the narrower field definitions.

2.2 Potential survey biases

Ideally, our respondents would report all answers accurately and their responses would reflect the preferences and characteristics of the full population. We can never formally test this, but we can take some steps to investigate the possibility of inattention and non-response bias and, in the case of non-response bias, possibly account for it.

First, we use our internally collected data on professors' fields and ranks to test for differences between the respondent sample and the population. Appendix A.2 shows that our respondents are slightly more likely to be full professors and less likely to be adjunct, clinical or other professors compared to the population. We also see a significant under-response from the medical and health sciences relative to other fields that is sizeable. Overall, this suggests that the results reported here may be less generalizable to the full spectrum of professors across medical schools or those in adjunct or clinical professor tracks.

[7] This response rate is more than twice what has been obtained from sourcing academic researcher contacts from the corresponding author data contained within the publication record (e.g., Myers et al. 2020).

[8] For instance, we group Medical School based professors with those in other Medicine- or Health-related fields, and we also assign sociology to the aggregate field of social science despite the market difference in average PCA score for this field.

Second, in Appendix A.2, we compare our respondents to the population according to the HERD survey, which reports on the amounts of R&D funding flowing to each of the institutions in our population. When we examine measures including total funding amounts, funding by source, and by type, we do find some evidence that professors who complete our survey are located at institutions with lower-than-average amounts of funding. However, graphical illustrations and statistical tests of these comparisons show that the difference is relatively small in economic magnitude. We have good representativeness over the full distribution, and the average difference in funding amounts between the institutions of respondents and non-respondents is generally in the range of 4–6%.

We are also able to merge roughly two-thirds of the population to their records in the Dimensions database (Digital Science 2018) using a fuzzy, name- and affiliation-based merging process. This allows us to compare individual-level grant input and publication output data across our population and respondents. This exercise is reported in Appendix A.2. Again, we find economically small differences between the full population and respondents, most of which are not statistically significant. Our respondents' average publication output rates, field-normalized citation rates, and grant receipt rates and amounts are all within a few percentages of what is observed in the full population. Likewise, graphical investigations show strong overlap in the support of these variables. Overall, our respondent sample appears very similar to the population along many observable dimensions.

As a test of researchers' attention and their willingness to report truthfully, we can compare their self-reported salaries to their publicly-reported salaries for the subset of researchers at institutions that make such data public. To do so, we manually traced respondents at 89 institutions with public salary data to their records in these public sources. During this match, we used our data on researchers' e-mail addresses and institutional affiliations to maximize fidelity of the match. Still, there is likely still nonzero measurement error due to both (1) manual errors in the name merging process, and (2) our inability to perfectly confirm that the self- and publicly-reported salaries were referring to the same year of employment.

Appendix A.2 plots the relationship between these two sources of salary data. The correlation between the two is 0.75, and for roughly 75% of observations the difference between the self- and publicly-reported salary is less than 30%. This suggests the vast majority of respondents are responding truthfully along this dimension.

2.3 Sample summary statistics

Table 2 reports summary statistics for a majority of the components solicited in the survey. The majority of respondents are full professors (41%) followed by roughly equal proportions of assistant (27%) and associate professors (23%) with the remainder being adjunct, clinical, or other types of professors (9%). The distribution across aggregate fields is relatively even, although we imposed these field groupings in a way that sought approximate balance in sub-sample sizes.[9] Most respondents are tenured (59%), with another 22% still on the tenure-track and 20% not on the tenure track at all. The average number of years since tenure for those who are tenured is approximately 15 years. For those not on the tenure track, the average contract length is about 2.5 years, and for those pre-tenure, the average years until their tenure evaluation is about 2.6 years.

On average, individuals work roughly 50 hours per week, which is nearly eight hours more than the national average for full-time workers (US Bureau of Labor Statistics 2023). The average annual income is approximately $165,000, which is approximately the 93rd percentile in the US (US Census Bureau 2023). The standard deviation is nearly $90,000, indicating a significant amount of heterogeneity, which we return to below.

The average 5-year guaranteed research funding is roughly $85,000 per year, but with a very high standard deviation ($200,000 per year), suggesting tremendous heterogeneity across researchers. The same is true of researchers' expected fundraising, which is on the same order of magnitude as their current guarantees.

The majority of the individuals report being White (79%), followed by Asian (12%), with these groups being slightly (Whites) and substantially (Asians) over-represented relative to their shares of the full US population (US Census Bureau 2023; note: we allowed respondents to report multiple race/ethnicities). The representation of Black, Hispanic, and other ethnicities is relatively low at 3%, 6%, and 4% respectively. Nearly 25% of professors are non-US born. This is nearly double the rate of the full US population (US Census Bureau 2023), which is yet another signal of the importance of immigration for the US research enterprise (e.g., Kerr 2020).

The survey also includes a battery of questions related to the nature of professors' research. First, respondents are asked to rate on a scale of 0 to 10 whether their research is more about generating or testing hypotheses.[10] In addition, respondents are asked to

[9] See Appendix Table A2 for the groupings of the twenty narrower fields into these aggregate fields. The classification of fields into these broader groupings is chosen out of simplicity and with an understanding that it would be impossible to please all professors in terms of how the groups are constructed.

[10] The specific question is: "Use the scale below to rate whether the overall objective of your current research is more to generate new theories and hypotheses, or is more to test existing theories and hypotheses, or is somewhere in between."

report the intended outputs and audiences for their research.[11] Table 3 reports the summary statistics for these variables describing the nature of professors' research.

3 Findings

Finding 1: *There is much more inequality in earnings within fields than across. This holds true at the household level because professors exhibit positive assortative matching.*

Variation within and across fields in academic researchers' earnings is a regular conversation topic within the halls of most universities. But the difficulties of systematically matching researchers' fields to their earnings has limited investigations into this variation except in specific cases (e.g., Mohanty et al. 1986; Hamermesh 1988; Bellas 1997; Ehrenberg et al. 1998; Baker et al. 2023).[12] Furthermore, most existing data on researchers' earnings leave the remainder of the household (e.g., spouses' earnings) untouched despite the importance of the household as an economic unit (Chiappori and Mazzocco 2017).[13]

Figure 1 Panel (a) reports field-level average self-reported total annual earnings. To simplify comparisons, we construct these averages using professors' predicted earnings if the full sample was at the associate professor rank.[14] Professors' individual earnings span lower values of approximately $130,000 per year in fields such as the humanities, communication, agriculture, and education. Earnings in the highest-paying fields, economics, medicine, law, and business are roughly $200,000 per year.

How much variation in earnings do these field-level averages hide? To get a sense as to the variation in earnings both across and within fields, Panel (b) of Figure 1 plots three alternative individual-level Gini coefficients for the sample and, for comparison, the Gini

[11] Options for intended output are: publications, books, tools (e.g. data, software, instruments), or practical applications (products, patents, policies). Options for intended audience are: other academics, policymakers, businesses, or the general public. Respondents whether each of these options is their intended audience or output "Never or rarely", "Sometimes", or "Most or all of the time". Options are coded as 0, 1 or 2.

[12] For some prior work on professors' earnings, with results that are relevant to this and other findings in this paper, see: Tuckman and Tuckman (1976); Marsh and Dillon (1980); Mohanty et al. (1986); Fairweather (1993); Langton and Pfeffer (1994); Hearn (1999); Perna (2002); Ehrenberg et al. (2006); Melguizo and Strober (2007).

[13] We're referring to pre-tax annual earnings of any sources here, decomposing the sources of earnings below. For simplicity, the survey question that solicited earnings did not belabor the distinction between "earned" and "non-earned" incomes, which leaves open the possibility that some respondents reported non-earned income. Still, such income likely would've been reported via the "other" source category, which accounts for only 5% of reported earnings on average.

[14] Furthermore, Figure 1 reports post-shrinkage means using the empirical Bayes shrinkage methodology of Chandra et al. (2016) in order to adjust for differences in sub-sample sizes across fields.

coefficient for the full US population of full-time workers per recent estimates from the US Census Bureau (Guzman and Kollar 2023). The overall Gini coefficient for professors' own earnings is roughly 0.27, which is significantly lower than the 0.43 observed across all US workers (Guzman and Kollar 2023) and is comparable to countries such as Belgium, the Netherlands, and Iceland (World Bank 2023). Appendix Figure B2 reports additional views of earnings variation. Interestingly, there is very little variation in the within-field Gini coefficient across fields, and there is no significant relationship between field-level average earnings and within-field inequality.

To compare inequality in earnings within fields, we estimate a separate Gini coefficient for each field and report the average of these within-field estimates. To compare inequality in earnings across fields, we take the average salary within each field (held constant at associate professor rank) and calculate the Gini coefficient assuming there is one representative professor in each field who earns this (average) amount. Inequality in professors' earnings is much higher when focusing on the within-field variation (average Gini coef.=0.24, s.d.=0.02) compared to the variation in field-level averages (Gini coef.=0.10).

As evidenced by the strong positive relationship between individual and household earnings (Panel c), there is clear evidence of positive assortative matching amongst research professors. Furthermore, it appears that the degree of this matching is relatively similar across fields; most field-level averages are relatively close to the linear fit of the relationship between own and household earnings. To illustrate the role of this assortative matching on household-level earnings variation, Figure 1 Panel (d) reports the same Gini coefficients as Panel (b), this time based on household-level earnings. As expected given the positive matching, the pattern remains the same. Compared to the US population of households (Gini coef.=0.49), earnings across professors' households is more equal (Gini coef.=0.31) with most of the variation driven by earnings differences within fields.[15]

There is much more inequality in earnings within fields than there is across fields, a novel fact that is of both inherent and policy relevance to the market for academic research labor. In many of the following analyses, we attempt to identify some of the sources of this variation based on how professors spend their time and what outputs they produce.

> **Finding 2** *Differences in institutions, faculty ranks, tasks, and sources of earnings can account for roughly 45% of the variation in earnings across professors.*

Understanding the different incentives researchers face is key to understanding how they

[15]The differences between household- and individual-level Gini coefficients in the full US population and professors is relatively similar on the scale of 5 p.p. and 15%. This suggests the positive assortative matching observed amongst professors is similar to that observed in the full population (Greenwood et al. (2014)).

allocate their time. How much of the variation in professors' earnings can be explained by observable differences in their work? As evidenced in Appendix Figures B1 and B2, there is considerable heterogeneity in professors' ranks, in how they spend their time, and in the sources of their earnings.

To better understand earnings differences, we first estimate rank- and institution-specific average earnings. The average annual earnings by rank are as follows: assistant, 131,670 (s.d.= 61,980); associate, 139,752 (s.d.= 64,879); full, 204,985 (s.d.= 98,970); other, 98,329 (s.d.= 81,345). Figure 2 Figure 2 Panel (a) shows that there is considerable variation in average earnings across institutions. There is some right skew to this distribution, but it is much less skewed than the distribution of firm-level average earnings in the US economy. Here, the 75th:25th percentile ratio is roughly 1.3 and the 90th:10th percentile ratio is roughly 1.8. In the broader economy, these ratios tend to be closer to 2 and 5, respectively (Song et al. (2019)). Notably, these are not adjusted for any observable differences across professors.

Next, we estimate observational regressions that relate professors' earnings to different features of their work. First, we regress their earnings on their time spent on six different tasks (research, fundraising, teaching, administration, clinical, or other). Next, we regress their total earnings on the source of their earnings (base salaries, grant-covered, supplemental earnings from their primary institution, clinical work, or other). These regressions are of the form: $Y_i = \alpha + \sum_j X_i^j \beta^j + \epsilon_i$, where Y_i is professor i's earnings and X_i^j is the vector of either the professor's hours spent on each tasks indexed by j, or the percentage points of the professor's earnings due to each source indexed by j. The results of these two regressions (the estimates of β^j) represent the implied marginal wages of each task (or income source) holding the amount of time spent on the other tasks (or share of income from other sources) fixed. Figure B3 illustrates the variation in these metrics across fields.

There are substantial differences in the implied returns to different tasks (Figure 2, Panel b). Clinical work is associated with the most earnings, nearly $75 per hour. This is to be expected since most "clinical" work performed by professors involves medical care delivery at academic medical systems. Research, administration, and "other" tasks have implied marginal wages of roughly $30–50 per hour. Fundraising is less positively associated with earnings (roughly $12 per hour), while additional time spent on teaching activities is associated with negative marginal earnings. Clearly this reflects a significant selection effect whereby positions with larger teaching responsibilities also have lower salaries. Still, this observational regression makes the importance of this selection very clear.

Overall, Figure 2 Panel (c) echoes the findings of Panel (b). In short, professors that have a larger fraction of their earnings coming from sources besides their base salaries have

larger earnings levels. This is especially true for those undertaking clinical work.

We decompose the variation in earnings across professors more formally in Figure 2 Panel (d).[16] Institution and rank averages explain a considerable amount of within-field variation, as do professors' tasks and earnings sources. Field-level averages alone describing roughly 10% of earnings variation (Column 1), with the full set of covariates explaining roughly 55% of earnings variation (Column 5). This suggests that nearly 45%(=55%-10%) of earnings variation is due to these institutional and position specific features of professors' work and how compensation varies across these dimensions. It also implies that 45%(=100%-55%) of the total variation may be due to other factors or idiosyncratic differences. We focus on one of those possible factors, research output, next.

> **Finding 3** *There are significant differences across fields in the implied payoff to producing observable research output (e.g., earnings per publication). However, research output and payoff differences can only account for 2–9% of the variation in earnings across professors.*

How do different fields implicitly reward different types of research progress? And how much of the earnings variation across fields is due to differences in research output or the different way that output is rewarded? With a large fraction of our sample (78%) matched to their grant and publication histories, we can explore these questions further.[17]

Figure 3 Panel (a) reports the results of an observational regression of earnings on research inputs and outputs. To allow for some heterogeneous returns across fields, we interact the research metrics with indicators for each of the five aggregate fields in our sample. For each of the three metrics, we standardize the variation within each aggregate field so that a unit increase in each metric corresponds to a one s.d. increase per the distribution within the aggregate field.[18]

We find significant differences across fields in terms of the implied earnings per research metric. All fields appear to implicitly reward grant funding, with a one s.d. increase being associated with somewhere between \$5,00–15,000 in additional annual earnings, although we only estimate a precise relationship within the medical and natural sciences. There are positive implied returns to publication output across all fields as well, although there

[16] Note that in a multivariate model, the sum of the partial-R^2 statistics for each variable need not sum to one.

[17] See Appendix A.5 for more on this matching process and comparisons of the matched and un-matched sub-samples.

[18] The publication and citation counts were already normalized within the narrower (20) fields, so this more aggregate standardization is less relevant for these two metrics, but preserves consistency across the three metrics.

are large differences in magnitudes and the relative weight placed on publications versus citations. Humanities and social sciences appear to put a much larger value on citations relative to publications — the implied returns to citation counts are 2–3× the returns to publication counts. Engineering, medical, and natural sciences appear to operate in the opposite way. Overall, these point to very different norms across fields in terms of how researchers are rewarded for output.

We again decompose earnings heterogeneity, this time focusing on these research metrics, in Figure 3 Panel (b).[19] Without accounting for any other covariates, differences in these research metrics appear roughly equally as important as field-level differences, with both explaining roughly 9% of earnings variation. However, when we include the full set of institution, faculty rank, task, and earnings source covariates that we explored in Figure 2 (Columns 6–7), we find that research output to be much less important. Now, research metrics can account for only 2%(=58%-56%) of earnings variation.

When paired, with the prior finding, it appears that the traditional metrics of research output often used in the science of science are much less related to professors' earnings than other attributes of their job. This doesn't necessarily imply anything about the validity of these metrics as indicators of scientific progress. Nor does this imply anything about the optimality of researchers' payoffs from conducting research. However, it illustrates that there is potentially a large gap between the way in which professors are financially compensated and the way in which scholars in the science of science field might characterize their performance. We turn more specifically to the notion of research productivity next.

> **Finding 4**: *Professors with higher gross output (i.e., annual publications) are not always more productive on a per-research-hour basis because of substantial variation in professors' time allocations. This is especially true for non-tenure-track professors.*

Despite time being a key scientific input, the difficulty of observing even a proxy of researchers' time allocations has severely limited our understanding of the labor component of the scientific production function. This has led most studies to assume that researchers all have access to the same amount of time per year and either explicitly or implicitly use researchers' gross output per year as a measure of productivity.

But of course, professors generally balance multiple roles in a university, only one of them being a researcher. Teaching and advising responsibilities, administrative duties, and

[19]Recall, this is based only on observations matched to the Dimensions data, which can lead to some discrepancies if compared to Figure 2 Panel (d).

grant-writing tasks all can loom large.[20] Table B2 reports the pairwise correlations for the main categories of time allocation we focus on: research (including supervising others); fundraising for research; teaching or advising (not as a part of their own research); clinical or medical practice; all other activities.[21] Except in the case of fundraising time, which appears to be a partial complement to research time, time spent on all other tasks is associated with a decline in time spent on research.

The key question is whether these other time constraints are allocated in a way that is correlated with researchers' actual scientific productivity. If time constraints (e.g., administrative duties) are often allocated to researchers with lower hourly scientific productivity (i.e., the two are negatively correlated), then researchers' annual and hourly output will be very closely aligned. However, to the extent certain researchers with high hourly productivity have fewer opportunities to conduct research because they face additional time constraints, then the alignment between annual and hourly productivity will begin to deteriorate.[22]

In order to understand how well traditional measures of gross output per year correlate with measures that account for differences in input levels (i.e., output per hour worked), we again focus on the Dimensions-matched sample where we can see professors' publication output. We calculate their publication output on both an annual and hourly basis (per their field-normalized publication counts). The correlation between the two measures is ρ=0.69, which suggests that annual output is indeed informative of hourly productivity, but it may be misleading for some. To get a better view, Figure 4 Panel (a) plots each individual's percentile of annual output on the x-axis and their percentile of hourly output on the y-axis, noting that this compression into percentiles hides the skewed nature of these measures. Researchers below the 45^o line have an annual output that overstates their hourly output, and vice versa.

To further investigate this issue, Figure 4 Panel (b) plots a histogram of the ratio of each professor's percentile on an hourly basis compared to the same on an annual basis. We find about 40% of professors have an annual output percentile within 10% their hourly output percentile (see: the grey shaded bar in the histogram). Another 35% have output

[20] For example, in grant-intensive fields, many have raised concerns that scientists devote too much time to unproductive activities in order to win grants (Gross and Bergstrom 2019; Collison et al. 2021); however, it is difficult to estimate the social value of these efforts since not all time and effort devoted to fundraising is necessarily wasteful (Ayoubi et al. 2019; Myers 2022).

[21] We ask respondents to forecast their weekly hours they will spend on each of these activities over the coming 5-year horizon in hopes of them estimating something close to their steady-state time allocations that are not driven by year-to-year idiosyncrasies.

[22] If the positive correlation between hourly productivity and time constraints was large enough, there could feasibly be no correlation between annual and hourly research output.

percentiles within 20% of each other. There are a significant number of researchers for whom annual output is not a strong proxy for their hourly output. Understanding who these researchers are and whether more resources ought to be allocated to them is a key policy question.

As a first look, we use the stability selection method of Meinshausen and Bühlmann (2010) to identify "important" predictors of the gap between researchers' annual and hourly output (per percentiles).[23] Figure 4 Panel (b) reports the stability selection share for the top ten covariates along with the univariate correlation between each covariate and the ratio of researchers' annual and hourly output. Unsurprisingly, the strongest predictors are variables related to researchers' work hours. Individuals with the most understated hourly productivity are in non-tenure-track, adjunct, or other position, presumably, because they have the largest constraints on their time. Interestingly, there is some evidence that researchers pursuing non-traditional research outputs (that is, not journal articles intended for academics) also appear to have understated hourly productivities. This again may be due to their time allocations focusing more on non-research-specific tasks. Overall, this new view of researchers' time indicates that gross output measures like publications per year may not provide an unbiased view into researchers' true underlying productivity in terms of their ability to convert their actual *research* time into scientific output.

> **Finding 5** *Three of the strongest predictors of professors' beliefs about the riskiness of their research are: (1) the share of their time spent fundraising; (2) their risk-taking in their personal lives; and (3) their orientation towards generating new hypotheses with their research (as opposed to testing hypotheses).*

Discourse about innovation and science policy often asserts that the system overly discourages scientists from taking risks, causing society to miss out on high-impact scientific discoveries and inventions (e.g. Alberts et al. 2014). Most empirical work has focused on this issue relies on ex-post measures of risk-taking based on bibliometric measures.[24] These ex-post measures of publications are clearly limited in their ability to proxy for ex-ante risk-taking by researchers.

To provide a new, alternative view of risk in science, the survey solicits professors' subjective beliefs of their own risk-taking behaviors.[25] Using questions structured in the same

[23] This method proceeds as follows: (1) a random 50% sub-sample is drawn; (2) a standard cross-validation Lasso is used to select the relevant covariates (we use the standard k-fold cross-validation `lasso` program in `Stata` with all default options); (3) Steps (1–2) are repeated 100 times, recording the share of samples each covariate is selected by the Lasso (i.e., the stability selection share).

[24] See Greenblatt et al. (2022) for an effort that includes a wide range of proxies for concepts underlying risk and novelty, and see Carson et al. (2022) and Carson et al. (2023) alternative approaches.

[25] Constructing any sort of field-agnostic (relatively) objective measure of risk-taking that mirrored those

format as the more general risk preference questions of Dohmen et al. (2011), researchers report how risky they think their own research is, as well as how risky they think their peers think their research is (on a scale from 0 to 10, with larger values indicating more risk).[26] Our preferred metric reported throughout this paper is the average of these two responses, which hopefully serves to reduce some of the measurement error inherent to either phrasing of the question. Figure B4 reports the distribution of this (averaged) risk score, which illustrates significant support across the full range of possible values except for the uppermost tail of risk-taking.

We use the ML-based approach of stability selection to identify the covariates that best predict researchers' perceptions about the riskiness of their research. Figure 5 Panel (a) reports the results of this exercise, with Panels (b–d) showing binned scatterplots of researchers' risk perceptions based on three of the top predictors we identify via stability selection: (Panel b) the share of time researchers spend on fundraising; (Panel c) researchers' willingness to take risks in their personal lives; and (Panel d) researchers orientation towards generating (as opposed to testing) hypotheses.

Figure 5 Panel (b) suggests that professors who undertake more fundraising may inherently perceive more risk in their research. This is interesting because "fundraising risk" has not traditionally received much attention due to the difficulties of observing professors' specific funding streams. This may be driven by the fact that professors who spend a large fraction of their time fundraising tend to be in "soft-money" positions where a portion of their salary is derived from their fundraising. [27] Alternatively, it may be that, projects that are more ambitious and risky tend to be projects that also require more resources and hence more fundraising. This points to the need for more work on understanding what risk means to researchers (e.g., Holzmeister et al. 2020): What are the outcomes they care about? How do they perceive risk?

Risk-taking in personal life is another strong predictor of researchers' perceptions of the riskiness in their science. Of course, the simplest explanation here is a survey response bias whereby individuals inflate their risk-taking in both questions and this generates the correlation. However, the pattern is also consistent with researchers' latent risk-aversion being a key determinant of how they pursue their science. To the extent this is true, it suggests that understanding the extensive margin of selection into science and how it may

commonly used in lab experiments (e.g., gambles over outcomes) proved extremely difficult in pilot tests due to the heterogeneity in relevant outcomes. Hence, our more subjective, but much simpler measure.

[26] Pilot interviews with scientists suggested that both approaches would prove useful avenues for soliciting scientists beliefs.

[27] Every percentage point increase in the share of researchers' time spent on fundraising is associated with an additional $1,095 (s.e.=$58) in annual earnings.

screen individuals with higher or lower levels of latent risk-aversion would be a fruitful avenue for future research.

Lastly, we see that researchers who report focusing on generating new hypotheses also tend to report higher perceived risk in their research. This aligns with the idea that it is inherently more difficult to capture the value of a good question compared to a good answer; in other words, hypotheses themselves have stronger public-good attributes than tests of hypotheses. Undertaking projects where your ability to capture the value of your efforts is less certain would likely be perceived as more risky. We dig further into researchers' strategies next.

> **Finding 6** *Older professors have different intended research outputs than younger professors (i.e., focusing on books as opposed to journal articles), but their intended audiences are the same. Administrative duties exhibit a rise and fall over professors' careers, often with a discontinuous increase after receiving tenure, which can explain a large fraction of the decline in research hours post-tenure.*

Besides taking (or avoiding) risks, what exactly are research professors intending to do with their research? Here, we dig into our questions related to the intended outputs and audience of professors' research (see Table 3).[28] Specifically, motivated by early work in the economics of science related to life-cycle effects (Levin and Stephan 1991), we focus on temporal changes across ages and professional experience. An important caveat to reiterate here is that the cross-sectional nature of the survey means that any temporal dynamics reflect both age- or experience-related effects in addition to any selection effects that occur over the life cycle and/or career cycle.

Figure 6 Panels (a–b) plot professors' intended outputs and audiences across the forty years of ages in our sample. Notably, there is a marked evolution in *what* professors are focusing on producing, shifting from a focus on journal articles, materials, or methods in their early decades into a focus on books, products, or services in their later decades (Panel a). However, there is no significant change in *who* professors are focusing their efforts towards (Panel b). This pattern suggests that professors' preferences over their audience are quite stable, but the optimal way of reaching this audience is not. Books as a scientific

[28] All of the questions related to professors' intended output and audience are solicited on a likert scale with three values of frequency ("Rarely", "Sometimes", "Most of the time"), which we convert into a variable valued {0,1,2} (see Table 3). These measures are intended to reflect the underlying *share* of professors' scientific production that is destined for a particular output or audience type. Thus, for all of the following analyses, we assume that professors have a fixed level of intentions in these two dimensions and re-scale each their responses into fractions that the sum to one for each dimension. For example, if a professor reports that all four output types are their intentions "Most of the time", then we assume that the share of their intended output of each type is 1/4.

output have not received much attention by the science of science community, likely because of data limitations.[29]

Figure 6 Panels (c–d) revisit professors' time use, now looking over the life cycle (Panel c) and more narrowly around the tenure evaluation process for those on the tenure track (Panel d).[30] As most experienced professors can attest, age and experience are associated with a clear increase in administrative duties. However, this provides one of the first views of this shift in task composition that allows us to quantify the relative increase in administrative duties relative to the change in professors' time spent on their research. A common result publicized by studies that can observe only publication output across researchers' careers is the marked decline in their publication output after their early years of work, and especially after receiving tenure (Brogaard et al. 2018). Focusing specifically on Panel (d), we can see that the receipt of tenure is associated with a marked increase in administrative duties and, to a much lesser degree, some increases in teaching and other duties. Aggregating these changes together indicates that, in the first ten years post tenure, approximately 80% of the decline in research and fundraising efforts can be explained by the increase in teaching, administrative, and other effort. Here again, much like our prior finding on the differences between annual and hourly output, our ability to observe researchers' time allocations indicates that the "post-tenure glut" in publication may not be any change in productivity on a publication-per-research-hour basis but, to a large degree, may more simply be a decline in input levels. This distinction is important because it speaks to the trade-offs of the institutions of academic science and professors' responsibilities therein.[31]

> **Finding 7**: *Professors' position on the basic–applied spectrum can be proxied with the intended output and audience of their research; more applied "Edison-like" professors, whose output is more likely to be tools and products and whose audience is more likely to be businesses and policymakers, report a higher willingness to take risks in their personal lives.*

Understanding the selection of research topics by professors and the rewards for these choices is crucial to understanding the direction of science. One of the most common

[29]Exceptions include Wanner et al. (1981); Sabharwal (2013); Giménez-Toledo et al. (2016), and in a more general sense Giorcelli et al. (2022).

[30]Here, we group research and fundraising time given their positive correlation as shown in the pairwise time-use correlations of Table B2.

[31]For instance, the ideal distribution of professors' administrative duties over the course of their career will depend on, among other things, how their productivity evolves over their life-cycle, which may follow field-specific patterns (Levin and Stephan 1991; Galenson and Weinberg 2001; Weinberg and Galenson 2019).

approaches to characterizing research is on a spectrum of "basic" to "applied". However, creating a quantitative measure solely based on existing data is challenging. For example, bibliometric measures such as patent citations may have limited validity in some fields where patenting is rare. To provide an alternative view, the survey includes multiple questions related to the nature of professors' research (see Table 3). Each of these questions were designed to both capture different types of scientific outputs and audiences, but also to reflect different dimensions of the basic–applied spectrum as often described.

To combine the information contained in all of these nature-of-research questions, we use Principal Components Analysis (PCA) to estimate a single-dimension, standardized index. Figure 7 Panel (a) reports the results from the PCA. On one end of this spectrum are professors focused on generating hypotheses and writing journal articles for academics, and on the other end of this spectrum are professors focused on testing hypotheses and making tools and products for policymakers, businesses, and the general public. Hence, in the spirit of Stokes's (1997) quadrants, we term this uni-dimensional index the "Bohr-Edison" score. More negative values indicate more "Bohr"-like basic science and more positive values indicate more "Edison"-like applied science.

Figure 7 Panel (b) reports the field-level average Bohr-Edison scores. The ranking is intuitive, with traditionally "theoretical" fields like mathematics and physics scoring more towards the Bohr end of the spectrum, and more applied and technically-oriented fields like agriculture, law, and medicine scoring more towards the Edison end of the spectrum.

One way to conceptualize the basic-applied spectrum through the lens of economics is the ability of the researcher to appropriate the value of their outputs. The application of more basic research is fundamental in nature and may be harder to appropriate, and vice versa as the research becomes more applied. In this vein, Figure 7 Panel (c) plots professors' earnings as a function of their Bohr-Edison score. As expected, we find a strong positive correlation. The correlation reverses at the extreme Edison-end of the spectrum, which may be driven by the fact that some of the most applied professors in the sample come from fields such as education, communication, and agriculture (Figure 7 Panel b) which are fields with some of the lowest average earnings (Figure 1 Panel b).

To understand what variables are most predictive of the Bohr-Edison score, we again use the stability selection approach. Figure 7 (d) reports the top 10 predictors per their stability selection share along with their univariate correlations with the Bohr-Edison score. A noteworthy finding here is that personal risk-taking is one of the best predictors of doing more applied, Edison-like work. This is an example of a pattern that would be difficult to find using ex-post measures of risk aversion (e.g., based on citations or text of publications and patents) because it is difficult to know whether cross-field differences

in such measures are due to underlying professor characteristics or simply reflect field differences (e.g., different citation norms). To the extent the Bohr-Edison score reflects Edison-like entrepreneurship, this finding echoes other work showing that entrepreneurs tend to be more risk tolerant (Van Praag and Cramer 2001; Cramer et al. 2002; Herranz et al. 2015). More generally, this approach to transforming researchers' intentions in terms of the outputs and audience of their science could prove useful in generating observable, ex-ante variation in researchers' positions along the basic-applied spectrum.

4 Discussion

The emergence of large, curated datasets derived from publication and grant records has facilitated a surge of new empirical studies of science. However, there are many important variables that even high-quality administrative datasets do not capture (Stantcheva 2022). In this paper, we document our survey efforts to solicit and codify some of these important but hard-to-observe variables for the academic research workforce: professors' time allocations, their earnings sources, the nature of their research, and their risk aversion. This new survey data, combined with existing datasets, yields new insights into variation amidst US academic researchers both *within* and *across* fields at a national scale.

We are certainly not the first to use survey methods to learn about the academic research workforce. But our approach provides one of the first broad views across the full spectrum of science in modern research universities. We do not report any causal effects here, and we are limited in our ability to precisely disentangle sources of heterogeneity across researchers in many dimensions. But we have highlighted a number of novel features of this workforce.

We take a narrow view of the broader academic research workforce here, focusing only on professors. The dramatic rise in contingent and part-time faculty (Colby 2023) and prevalence of "staff scientists" (Carpenter 2012) at US universities suggest that we clearly are missing some important workers. Targeting these researchers would likely require alternative outreach techniques and new survey instruments, but would certainly be worthwhile. Future survey work can build on our efforts more broadly by investing more resources into recruiting respondents, eliciting preferences with more precise methods, or eliciting a wider spectrum of preferences. Such efforts can continue to provide a complementary view of the science of science.

References

Alberts, B., Kirschner, M. W., Tilghman, S., and Varmus, H. (2014). Rescuing US biomedical research from its systemic flaws. *Proceedings of the National Academy of Sciences*, 111(16):5773–5777.

Arora, A., Belenzon, S., Patacconi, A., and Suh, J. (2020). The changing structure of American innovation: Some cautionary remarks for economic growth. *Innovation Policy and the Economy*, 20(1):39–93.

Ayoubi, C., Pezzoni, M., and Visentin, F. (2019). The important thing is not to win, it is to take part: What if scientists benefit from participating in research grant competitions? *Research Policy*, 48(1):84–97.

Azoulay, P., Graff-Zivin, J., Uzzi, B., Wang, D., Williams, H., Evans, J. A., Jin, G. Z., Lu, S. F., Jones, B. F., Börner, K., et al. (2018). Toward a more scientific science. *Science*, 361(6408):1194–1197.

Baker, M., Halberstam, Y., Kroft, K., Mas, A., and Messacar, D. (2023). Pay transparency and the gender gap. *American Economic Journal: Applied Economics*, 15(2):157–183.

Bellas, M. L. (1997). Disciplinary differences in faculty salaries: Does gender bias play a role? *Journal of Higher Education*, 68(3):299–321.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–1144.

Brogaard, J., Engelberg, J., and Van Wesep, E. (2018). Do economists swing for the fences after tenure? *Journal of Economic Perspectives*, 32(1):179–194.

Carpenter, S. (2012). A Hidden Academic Workforce. Science Magazine. Available from https://www.science.org/content/article/hidden-academic-workforce. Accessed in 2023.

Carson, R. T., Graff Zivin, J., Louviere, J. J., Sadoff, S., and Shrader, J. G. (2022). The risk of caution: Evidence from an experiment. *Management Science*, 68(12):9042–9060.

Carson, R. T., Zivin, J. S. G., and Shrader, J. G. (2023). Choose your moments: Peer review and scientific risk taking. *Mimeo*.

Chandra, A., Finkelstein, A., Sacarny, A., and Syverson, C. (2016). Productivity dispersion in medicine and manufacturing. *American Economic Review*, 106(5):99–103.

Chiappori, P.-A. and Mazzocco, M. (2017). Static and intertemporal household decisions. *Journal of Economic Literature*, 55(3):985–1045.

Cohen, W. M. (2010). Fifty years of empirical studies of innovative activity and performance. *Handbook of the Economics of Innovation*, 1:129–213.

Cohen, W. M., Sauermann, H., and Stephan, P. (2020). Not in the job description: The commercial activities of academic scientists and engineers. *Management Science*, 66(9):4108–4117.

Colby, G. (2023). Data Snapshot: Tenure and Contingency in US Higher Education. American Association of University Professors. Available from https://www.aaup.org/article/data-snapshot-tenure-and-contingency-us-higher-education. Accessed in 2023.

Collison, P., Cowen, T., and Hsu, P. (2021). What we learned doing fast grants. Available from https://future.a16z.com/what-we-learned-doing-fast-grants/. Accessed in 2023.

Cramer, J. S., Hartog, J., Jonker, N., and Van Praag, C. M. (2002). Low risk aversion encourages the choice for entrepreneurship: An empirical test of a truism. *Journal of Economic Behavior & Organization*, 48(1):29–36.

Curty, R. G., Crowston, K., Specht, A., Grant, B. W., and Dalton, E. D. (2017). Attitudes and norms affecting scientists' data reuse. *PLoS ONE*, 12(12):e0189288.

Digital Science (2018). Dimensions [software]. Available from https://app.dimensions.ai. Accessed in 2023, under licence agreement.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Ehrenberg, R. G., McGraw, M., and Mrdjenovic, J. (2006). Why do field differentials in average faculty salaries vary across universities? *Economics of Education Review*, 25(3):241–248.

Ehrenberg, R. G., Pieper, P. J., and Willis, R. A. (1998). Do economics departments with lower tenure probabilities pay higher faculty salaries? *Review of Economics and Statistics*, 80(4):503–512.

Fairweather, J. S. (1993). Faculty reward structures: Toward institutional and professional homogenization. *Research in Higher Education*, 34:603–623.

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., et al. (2018). Science of science. *Science*, 359(6379):eaao0185.

Fox, M. F. and Stephan, P. E. (2001). Careers of young scientists: Preferences, prospects and realities by gender and field. *Social Studies of Science*, 31(1):109–122.

Galenson, D. W. and Weinberg, B. A. (2001). Creating modern art: The changing careers of painters in france from impressionism to cubism. *American Economic Review*, 91(4):1063–1071.

Ganguli, I. and Gaulé, P. (2019). Will the us keep the best and the brightest (as post-docs)? career and location preferences of foreign stem phds. In *The Roles of Immigrants and Foreign Students in US Science, Innovation, and Entrepreneurship*, pages 49–69. University of Chicago Press.

Giménez-Toledo, E., Mañana-Rodríguez, J., Engels, T. C., Ingwersen, P., Pölönen, J., Sivertsen, G., Verleysen, F. T., and Zuccala, A. A. (2016). Taking scholarly books into account: Current developments in five european countries. *Scientometrics*, 107:685–699.

Giorcelli, M., Lacetera, N., and Marinoni, A. (2022). How does scientific progress affect cultural changes? a digital text analysis. *Journal of Economic Growth*, 27(3):415–452.

Greenblatt, W. H., Maity, S. K., Levy, R. P., and Azoulay, P. (2022). Does grant peer review penalize scientific risk taking? Evidence from the NIH. *Mimeo*.

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2014). Marry your like: Assortative mating and income inequality. *American Economic Review*, 104(5):348–353.

Gross, K. and Bergstrom, C. T. (2019). Contest models highlight inherent inefficiencies of scientific funding competitions. *PLoS Biology*, 17(1):e3000065.

Guzman, G. and Kollar, M. (2023). Income in the United States: 2022. *United States Census Bureau: Current Population Reports*.

Hamermesh, D. S. (1988). Salaries: Disciplinary differences and rank injustices. *Academe*, 74(3):20–24.

Hearn, J. C. (1999). Pay and performance in the university: An examination of faculty salaries. *Review of Higher Education*, 22(4):391–410.

Herranz, N., Krasa, S., and Villamil, A. P. (2015). Entrepreneurs, risk aversion, and dynamic firms. *Journal of Political Economy*, 123(5):1133–1176.

Holzmeister, F., Huber, J., Kirchler, M., Lindner, F., Weitzel, U., and Zeisberger, S. (2020). What drives risk perception? A global survey with financial professionals and laypeople. *Management Science*, 66(9):3977–4002.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Review of Economic Studies*, 76(1):283–317.

Jones, C. I. (2022). The past and future of economic growth: A semi-endogenous perspective. *Annual Review of Economics*, 14:125–152.

Kerr, W. R. (2020). The gift of global talent. *Innovation Policy and the Economy*, 20(1):1–37.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

Langton, N. and Pfeffer, J. (1994). Paying the professor: Sources of salary variation in academic labor markets. *American Sociological Review*, pages 236–256.

Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., and Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy*, 46(4):868–879.

Levin, S. G. and Stephan, P. E. (1991). Research productivity over the life cycle: Evidence for academic scientists. *American Economic Review*, 81(1):114–132.

Lin, Z., Yin, Y., Liu, L., and Wang, D. (2023). Sciscinet: A large-scale open data lake for the science of science research. *Scientific Data*, 10(1):315.

Marsh, H. W. and Dillon, K. E. (1980). Academic productivity and faculty supplemental income. *Journal of Higher Education*, 51(5):546–555.

Marx, M. and Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal*, 41(9):1572–1594.

Meinshausen, N. and Bühlmann, P. (2010). Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(4):417–473.

Melguizo, T. and Strober, M. H. (2007). Faculty salaries and the maximization of prestige. *Research in Higher Education*, 48:633–668.

Mohanty, D. Q., Dodder, R. D., and Karman, T. A. (1986). Faculty salary analyses by region, rank, and discipline from 1977–1978 to 1983–1984. *Research in Higher Education*, 24:304–317.

Myers, K. R. (2022). The potential benefits of costly applications in grant contests. *Mimeo*.

Myers, K. R., Tham, W. Y., Yin, Y., Cohodes, N., Thursby, J. G., Thursby, M. C., Schiffer, P., Walsh, J. T., Lakhani, K. R., and Wang, D. (2020). Unequal effects of the COVID-19 pandemic on scientists. *Nature Human Behaviour*, 4(9):880–883.

National Science Foundation (2021). Survey of Doctoral Recipients (SDR). Available from https://ncses.nsf.gov/surveys/doctorate-recipients/2021. Accessed in 2023.

National Science Foundation (2023). Higher Education Research and Development Survey (HERD). Available from https://www.nsf.gov/statistics/srvyherd/. Accessed in 2023.

Perkmann, M., Salandra, R., Tartari, V., McKelvey, M., and Hughes, A. (2021). Academic engagement: A review of the literature 2011-2019. *Research Policy*, 50(1):104114.

Perna, L. W. (2002). Sex differences in the supplemental earnings of college and university faculty. *Research in Higher Education*, 43:31–58.

Philipps, A. (2022). Research funding randomly allocated? A survey of scientists' views on peer review and lottery. *Science and Public Policy*, 49(3):365–377.

Popper, K. (1934). *The Logic of Scientific Discovery*. Julius Springer, Hutchinson, and Co.

Sabharwal, M. (2013). Comparing research productivity across disciplines and career stages. *Journal of Comparative Policy Analysis: Research and Practice*, 15(2):141–163.

Sauermann, H. and Cohen, W. M. (2010). What makes them tick? Employee motives and firm innovation. *Management Science*, 56(12):2134–2153.

Shortlidge, E. E. and Eddy, S. L. (2018). The trade-off between graduate student research and teaching: A myth? *PLoS ONE*, 13(6):e0199576.

Song, J., Price, D. J., Guvenen, F., Bloom, N., and Von Wachter, T. (2019). Firming up inequality. *Quarterly Journal of Economics*, 134(1):1–50.

Stantcheva, S. (2022). How to run surveys: A guide to creating your own identifying variation and revealing the invisible. *National Bureau of Economic Research Working Paper 30527*.

Stern, S. (2004). Do scientists pay to be scientists? *Management Science*, 50(6):835–853.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Brookings Institution Press, Washington, DC.

Tuckman, B. H. and Tuckman, H. P. (1976). The structure of salaries at American universities. *Journal of Higher Education*, 47(1):51–64.

US Bureau of Labor Statistics (2023). American Time Use Survey (ATUS). Available from https://www.bls.gov/tus/charts.htm. Accessed in 2023.

US Census Bureau (2023). Current Population Survey (CPS). Available from https://www.census.gov/data/datasets/time-series/demo/cps/cps-asec.html. Accessed in 2023.

US Department of Education (2021). National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS). Available from https://nces.ed.gov/ipeds/. Accessed in 2023.

Van Praag, C. M. and Cramer, J. S. (2001). The roots of entrepreneurship and labour demand: Individual ability and low risk aversion. *Economica*, 68(269):45–62.

Wanner, R. A., Lewis, L. S., and Gregorio, D. I. (1981). Research productivity in academia: A comparative study of the sciences, social sciences and humanities. *Sociology of Education*, pages 238–253.

Weinberg, B. A. and Galenson, D. W. (2019). Creative careers: The life cycles of nobel laureates in economics. *De Economist*, 167(3):221–239.

World Bank (2023). World Development Indicators: Gini index. Available from https://data.worldbank.org/indicator/SI.POV.GINI. Accessed in 2023.

Tables and Figures

Table 1: Institution-level population summary statistics

	type	count	mean	sd
prof. share, by rank				
assistant	$[0, 1]$	158	0.25	0.10
associate	$[0, 1]$	158	0.23	0.06
full	$[0, 1]$	158	0.37	0.11
adjunct, clinical, other	$[0, 1]$	158	0.15	0.11
prof. share, by field				
engineering, math & related sciences	$[0, 1]$	158	0.17	0.13
humanities & related sciences	$[0, 1]$	158	0.17	0.10
medicine & health sciences	$[0, 1]$	158	0.36	0.30
social sciences	$[0, 1]$	158	0.13	0.09
natural sciences	$[0, 1]$	158	0.17	0.11
num. prof., total	$[0, \infty)$	158	843.60	555.92
num. schools	$[0, \infty)$	158	11.73	5.76
num. departments	$[0, \infty)$	158	80.28	40.55
HERD data				
total R&D funding, M-$	$[0, \infty]$	158	469.55	414.93
federal R&D funding, M-$	$[0, \infty)$	158	249.60	273.42
non-federal R&D funding, M-$	$[0, \infty)$	158	219.96	182.89
total wage bill, M-$	$[0, \infty)$	158	205.98	183.85
num. principal investigators	$[0, \infty)$	158	862.89	729.41

Note: All variables are from the NSAR population, except for the last rows, which are sourced from HERD.

Table 2: Professor-level sample summary statistics — Professional and socio-demographic

	type	count	mean	sd
rank				
assistant	$\{0,1\}$	4,145	0.25	0.43
associate	$\{0,1\}$	4,145	0.25	0.44
full	$\{0,1\}$	4,145	0.41	0.49
adjunct, clinical, other	$\{0,1\}$	4,145	0.09	0.28
aggregate field				
engineering, math & related sciences	$\{0,1\}$	4,145	0.17	0.37
humanities & related sciences	$\{0,1\}$	4,145	0.19	0.39
medicine & health sciences	$\{0,1\}$	4,145	0.28	0.45
social sciences	$\{0,1\}$	4,145	0.16	0.36
natural sciences	$\{0,1\}$	4,145	0.21	0.41
tenure status				
not on tenure track	$\{0,1\}$	4,145	0.20	0.40
contract length (non tenure track)	$[1,\infty)$	788	2.51	2.02
pre-tenure	$\{0,1\}$	4,145	0.22	0.41
years until tenure eval.	$[1,\infty)$	875	2.63	1.93
tenured	$\{0,1\}$	4,145	0.59	0.49
years since tenure	$(-\infty,0)$	2,364	14.86	11.51
work hrs., earnings, research funding				
work hours per week	$[1,\infty)$	4,145	49.86	13.01
work-hrs. share, research	$[0,1]$	4,145	0.39	0.20
work-hrs. share, fundraising	$[0,1]$	4,145	0.09	0.11
work-hrs. share, teaching	$[0,1]$	4,145	0.27	0.17
work-hrs. share, administration	$[0,1]$	4,145	0.15	0.14
work-hrs. share, clinical	$[0,1]$	4,145	0.04	0.14
work-hrs. share, other	$[0,1]$	4,145	0.06	0.09
own annual earnings	$[0,\infty)$	4,145	160,819.06	89,752.57
earnings share, base salary	$[0,1]$	4,145	0.74	0.31
earnings share, grant-sponsored	$[0,1]$	4,145	0.15	0.24
earnings share, supplemental	$[0,1]$	4,145	0.03	0.07
earnings share, other	$[0,1]$	4,145	0.04	0.12
earnings share, clinical	$[0,1]$	4,145	0.02	0.10
5-year guaranteed research funding	$[0,\infty)$	4,145	431,378.17	1,087,730.32
5-year fundraising expectations	$[0,\infty)$	4,145	549,068.76	1,056,442.72
socio-demographics				
asian	$\{0,1\}$	4,081	0.12	0.33
black	$\{0,1\}$	4,081	0.03	0.18
hispanic	$\{0,1\}$	4,081	0.06	0.23
other	$\{0,1\}$	4,081	0.04	0.19
white	$\{0,1\}$	4,081	0.79	0.41
US born	$\{0,1\}$	4,060	0.73	0.45
non-US born	$\{0,1\}$	4,060	0.25	0.43
married or in partnership	$\{0,1\}$	4,015	0.84	0.36
num. dependents	$\{0,1\}$	4,075	0.99	1.13
household annual earnings	$[0,\infty)$	4,016	253,644.17	158,480.32

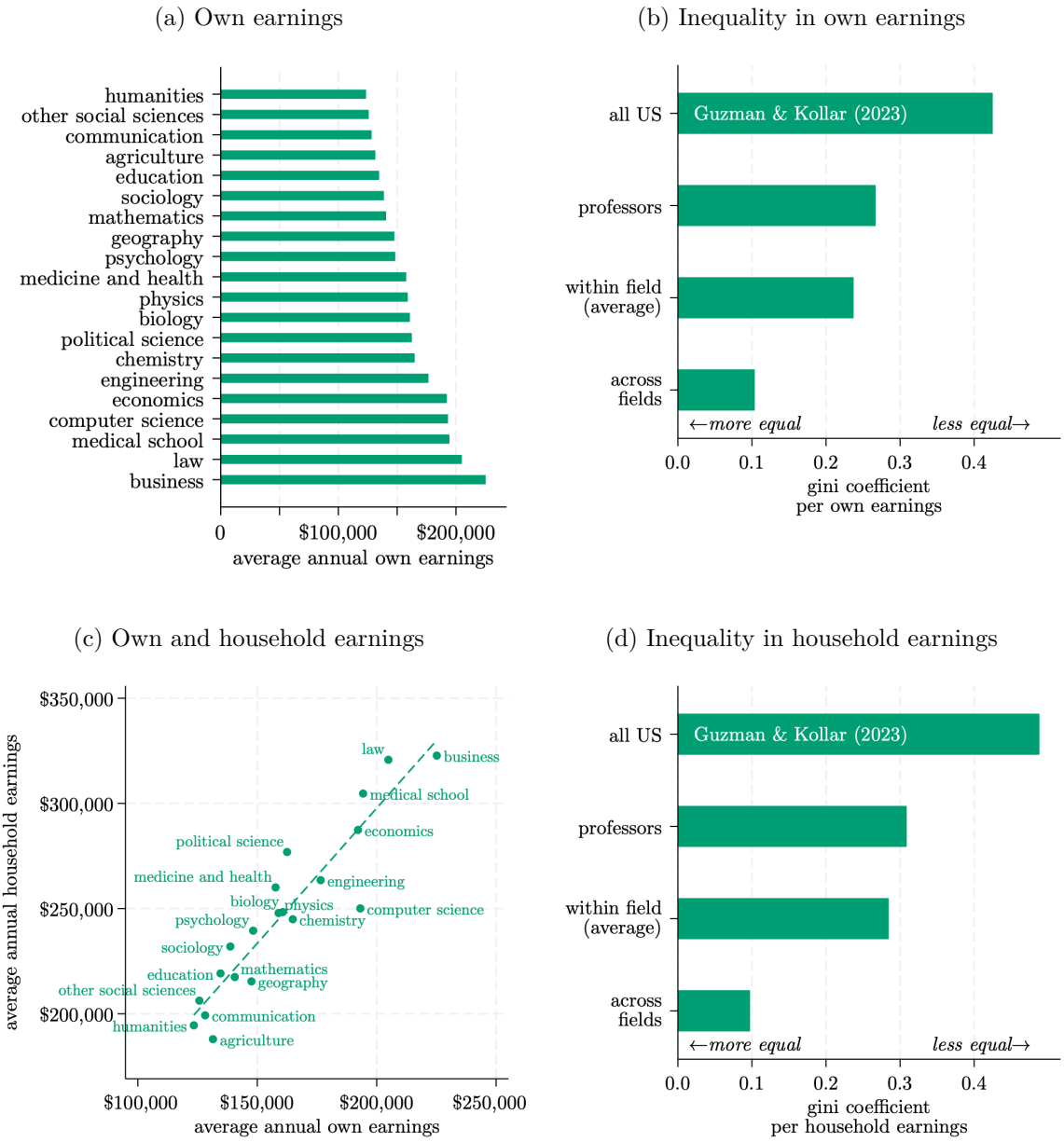
Note: All variables are from the NSAR sample. Any counts that do not equal the sample size (4,145) are due to the question either being not applicable (i.e., the tenure status questions) or were socio-demographic questions where respondents were not forced to answer.

Table 3: Professor-level sample summary statistics — Nature of research

	type	count	mean	sd
intended output				
journal articles	$\{0, 1, 2\}$	4,145	1.87	0.37
books	$\{0, 1, 2\}$	4,145	0.53	0.69
research materials, tools, etc.	$\{0, 1, 2\}$	4,145	0.70	0.68
products	$\{0, 1, 2\}$	4,145	0.47	0.64
intended audience				
academics	$\{0, 1, 2\}$	4,145	1.88	0.36
policymakers	$\{0, 1, 2\}$	4,145	0.84	0.69
businesses and organizations	$\{0, 1, 2\}$	4,145	0.54	0.62
general public	$\{0, 1, 2\}$	4,145	0.83	0.62
other				
objective is to test (vs. generate) hypotheses	$[0, 10]$	4,145	4.88	2.54
risk of research	$[0, 10]$	4,145	4.61	2.24

Note: All variables are from the NSAR sample. The variables of type "{0,1,2}" are conversions of a three-point likert scale increasing in intensity. The variables of type "$[0, 10]$" were solicited directly as real numbers on a scale from 0 to 10.

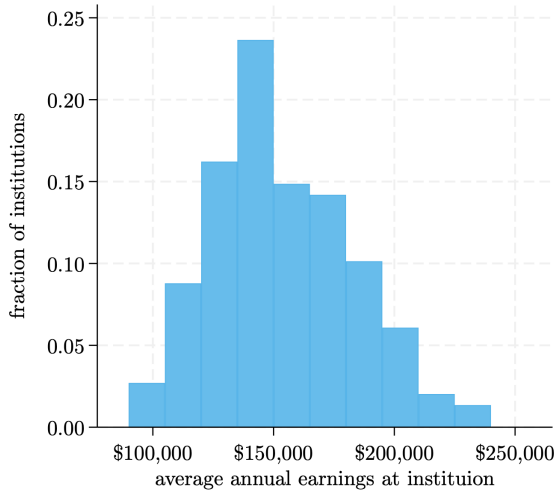
Figure 1: Individual and household earnings across fields



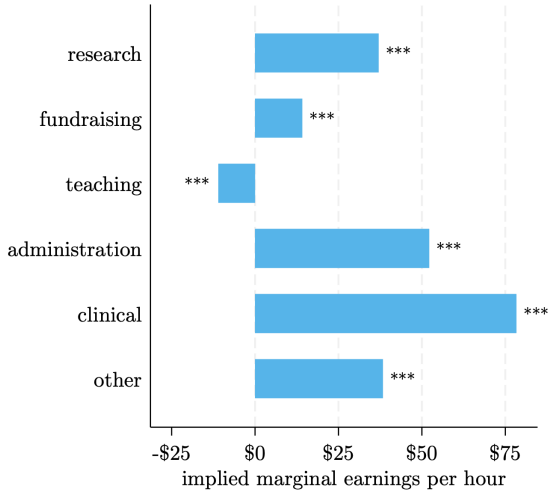
Note: Based on 4,016 observations reporting both own and household earnings. All earnings estimates hold professors' ranks fixed at the associate professor level. All averages reported are post-shrinkage to adjust for differences in sub-sample sizes across fields. Panels (b) and (d) includes estimates from the US Census Bureau (Guzman and Kollar 2023).

Figure 2: Earnings per institutions, ranks, tasks, and sources

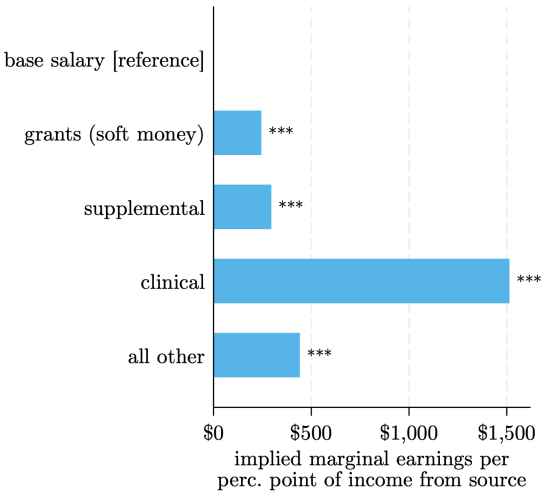
(a) Earnings across institutions



(b) Observational earnings regression, by task



(c) Observational earnings regression, by source



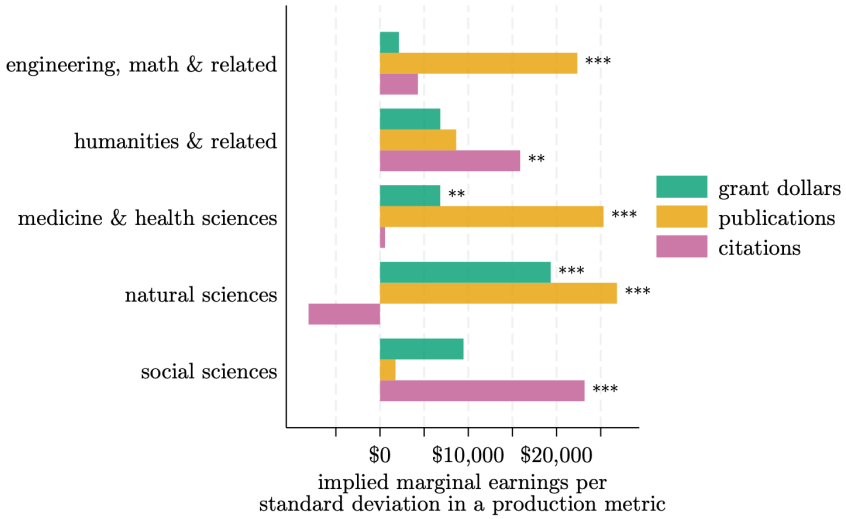
(d) Decomposing heterogeneity

	(1)	(2)	(3)	(4)	(5)
partial R^2:					
field	0.10	0.15	0.09	0.12	0.15
inst. & rank		0.46			0.50
tasks			0.14		0.11
sources				0.18	0.22
total R^2	0.10	0.38	0.21	0.23	0.55
N obs.	4,145	4,145	4,145	4,145	4,145

Note: Based on 4,145 observations. Observational earnings regressions are of the form: $Y_i = \alpha + \sum_j X_i^j \beta^j + \epsilon_i$, where Y_i is professor i's earnings and X_i^j is the vector of either (Panel b) the professor's hours spent on each tasks indexed by j, or (Panel c) the percentage points of the professor's income due to each source indexed by j; the bar graphs in Panels (b–c) report the resulting estimates of the β^j coefficients; stars indicate statistical significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.. In Panel (c), since the source percentages sum to 1, the base salary category is set to be the refernce group. Institutional averages reported are post-shrinkage to adjust for differences in sub-sample sizes across institutions.

Figure 3: Earnings per research metrics

(a) Observational earnings regression, by research metric



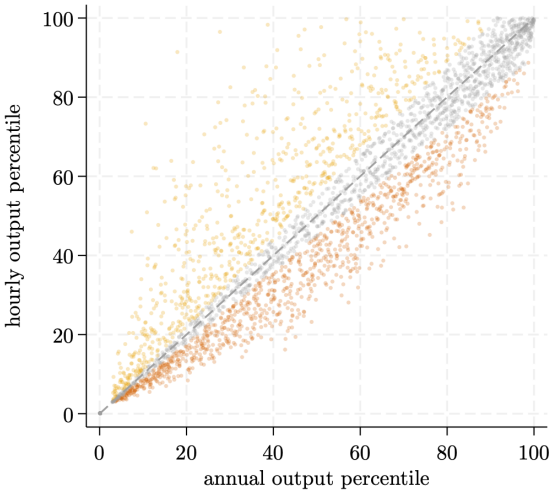
(b) Decomposing heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
partial R^2:							
field	0.09	0.10	0.11	0.10	0.11	0.14	0.15
grant dollars		0.04			0.01		0.02
publications			0.09		0.03		0.02
citations				0.07	0.01		0.01
inst. & rank						0.52	0.39
tasks						0.11	0.11
sources						0.24	0.27
total R^2	0.09	0.12	0.16	0.15	0.19	0.55	0.58
N obs.	3,197	3,197	3,197	3,197	3,197	3,197	3,197

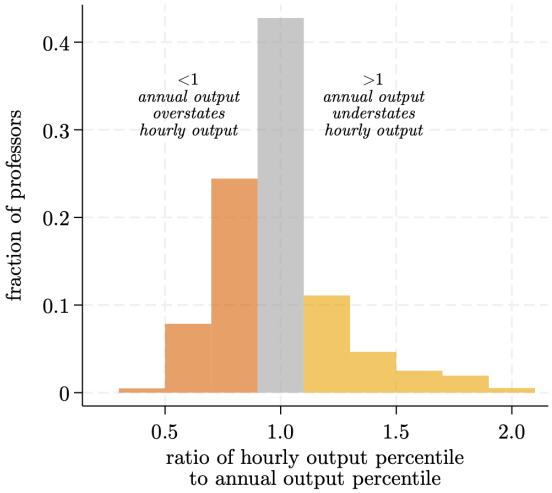
Note: Based on 3,197 observations matched to the publication/grant database. Observational earnings regression is of the form: $Y_i = \alpha_{f(i)} + \sum_j X_i^j \beta_{f(i)}^j + \epsilon_i$, where Y_i is professor i's earnings and X_i^j is the vector of their standardized inputs and outputs indexed by j and where f indexs fields; the bar graph in Panel (a) reports the resulting estimates of the β_f^j coefficients; stars indicate statistical significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 4: Research productivity and time allocations

(a) Correlation between hourly and annual output



(b) Ratio of hourly and annual output



(c) Lasso-selected predictors of output ratio

	stability selection share	univar. corr.
share of hrs., research	1.00	–0.59***
share of hrs., clinical	1.00	0.39***
non-tenure-track	1.00	0.26***
total work hrs.	1.00	–0.23***
adj. or clin. prof.	1.00	0.19***
share of hrs., fundrais.	1.00	–0.13***
output: journal articles	0.82	–0.17***
earnings, other sources	0.51	0.04**
audience: academics	0.50	–0.15***
output: materials, methods	0.44	–0.08***

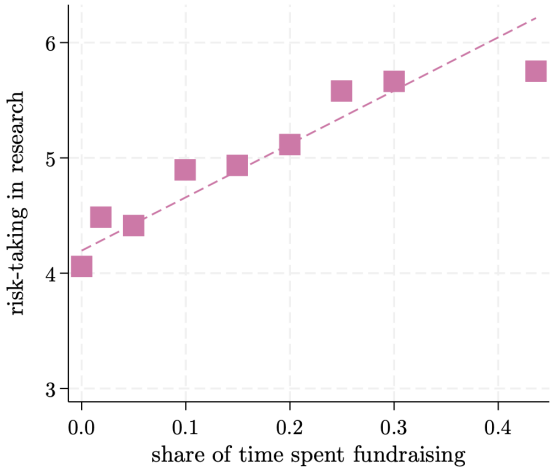
Note: Based on 3,029 observations matched to the publication database. Output is based on field-normalized publication rates per year or per research-hour. In Panel (c), stars indicate the significance of the unconditional correlation: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 5: Correlates of research risk-taking

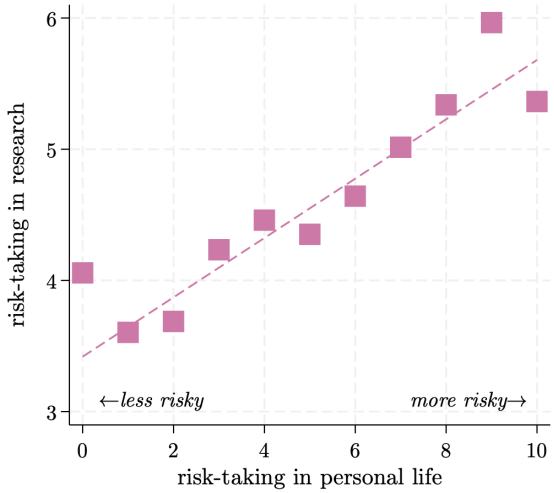
(a) Lasso-selected predictors of research risk-taking

	stability selection share	univar. corr.
share of hrs., fundrais.	1.00	0.22***
risk-taking in personal life	1.00	0.21***
test (vs. generate) hypoth.	1.00	–0.20***
share of hrs., research	1.00	0.18***
non-tenure-track	1.00	–0.14***
output: materials, methods	1.00	0.13***
female	1.00	–0.13***
household earnings	1.00	0.09***
total work hrs.	0.99	0.10***
non-recent immig. heritage	0.98	–0.10***

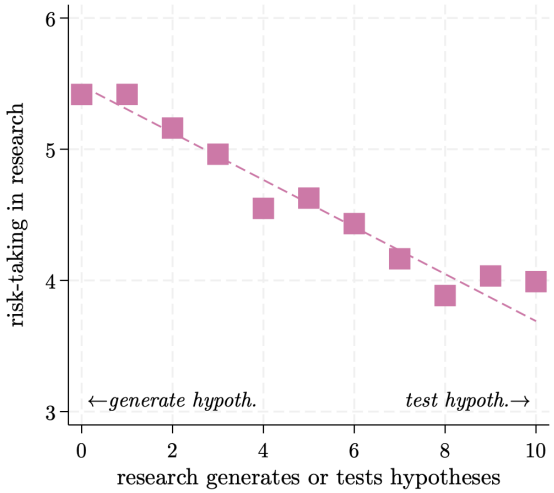
(b) Research risk and fundraising



(c) Research risk and personal risk-taking

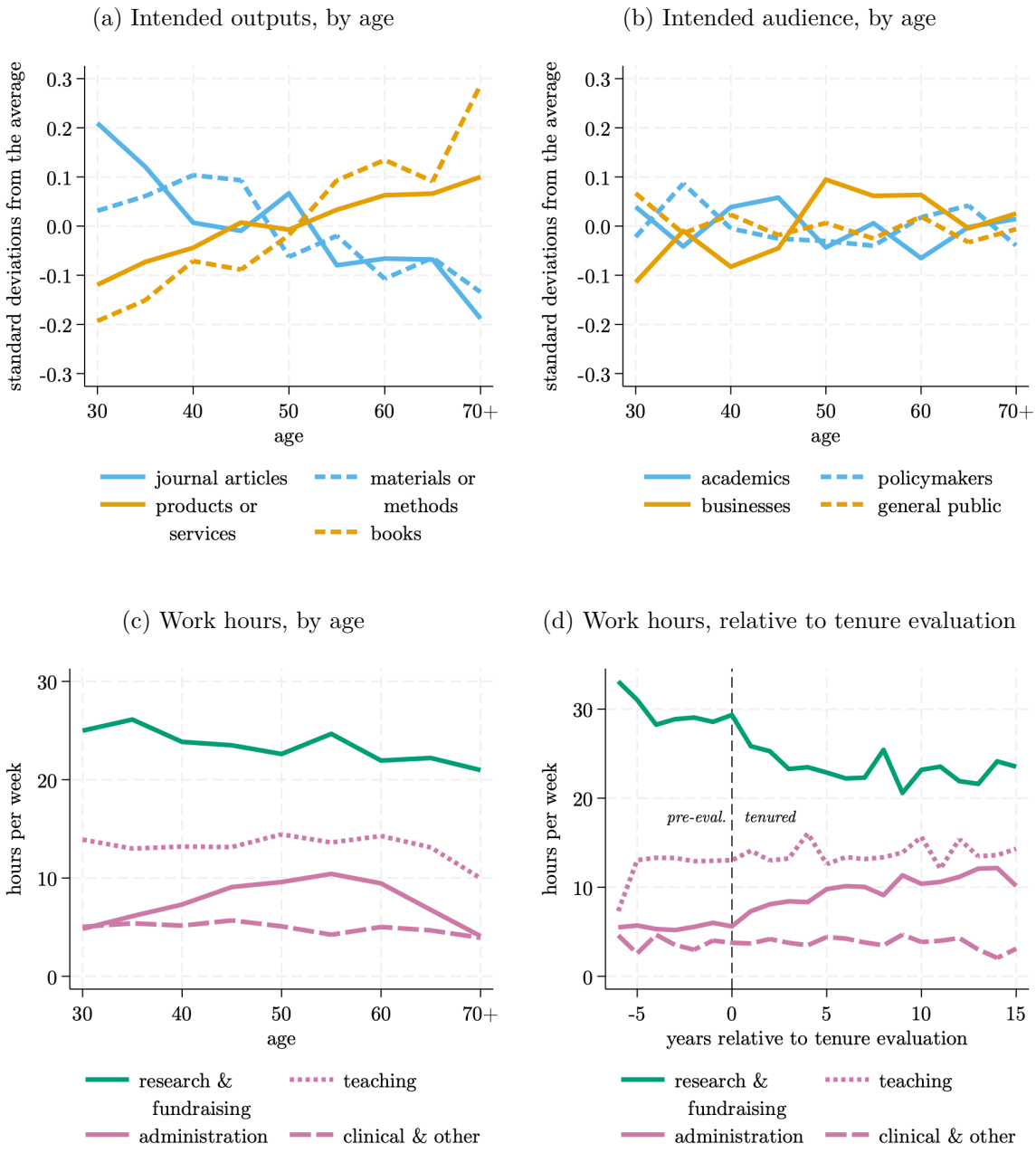


(d) Research risk and strategy



Note: Based on 4,145 observations. Panel (a) reports the results from lasso regressions predicting research risk with the full set of covariates from the sample; the stability selection share column reports the share of 100 bootstrap sub-samples that the covariate is selected in, and the univariate correlation column reports that correlation for the full sample; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Panels (b–c) shows binned scatterplots relating researchers' perceptions of the riskiness of their research (y-axis) to select covariates (x-axis).

Figure 6: Nature of research, lifecycle, and tenure



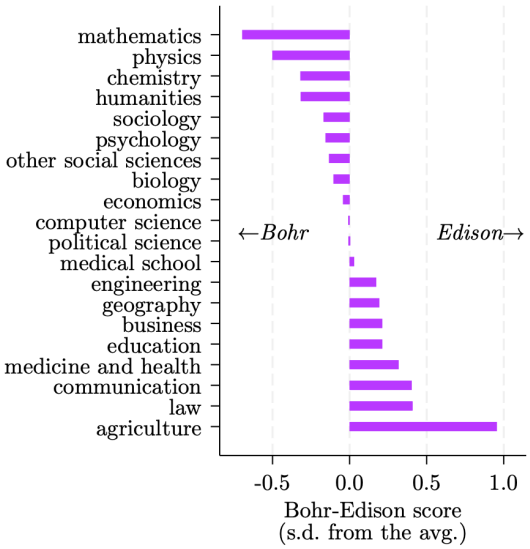
Note: Panels (a–c) are based on 4,055 observations reporting their age; Panel (d) is based on 2,161 observations from tenure-track professors at most 6 year before, or 15 years after, their tenure evaluation and includes only tenured professors post-evaluation. See the main text for details on how the intended output and audience scales are constructed.

Figure 7: Correlates of the Bohr–Edison ("basic–applied") spectrum

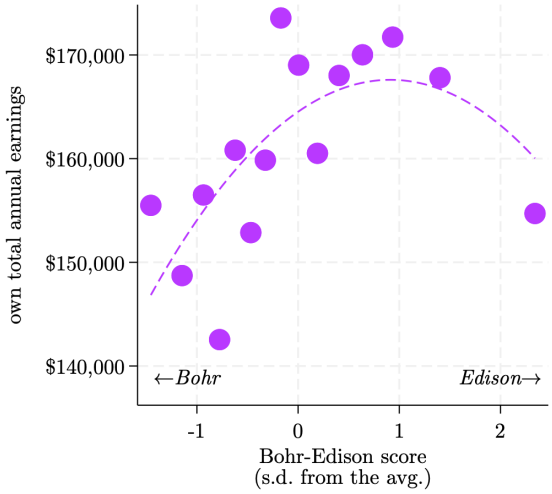
(a) Principal component analysis of nature-of-research questions

	loading
objective is to generate (0) versus test hypotheses (10)	0.18
intended output	
journal articles	–0.25
books	–0.07
research materials, tools, etc.	0.24
products	0.51
intended audience	
academics	–0.37
policymakers	0.38
businesses and organizations	0.43
general public	0.35

(b) Bohr-Edison scores, by field



(c) Bohr-Edison scores and earnings



(d) Lasso-selected predictors of Bohr-Edison score

	stability selection share	univar. corr.
share of hrs., other	1.00	0.13***
personal risk-taking	1.00	0.13***
tenured	1.00	–0.12***
earnings, other sources	1.00	0.11***
share of hrs., fundrais.	1.00	0.08***
share of hrs., research	1.00	–0.08***
has dependents	1.00	0.07***
age	1.00	–0.00
non-tenure-track	0.99	0.12***
funding expectations	0.99	0.10***

Note: Based on 4,145 observations. Panel (a) reports results from the Principal Component Analysis (PCA) used to generate the Bohr-Edison score; more positive values indicate more applied, science, and more negative values indicate more basic, science. Panels (b-c) show field-level averages and earnings relationship, respectively. Panel (d) reports the results from lasso regressions predicting the Bohr-Edison score with the full set of covariates from the sample; the stability selection share column reports the share of 100 bootstrap sub-samples that the covariate is selected in, and the univariate correlation column reports that correlation for the full sample; stars indicate the significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A Additional information on methods

A.1 Population data collection

Our process for the population data collection was as follows. We hired workers to collect publicly available contact information. The workers received detailed instructions on the layout of institutions' websites, the required fields, and data formatting guidelines. After a first round of data collection, the research team manually checked each submitted file by navigating to the relevant institutions' webpage and using random spot-checking to ensure the collected information was complete, accurate, and correctly formatted. If discrepancies were identified, workers were then instructed to add any missing information to ensure a complete record. This process repeated as many times as necessary to complete each file. Finally, the research team cleaned the files to ensure consistent formatting.

A.2 Population and sample comparisons

A key challenge for all surveys is ensuring representativeness. Here we report three tests of representativeness using data that is observable for both respondents and non-respondents. First, we test for differences in means of the few observable variables extracted from professors' online information: their aggregate field and their rank. Table A1 reports t tests showing both that our respondents includes a slightly larger share of full professors with a relatively equal smaller share of adjunct, clinical or other professors. We also see under-responding from the medical and health sciences relative to all other fields that is sizeable; the recruited share in this field is nearly 45% while the respondent share accounts for just under 30%. Our best hypothesis as to this difference is that is driven in part by the unique structure of medical schools, which account for the vast number of professors in this field. The title of "professor" can often imply notably different duties and positions at medical schools relative to the rest of academia, with professors at medical schools often carrying a significantly larger load of clinical duties. Thus, to some individuals we recruit from these schools, a survey soliciting responses about one's "research" may seem disproportionately irrelevant. This, combined with the relatively less flexible job arrangements given their clinical duties, may be the cause for this discrepancy. Altogether, this implies that caution should be had when interpreting results for this particular field.

Figure A1 shows this comparison based on data from the National Science Foundation's HERD survey (National Science Foundation 2023), which reports institutional-level data on the total amount of funding flows into all of the institutions in our population (recall, our population was constructed using the HERD, which is why this data is available for

Table A1: Test of differences between recruited and respondent samples

	mean		
	respondents	non-respondents	diff.
assistant	0.26	0.25	0.01
associate	0.23	0.22	0.01
full	0.40	0.35	0.05***
adjunct, clinical, other	0.11	0.18	–0.07***
engineering, math & related sciences	0.17	0.15	0.02***
humanities & related sciences	0.19	0.15	0.04***
medicine & health sciences	0.28	0.44	–0.15***
social sciences	0.16	0.11	0.04***
natural sciences	0.20	0.15	0.05***
N obs.	4,357	126,428	

Note: Based on 130,785 e-mail-based observations in the 50% of the population sampled for recruitment. t test of differences in means; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

the full population).

Figure A1 reports the distributions and regression tests for mean differences in six metrics of institution-level research funding that compare respondents to the full set of researchers invited to participate ("Sample e-mailed"). The distributions overlap to a large degree (see Panel a). We do estimate statistically significant differences in means (see Panel b), but the magnitudes of these differences are all in the range of approximately 4–6%.

Figure A2 reports the results of a similar exercise, instead using individual-level data on researchers' publication output and grant receipts. This data was obtained by performing a fuzzy match of our population (i.e., using names and institution data) to the Dimensions database (Digital Science 2018), which includes disambiguated researcher-level records. We focus on researchers' publications and grants during the twenty years prior (2003-2022) and see very little differences between our respondents and the full set of individuals invited to the survey. The distributions have strong overlap over the full support (see Panel a), and the mean differences are all insignificant and/or smaller than 7.5%.

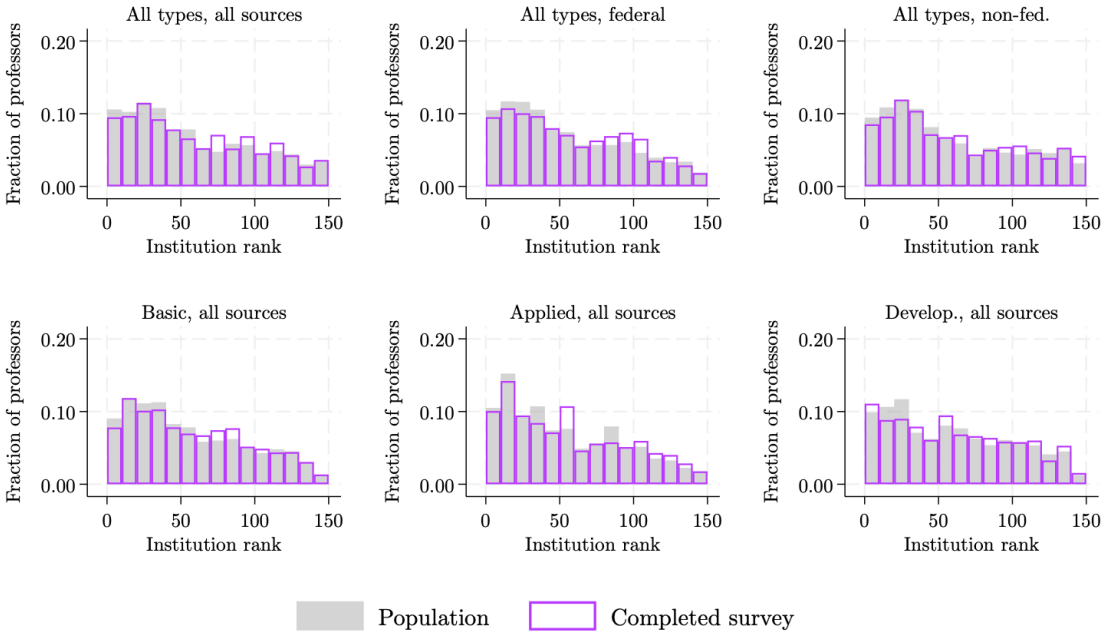
A.3 Self- and publicly-reported salary comparisons

One limitation of our survey-based methodology is that respondents do not report truthfully due to inattention or some bias such social desirability or an experimenter demand effect. Testing for such issues is notoriously challenging. But one of the survey questions, which solicits respondents' annual salary, provides one test of the degree to which respondents are reporting truthfully since a sub-sample of respondents' salaries are publicly observable.

A number of, primarily publicly owned, institutions in our population publicly post the salaries of their workforce. Using this data we can estimate the extent to which researchers'

Figure A1: Recruitment versus completion sample comparison per HERD metrics

(a) Distributions per HERD R&D measures



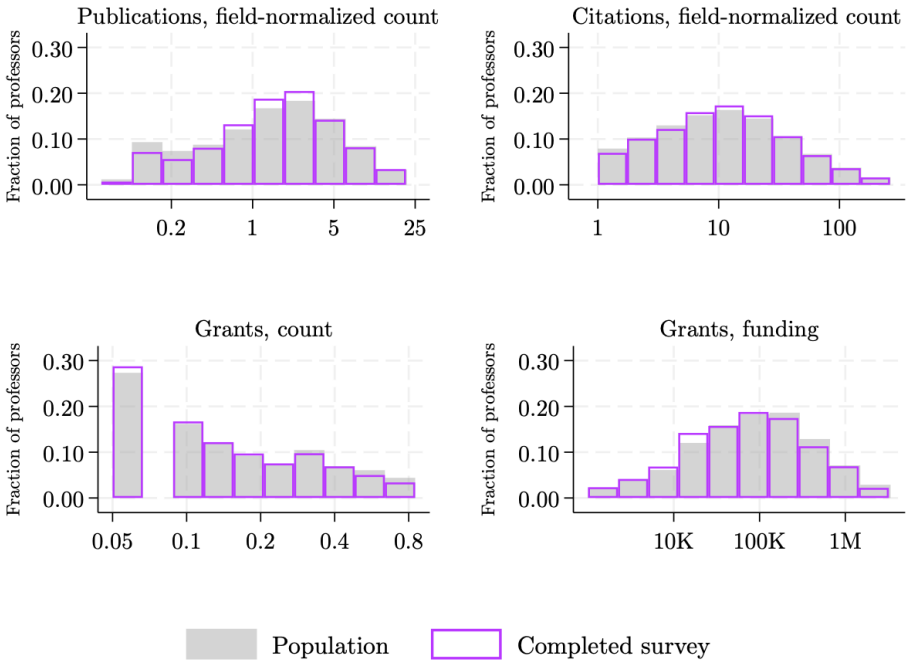
(b) Regression estimates of mean differences

	All sources, by type			All types, by source		
	All	Federal	Non–fed.	Basic	Applied	Develop.
	(1)	(2)	(3)	(4)	(5)	(6)
Completed survey	–35.01***	–21.30***	–13.71***	–24.31***	–7.170***	–3.812***
	(6.953)	(4.511)	(3.062)	(4.867)	(2.492)	(1.328)
Constant	649.1***	351.7***	297.4***	413.2***	176.7***	60.46***
	(1.308)	(0.863)	(0.578)	(0.917)	(0.459)	(0.254)
% diff.	–5.4%	–6.1%	–4.6%	–5.9%	–4.1%	–6.3%
N obs.	130,785	130,785	130,785	130,785	130,735	128,169

Note: Panel (a) compares the distributions of e-mailed professors and respondents per the rank of their institution along each dimension of R&D funding. Panel (b) reports estimates from a regression of each sampled professor's institutional R&D funding (in 2019 $-M) on a dummy for whether the sampled individual completed the survey; the "% diff." row reports the mean difference in the measure as a percentage of the non-respondent average (i.e., it is the ratio of the two coefficients); robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All institutional data is from the 2019 NSF HERD (National Science Foundation 2023).

Figure A2: Recruitment versus completion sample comparison per Dimensions metrics

(a) Distributions per Dimensions publications and grants



(b) Regression estimates of mean differences

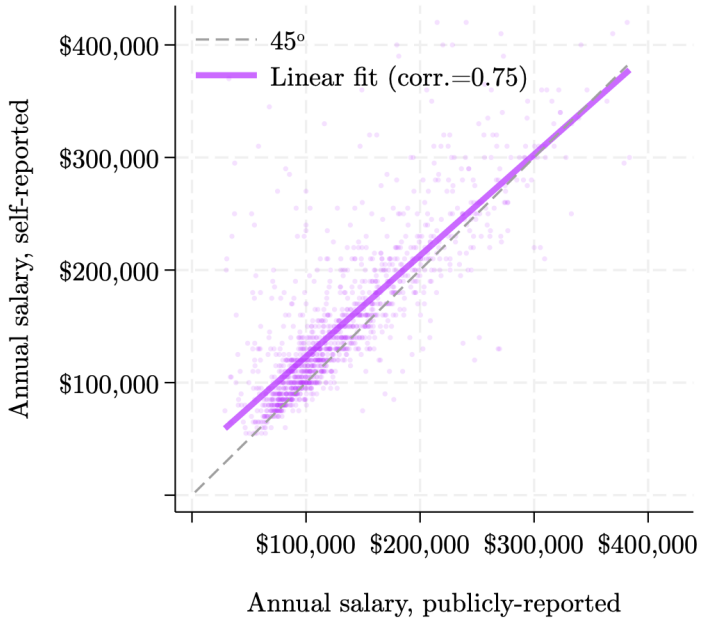
	Pub., count (1)	Pub. cites, normalized (2)	Grant, count (3)	Grant, total $ (4)
Completed survey	–0.116*	–0.573	0.00857**	–5250.7
	(0.0689)	(0.810)	(0.00344)	(8556.2)
Constant	2.902***	20.56***	0.117***	121115.1***
	(0.0158)	(0.186)	(0.000717)	(2100.2)
% diff.	–4.0%	–2.8%	7.3%	–4.3%
N obs.	86,504	86,504	86,504	86,504

Note: Panel (a) compares the distributions of e-mailed professors and respondents per each dimension of individual-level publication output and grant funding per year (2003–2022). Panel (b) reports estimates from a regression of each sampled professor's publication or grant metric on a dummy for whether the sampled individual completed the survey; the "% diff." row reports the mean difference in the measure as a percentage of the non-respondent average (i.e., it is the ratio of the two coefficients); robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All publication and grant data is from the Dimensions database (Digital Science 2018).

self- and publicly-reported salaries align.

We used a manual approach to match respondents to their public records (when possible), relying on researchers' names and affiliations to make as high-fidelity of a match as possible. Across and within the 89 institutions that had public salary data, there was a significant amount of heterogeneity in how salaries were reported. In the matching process we made note of any irregularities or confusion for each record matched. Below, we focus only on matches where both we had a high degree of confidence that both (1) the respondent was in fact the individual listed in the public record, and (2) the salary listed represented the full annual salary of that individual.

Figure A3: Self- and publicly-reported salaries



Note: Shows a scatterplot of salaries by source along with a linear fit and 45^o line for comparison.

We expect a high degree of alignment between these two measures, but not perfect alignment. Beyond the traditional concern of misreporting by the respondents, misalignment may occur for three other possible reasons: (1) human error in the matching process; (2) the publicly reported salary was for a different fiscal- or calendar-year than what the respondent was reporting or did not in fact reflect the individuals' full annual salary; or (3) we solicited salaries in the survey using a pre-populated menu of discrete values to minimize effort, which in turn creates "lumpiness" in the self-reported values.

Figure A3 reports a scatterplot of respondents self- and publicly-reported salaries, which

have a correlation of approximately 0.75. Overall, the figure indicates a high degree of alignment between the two measures, suggesting that respondents are answering this question truthfully. Given the traditional sensitivity surrounding salaries amongst this population, especially relative to the questions asked in our thought experiments, this alignment suggests that most respondents were in fact attentive and truthful.

A.4 Narrow and broad fields

Table A2: Field composition of respondents

	type	count	mean	sd
humanities & related				
communication	[0, 1]	4,145	0.03	0.17
education	[0, 1]	4,145	0.04	0.20
humanities	[0, 1]	4,145	0.11	0.31
law	[0, 1]	4,145	0.02	0.13
medicine & health				
medical school	[0, 1]	4,145	0.21	0.40
medicine and health	[0, 1]	4,145	0.07	0.26
natural	[0, 1]			
agriculture		4,145	0.02	0.15
biology	[0, 1]	4,145	0.05	0.22
chemistry	[0, 1]	4,145	0.02	0.14
engineering	[0, 1]	4,145	0.06	0.24
geography	[0, 1]	4,145	0.05	0.21
physics	[0, 1]	4,145	0.04	0.19
social & math	[0, 1]			
business	[0, 1]	4,145	0.05	0.21
computer science		4,145	0.02	0.13
economics	[0, 1]	4,145	0.02	0.15
mathematics	[0, 1]	4,145	0.03	0.18
other social sciences	[0, 1]	4,145	0.04	0.19
political science	[0, 1]	4,145	0.04	0.19
psychology	[0, 1]	4,145	0.04	0.20
sociology	[0, 1]	4,145	0.04	0.19

A.5 Match to grant and publication database

Professors in the full population data were matched to their corresponding records in the Dimensions grant and publication database (Digital Science 2018) using a iterative, fuzzy-merge process based on their name (e.g., transformations and abbreviations thereof) and institution. To minimize false positives, only matches where one record in our population

data were matched to one record in the Dimensions data were used. Overall, roughly 78% of observations are successfully matched.

Table A3 reports t tests using a select set of covariates observable for the full sample to test for differences between the full sample and those professors merged to Dimensions. Clearly, it is not a random sub-sample. Overall, matched professors seem to be more experienced, perform more research, and have higher earnings. This is consistent with the observation (based on our exploration of the Dimensions database) that Dimensions' dismabiguation and record-making processes appear to be positively correlated with research output.

Table A3: Comparison of sub-sample matched to grants and publication data

	mean		
	matched	un-matched	diff.
assistant	0.25	0.25	–0.00
associate	0.24	0.30	–0.06***
full	0.44	0.31	0.12***
adjunct, clinical, other	0.07	0.13	–0.06***
engineering, math & related sciences	0.18	0.15	0.03**
humanities & related sciences	0.16	0.29	–0.13***
medicine & health sciences	0.28	0.27	0.01
social sciences	0.17	0.10	0.07***
natural sciences	0.21	0.19	0.02
not on tenure track	0.18	0.26	–0.08***
pre-tenure	0.21	0.25	–0.04***
tenured	0.62	0.50	0.12***
work hours per week	49.73	50.29	–0.56
work-hrs. share, research	0.40	0.34	0.07***
work-hrs. share, fundraising	0.09	0.08	0.02***
work-hrs. share, teaching	0.25	0.32	–0.07***
work-hrs. share, administration	0.15	0.16	–0.01**
work-hrs. share, clinical	0.04	0.04	–0.00
work-hrs. share, other	0.06	0.07	–0.00
own annual earnings	165,218.96	145,981.01	19,237.94***
earnings share, base salary	0.72	0.79	–0.06***
earnings share, grant-sponsored	0.16	0.10	0.06***
earnings share, supplemental	0.03	0.03	–0.01***
earnings share, other	0.04	0.04	–0.01
earnings share, clinical	0.02	0.01	0.00
5-year guaranteed research funding	479,314.98	269,717.83	209,597.16***
5-year fundraising expectations	598,211.60	383,341.24	214,870.36***
N obs.	3,197	948	

B Additional summary statistics and results

Table B1: Full-survey one-dimensional PCA, by field

field	aggregate field	count	mean	sd
law	social sciences	71	0.64	2.88
political science	social sciences	163	0.57	1.93
psychology	social sciences	177	0.28	1.81
business	social sciences	198	0.27	2.42
economics	social sciences	99	0.25	2.20
chemistry	natural sciences	85	0.20	2.15
medical school	medicine & health sciences	853	0.20	2.38
physics	natural sciences	158	0.15	2.08
biology	natural sciences	211	0.10	2.36
geography	natural sciences	201	0.07	2.03
sociology	medicine & health sciences	155	0.02	2.03
medicine and health	social sciences	292	-0.03	2.12
other social sciences	humanities & related sciences	157	-0.09	1.88
education	humanities & related sciences	170	-0.29	2.06
humanities	humanities & related sciences	456	-0.30	2.14
agriculture	engineering, math & related sciences	102	-0.31	1.88
communication	engineering, math & related sciences	117	-0.32	2.18
engineering	engineering, math & related sciences	260	-0.36	2.57
computer science	engineering, math & related sciences	76	-0.59	2.76
mathematics	engineering, math & related sciences	144	-0.67	2.40

Note: Reports one-dimensional PCA score using all relevant survey questions, averaged at the field level, alongside the aggregate field assigned.

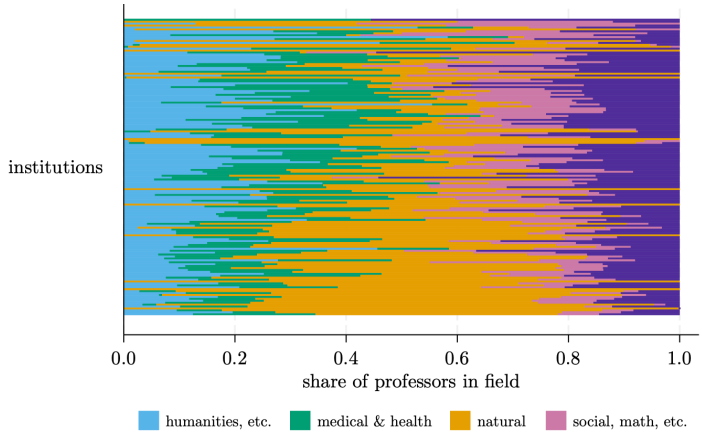
Table B2: Time use correlations

	research	fundraising	teaching	administ.	clinical	other
research	1					
fundraising	0.09***	1				
teaching	–0.22***	–0.13***	1			
administ.	–0.27***	–0.07***	0.04**	1		
clinical	–0.23***	–0.15***	–0.19***	–0.07***	1	
other	–0.11***	0.02	–0.00	–0.05***	–0.09***	1

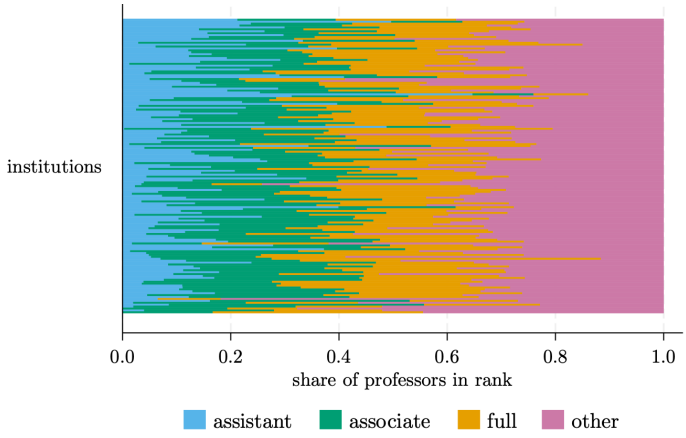
Note: Reports the pairwise unconditional correlations in hours worked per week on each task category.

Figure B1: Distributions of fields, institutions, and ranks

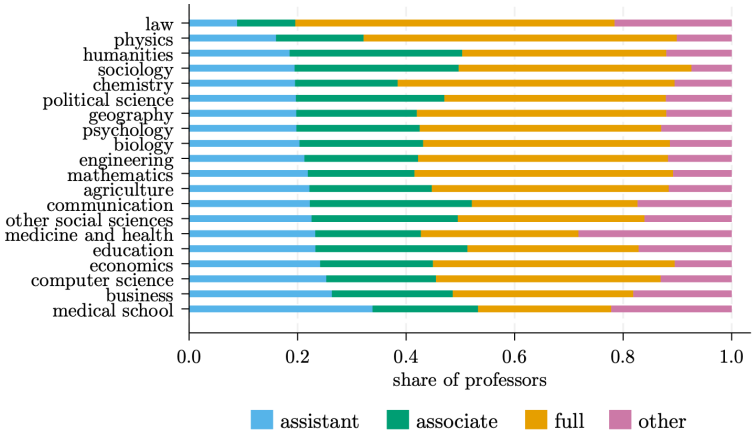
(a) Field composition of institutions



(b) Rank composition of institutions



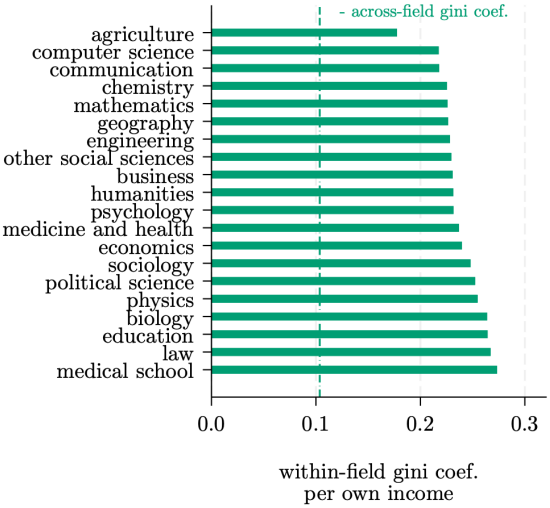
(c) Rank composition of fields



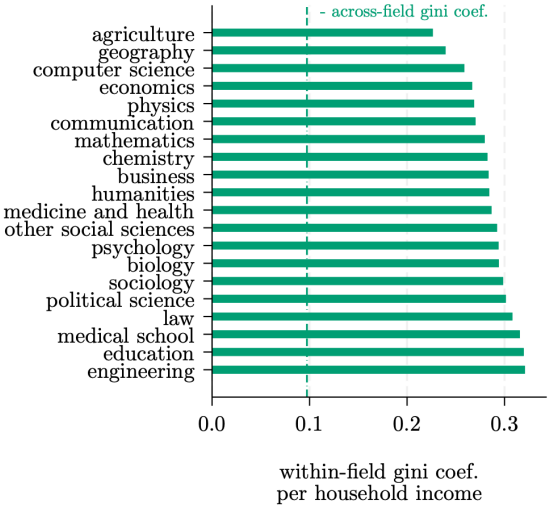
Note: Panels (a–b) plot the share of each institution per field (Panel a) and rank (Panel b); institutions are not labeled, but are sorted by size in terms of number of professors. Panel (c) plots the share of professors in each field by rank.

Figure B2: Additional earnings inequality results

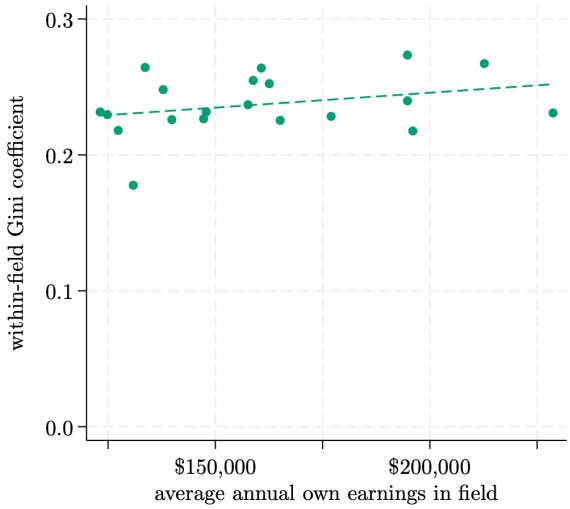
(a) Own earnings inequality, by field



(b) Household earnings inequality, by field



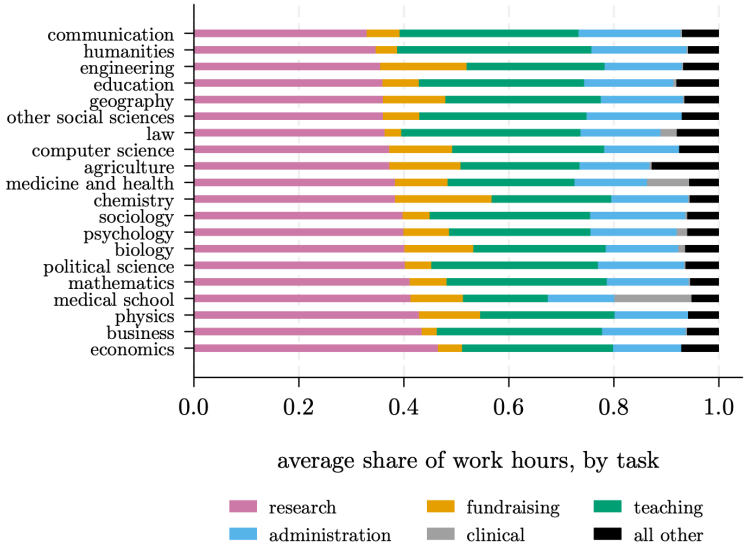
(c) Field level earnings and inequality



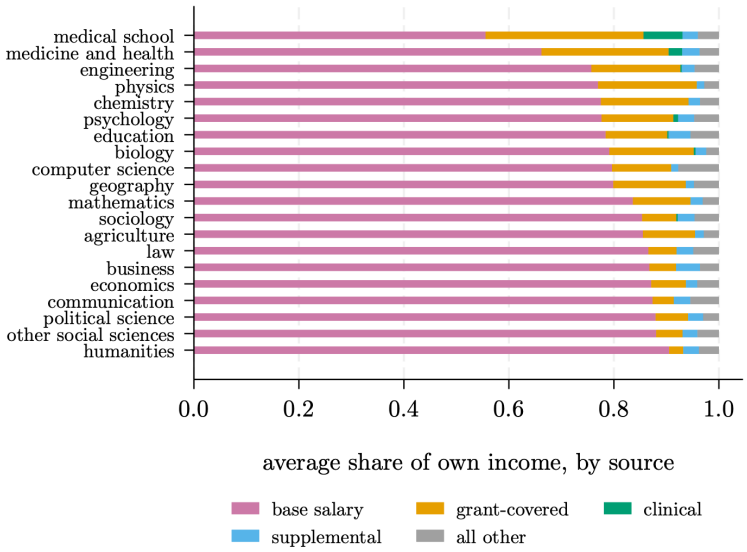
Note: Panels (a–b) plot the within-field Gini coefficients using either own (a) or household (b) earnings. Panel (c) is a scatterplot of within-field Gini coefficients per average own earnings in the same field, showing no significant relationship between the two (i.e., inequality within a field is not correlated with the average earnings in the field).

Figure B3: Tasks and earnings sources, by field

(a) Work tasks, share of total

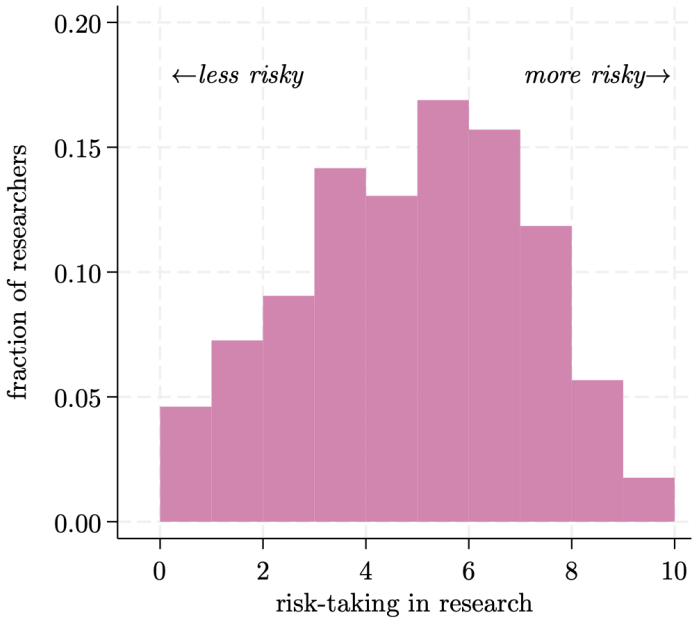


(b) Source of earnings, share of total



Note: Plots the average share of professors' total work hours spent on each task (Panel a) and the share of their total annual earnings per each source (Panel b) averaged at the field level.

Figure B4: Distribution of research risk beliefs



Note: Histogram of the average response to the two questions about researchers' own beliefs and second-order beliefs about peers' beliefs as to the riskiness of their research on a scale from 0 to 10.