The Impact of Research Funding on Publications Across Science Fields

Joshua L. Rosenbloom[a]
Donna K. Ginther[b]

[a]Iowa State University, Department of Economics
and National Bureau of Economic Research

[b]University of Kansas, Department of Economics
and National Bureau of Economic Research

January 8, 2024

Abstract

This article examines the relationship between Research & Development (R&D) funding and articles published. Using data from Academic Analytics from 2009-2019, we find that publications increased more rapidly than R&D funding, and that across universities there is a strong positive association between funding and the number of publications produced, even after controlling for faculty numbers. However, fixed-effects analysis – which controls for cross-institution variation – shows that incremental increases in funding are only weakly associated with increased publications. In the absence of true experimental variation in funding it is difficult to identify causal relationships between funding and scientific outputs. But our results do confirm that the purposive allocation of federal R&D funds results in resources being directed to faculty who convert these funds into scientific insights. We conclude with suggestions for future research.

Acknowledgements: We thank Alfred Arthur, Seongyeob Kim, Daria Milakina, and Davut Ayan for their research assistance. Fundamental Data provided by Academic Analytics, LLC. Academic Analytics, LLC management were not involved in preparation or review of the manuscript. The results presented do not necessarily reflect the opinions or positions of Academic Analytics, LLC or Academic Analytics Research Center AARC. This material is based upon work supported by the National Science Foundation under Grant No. SMA-1854849. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

In 2022 the federal government supported over $54 billion dollars of university research and development (R&D) activities. Other sources of funding – including industry, state government and the universities themselves – provided another $43 billion in support for university R&D.[1] Despite these significant expenditures, there is limited knowledge about the relationship R&D between funding and the production of scientific knowledge. Our study examines this relationship, using publications as the primary metric of scientific knowledge produced by researchers across the fields of biomedical sciences, chemistry, chemical engineering, computer science, electrical engineering, economics, and psychology. Our results show that publications increased dramatically while R&D funding grew at a much slower rate in most fields.

We focus on the period from 2009-2019. During these years the aggregate level of R&D funding fluctuated significantly, with large increases in the early years, due to the American Recovery and Reinvestment Act (ARRA), followed by flat funding between 2011 and 2015 in most fields. R&D funding began to grow again after 2015 through the end of our study period. The magnitude of these variations differed considerably between disciplines, however, with funding growing more than 50 percent in Computer Science, while funding in Chemistry actually declined slightly. In addition to variation across disciplines, there was considerable variation in the level of funding secured by faculty at different universities. While funding fluctuated, our analysis also reveals that there was a rapid acceleration in publications relative to R&D expenditures in all disciplines except Economics. Put differently, our analysis shows that the cost per article written fell substantially in most fields considered. This stands in contrast to

[1] https://ncses.nsf.gov/pubs/nsf24307 .

earlier research, which found that numbers of publications grew at roughly the same rate as funding inputs in the 1980s.

A central question for federal research funders as well as for university administrators is the causal relationship between funding received and scientific outputs. Since research funding is allocated competitively on the basis of merit, it is nearly impossible to tease out any general estimate of the causal effects of funding on the production of knowledge. In other words, no funding agency or university has run the experiment of randomly assigning additional funding to one group of researchers to see whether it boosts productivity. At the aggregate level, there are certainly year-to-year shocks to federal funding – for example the American Recovery and Reinvestment Act of 2009 resulted in an infusion of research funds that supported projects that otherwise might not have been funded. But there are good reasons to expect that the impact of small marginal variations in funding will not reveal much about the overall effect of funding for the system. The regular flow of funding from federal and non-federal funders is essential to supporting many of the activities of the scientific enterprise and its cessation or major cuts would almost surely cause a drastic reduction in knowledge production. Similarly, as illustrated by the NIH doubling that occurred between 1998 and 2003, the elasticity of the system in response to major infusion of funds is not especially high.

While we cannot say what the marginal impact of additional R&D funding, we believe that there is still a great deal to be learned from a more descriptive analysis of patterns of funding, publication, and staffing across scientific disciplines and universities over time. In particular we document that over the decade beginning in 2009, publications in most disciplines grew more quickly that funding, implying a falling average cost per publication. We also show that with a few notable exceptions, non-federal funding grew more quickly than total funding,

resulting in a shift in funding toward non-federal sources. Looking across universities we find that there is a strong positive association between the number of publications and both the number of faculty and the number of postdoctoral researchers. Finally, although there is in general a strong cross-university correlation between R&D funding and publication outputs, once one controls for persistent cross-university differences, temporal variations in funding have little or no consistent relationship with scientific outputs.

We begin with a brief overview of previous research on the relationship between funding and knowledge production. We then describe our data sources and aggregate trends before turning to an econometric analysis of the relationship between publications and R&D funding. In the final section of the paper we discuss the interpretation of our results.

The Relationship between Research Funding and Knowledge Production

A number of previous researchers have examined the relationship between R&D funding and knowledge production. Approaches to modeling the relationship between funding and publications (either raw counts or adjusted for citations) vary as do the identification strategies adopted. One approach, exemplified by Jacob and Lefgren (2011), Berg (2012), Lorsch (2015), Lauer et al. (2017) looks at the impact of receiving a grant on the productivity of an individual researcher. For example, Jacob and Lefgren (2011) use administrative data from the National Institutes of Health (NIH) on all training and standard grant proposals submitted between 1980 and 2000 to estimate the effects of being awarded a grant versus not receiving the grant on a scholar's subsequent publication. They find a positive effect of funding on subsequent publication, but the effect of funding is small, resulting in approximately one additional

publication in the subsequent 5 years.[2] In related work, researchers at the National Institutes of Health (NIH) have found pronounced diminishing marginal returns to NIH funding at the level of individual investigators (Berg (2012), Lorsch (2015), Lauer et al. (2017)).

These estimates may be biased downwards. In particular, Jacob and Lefgren's strategy relies on the fact that NIH first ranks proposals according to a priority score and then determines a "pay line", that separates funded from unfunded proposals. Reflecting this institutional detail they use a regression discontinuity approach, comparing the productivity of investigators whose applications are just above and just below the funding cut-off. As a result, their control group consists of investigators with strong, but not strong-enough, proposals. It is possible that the control group investigators may have received funding on resubmission to NIH or been able to secure funding from other sources, both of which would result in an understatement of the impact of funding on publications. Another concern with this focus on individual investigators is that it ignores the potential for spillover effects beyond the funded investigators. R&D funding to individual academic researchers includes the direct costs of the research, and sizeable payments to their university employers in the form of "overhead" or indirect costs (Facilities & Administration) in order to support the broader scientific enterprise. Facilities & Administration support is reallocated by the university to support other researchers and the research infrastructure. Further, Sattari et al (2022) found that the resources of funded investigators create "ripple effects" that increase the research output of graduate students and postdoctoral researchers that are not the principal investigators on the project.

[2] Although it remains unpublished, work by Arora and Gambardella (1998) adopts a similar methodology in examining the effects of grants from the National Science Foundation's (NSF) Economics program. Like Jacob and Lefgren, they find a relatively modest effect of funding on subsequent productivity, although this effect varies by researcher seniority.

To capture these spillover effects, several researchers have examined how R&D funding affects outputs at the level of universities or of the university system as a whole. Blume-Kohout, Kumar and Sood (2009) examined whether federal R&D life science funding was a complement or a substitute for non-federal R&D funding. They found that these funding sources were complements. Older work by Adams and Griliches (1998) examined the relationship between R&D funding and research publications in the 1980s and found evidence estimates were biased downwards when using universities as the unit of analysis as compared with aggregate estimates. They suggested that this discrepancy could be attributed to spillovers.

More recently Payne and Siow (2003) aggregate all disciplines to examine the effect of federal R&D funds on research output using a panel of 57 universities for the period 1981-1998. Using instrumental variables estimates, they found that an additional $1 million in federal R&D funding results in an increase of approximately 10 publications. Whalley and Hicks (2014) examined the relationship between university research expenditures and publications, citations and patents for a panel of 96 research universities between 1985-1996. They found that increases in research expenditures increase the number of publications but has no causal impact on citations per paper or patents.

The present study extends and broadens our previous work in the fields of chemistry and chemical engineering (Rosenbloom et al. 2015). In that earlier work, we found that the number of publications produced by researchers at a fixed sample of universities of 147 universities with the highest levels of R&D expenditures increased substantially more quickly than did the flow of funds into the system. The relatively rapid growth of publications compared to funding that we have documented for chemistry since 1990 stands in sharp contrast with earlier data for chemistry and with the trend for all sciences in the 1990s. James Adams and Zvi Griliches

(1998) found that between 1981 and 1989, funding and publication numbers in chemistry grew at close to the same rate. Javitz et al. (2010) report that for Science and Engineering as a whole, R&D expenditures per article rose by at least 10 percent between 1990 and 2001. This finding also stands in contrast to Berg (2012), Lorsch (2015), Lauer et al. (2017) who find diminishing returns to NIH funding.

Data

We will use data from Academic Analytics (https://academicanalytics.com/) to examine the returns to scale in science funding for the selected disciplines using university departments as the units of analysis. A distinctive feature of the Academic Analytics data is that this organization works closely with universities to ensure accurate lists of faculty and links them to their departmental affiliations. Academic Analytics combines automated author disambiguation with hand reconciling of data to ensure accurate publication and citation counts. Academic Analytics has collected data on individual faculty members, their publications, citations, and grants since 2009. These data are available to institutions that license the data to facilitate strategic planning. For each faculty member we have measures of journal articles, citations, grants, awards, faculty rank, faculty age, and field.

We aggregated the individual faculty counts of publications at the level of relatively homogeneous disciplines (at most universities these would comprise a department or school).[3] If a paper is coauthored with another faculty member in the same department, it is counted only once. If a paper is coauthored with another faculty member in a different department it is counted as one-half of a publication. In general, a coauthored publication is assigned to a

[3] As we noted earlier, there are good reasons to think that funding effects may spillover from individual investigators to others investigators (Sattari et al 2022).

department as (1 / number of authors in Academic Analytics). We argue that this measure is a more accurate count of publications because it is adjusted for the number of faculty coauthors. We focus on seven disciplines—biomedical sciences, chemistry, chemical engineering, computer science, electrical engineering, economics, and psychology—spanning the biological and physical sciences, engineering, and social and behavioral sciences. We combined the faculty and publication data with information on graduate students and postdocs from the NSF's Survey of Graduate Students and Postdocs and federal and non-federal research expenditures from the Higher Education Research and Development Survey. We further restrict our sample to the largest research performers by limiting our analysis sample to universities that accounted for 90% of R&D expenditures between 2009-2017.[4] Sample sizes vary across fields because some universities may not have reported having faculty in the fields we are considering.

Table 1 provides average characteristics of the full sample by field for the years 2009-2019. Biomedical fields are the largest in terms of R&D funding, faculty, graduate students, and postdocs. Biomedical fields received an average of $76 million per department with electrical engineering coming in second at an average of $20 million. Economics has the lowest amount of R&D funding at $3 million. Average publications are also the highest in biomedical fields, followed by computer science and then electrical engineering. As with research funding, economics has the lowest number of publications.

[4] At the time the sample was constructed we had data only through 2017. Subsequent updates have allowed us to extend our analysis through 2019, but we have not updated our selection criteria. But it is unlikely that there would be much change in the sample composition if we did.

Aggregate Trends in Knowledge Production Inputs

Before analyzing the relationship of research funding with knowledge production, we briefly review the evolution of aggregate (discipline-wide) measures of research inputs and outputs across the seven disciplines represented in our study. One important insight from this comparison is that circumstances varied considerably across fields, with some growing slowly while others expanded much more rapidly. In successive panels of Figure 1 we plot the growth of personnel – faculty, postdocs, grad students – financial inputs --R&D Funding – and research outputs –publications by discipline. To facilitate comparison each series is graphed as an index (set equal to 100 in 2009).

Faculty numbers increased most rapidly in Computer Science and Chemical Engineering, growing by almost 1/3 over the decade, followed closely by Biomedical Sciences. In contrast Chemistry and Psychology faculty numbers were nearly flat over the period, and Economics and Electrical Engineering fell between these two extremes. Postdocs increased most rapidly in Computer Science, Economics and Electrical Engineering, but it is important to note that the growth in these fields was from a very small base, so even small increases in absolute numbers look large. In contrast, postdoc numbers in Biomedical Sciences and Chemistry, two of the major employers of postdocs were stable or falling. Turning to graduate students, the increases in Computer Science dwarfed those in all other fields, increasing 2.5-fold. While this makes it hard to see the growth in other fields, Chemical and Electrical Engineering also achieved relatively rapid growth, while graduate student numbers in Economics and Chemistry grew very little.

Research funding followed quite different paths across the disciplines in our sample. Panels d-f show aggregate R&D funding trends as well as the evolution of federal and non-federal funding. With the exception of a drop between 2012 and 2013 R&D funding for Computer Science grew in every year, increasing more than 80 percent in the decade. The slowest growth was recorded in Chemistry, where funding was flat until 2016 before increasing slightly, and Biomedical Sciences, was essentially constant through 2014 before beginning to rise. Total funding for Economics was markedly more volatile, displaying a sharp drop between 2014 and 2015 and a pronounced spike upward between 2018 and 2019. As the next two panels make clear, much of this variation was due to differential trends in non-federal funding. Federal R&D funding did increase consistently for Computer Science and Electrical Engineering, but other disciplines saw inconsistent trends and relatively slow overall growth. Notably there was essentially no increase in federal R&D funding for Economics, Psychology or Chemistry. In contrast, non-Federal R&D funding increased more and more consistently in all of the fields we are considering.

The final panel of Figure 1 traces the growth of publication outputs, documenting an impressive growth in these scholarly products over time. Economics, which had the slowest growth still saw a nearly 30 percent increase in publications during 2009-2019. Meanwhile, the number of Chemical Engineering publications approximately doubled. The number of publications produced by Psychology faculty was almost as rapid, growing by close to 80 percent. Referring to panel a, the number of Psychology faculty increased only about 10 percent in this period (implying much higher rates of publication).

Across all seven disciplines, aggregate rates of publication outpaced the growth of research inputs. To look more closely at the relationship between publications and research

funding we calculate the average cost per publication at each university in our sample. That is we divide the number of publications in a discipline at each university by the discipline R&D funding reported at that university, to obtain the cost per publication. Figure 2 plots the

The data on outputs and expenditures shown in Figure 2 suggest that at the department level the cost per publication was declining over time. We divided total R&D expenditures at each institution by the total number of publications attributed to that institution to calculate the average cost per publication. Figure 2 plots median of the resulting distribution of cost per publication in each year by field. In most fields costs per publication were trending downwards, however the decline in costs per article were much larger in biomedical fields, chemical engineering, and psychology. The highest average cost per article was in biomedical fields at just over $200,000 in 2018 compared to a low of $50,000 in economics.

Knowledge Production and Research Funding at the University Level

Estimation

Universities produce new scientific knowledge by combining labor (faculty, postdoctoral researchers, graduate students), with capital (buildings, laboratories, infrastructure, and specialized equipment). Faculty salaries are paid for by universities, but support for postdoctoral researchers, graduate students, and most research equipment comes primarily R&D funding. We can think of buildings and other capital equipment used in knowledge production as institution-specific fixed effects.[5] To explore more fully the determinants of knowledge production, we estimated a knowledge production function of the form:

[5] For additional discussion of approaches to modeling knowledge production functions, see Hare and Wyatt (1988). de Groot (1991), which seeks to estimate university cost function is also relevant.

$$y_{it} = f\left(A_t, L_{it-l}, R^f_{it-l}, R^n_{it-l}, \alpha_i\right) \quad (1)$$

Where *i* indexes institutions, *t* denotes time periods, *L* is faculty labor input, R^f and R^n denote, respectively, federal and non-federal funding α_i is an institution specific effect, and A_t is included to capture technical changes (and other factors) that are common across all universities at a point in time. The subscripts *t-l* denote the fact that publications occur with a delay so articles published in year t reflect the results of work done in an early period. An alternative formulation replaces research funding measures with the primary labor inputs – graduate students and postdocs -- that this funding supports. That is:

$$y_{it} = f(A_t, L_{it-1}, G_{it-1}, P_{it-1}, \alpha_i) \quad (2)$$

Where G is the number of graduate students and P is the number of postdocs (both entered with a lag to reflect delays in the publication process.

In practice, we estimate a simple linear approximation to these functions, where quantities are measured in logs. That is equation (1) becomes:

$$lny_{it} = \alpha_i + A_t + \delta lnL_{it-1} + \phi lnR^f_{t-1} + \gamma lnR^n_{t-1} + e_{it} \quad (3)$$

And equation (2) becomes:

$$lny_{it} = \alpha_i + A_t + \delta lnL_{it-1} + \mu lnG_{t-1} + \sigma lnP_{t-1} + e_{it} \quad (4)$$

where e_{it} is a stochastic error term that is assumed to be independently distributed over time for each institution. Since the relationship is estimated in log-log form, the estimated coefficients can be interpreted as elasticities—the percentage change in publications given a one percent increase in inputs. We also estimate an alternative specification with total R&D funding, which would be appropriate if these two sources of funding act largely as substitutes for one another.

Given the panel nature of the data there are two sources of variation that affect the estimated coefficients: (1) variation across universities, and (2) variation over time in funding within universities. There are large and persistent differences across universities in faculty size, graduate student and postdoc cohorts and research funding that reflect differences in research orientation, institutional prestige, and other factors. The systematic portion of this cross-institutional variation is absorbed in the institution fixed-effects included in our model. Variations over time reflect both institution specific and larger national trends, as well as the vagaries of funding mechanisms which result in different levels of financial resources flowing from funders to different institutions. The A_t terms in equation (3) will capture common trends in funding variation over time, as well as common trends in publication practices or productivity. As a consequence, the coefficient estimates for faculty, research funding, graduate students and postdocs, reflect any systematic patterns arising from short-run (year-to-year) variation within universities in these variables.

To be clear, these coefficients should be interpreted as descriptive, not causal. In particular, because R&D funding is allocated purposefully and largely in response to the quality of faculty research funding applications. Thus, we cannot interpret a positive relationship between these variables as reflecting the effect of randomly allocating additional funding to a researcher. Rather it is a reflection of the fact that the funding system distributes resources to researchers who are able to translate these resources into new publications.

OLS Fixed Effect Results

We report here estimates from three versions of the panel-data model. Table 2 includes total research funding, Table 3 separates out federal and non-federal research funding, and Table

4 substitutes graduate students and postdocs for funding. The coefficients on faculty are positive and statistically significant for all fields, and their magnitudes are relatively consistent both across fields and between specifications. They imply that a ten percent variation in faculty numbers is associated with an approximately 5 percent variation in the number of publications produced. While the coefficients on R&D expenditures (both aggregated and disaggregated) are all positive, they are consistently quite small in magnitude, and statistically insignificant in most cases. At the sample mean values for Biomedical Sciences, for example, a 10 percent increase in aggregate R&D funding would be $7.6 million. The estimated coefficient in Table 2 of 0.027 implies that the marginal effect on publications of this added funding would result in just 1 (= 358 publications x 0.1 x 0.028) more publications in the next year. In Chemical Engineering, where a 10 percent increase in funding for the average department would be $820 thousand, this added funding would result in just 0.6 additional publications.

Turning to Table 4, we again find that short-run variations in numbers of graduate students and postdocs appear to have little systematic relationship with publication output in a department. The estimated coefficients on these two variables are small in magnitude, occasionally negative and rarely statistically significant.

In Figure 3 we have plotted the estimated year effects (A_t). Since we have controlled for the primary inputs used in producing scientific knowledge, one interpretation of these coefficients is that they reflect technological changes in this production relationship. Of course, an alternative view is that they reflect a reduction in quality of publications as the publish-or-perish imperative stimulates efforts to squeeze more publications out of each discovery. Consistent with our earlier finding that the average cost per publication was falling at the median university department, Figure 3 depicts a generally rising trend in publication numbers in all the

disciplines in our sample. The rate of growth was greatest in Psychology, where holding inputs constant, publications increased by almost 80 percent, while it was smallest in Economics, where the increase was under 20 percent.

Pooled OLS Regressions

The absence of a statistically significant relationship between research funding and publication outputs in the regressions discussed above does not imply that R&D funding plays no role in determining the level of research outputs within a discipline. Recall that the relationship between funding and publications in these regressions is identified based on year-to-year variations in levels of funding and publication around long-run institutional levels that are absorbed into the institutional fixed effects.[6] There are pronounced and relatively persistent variations in funding, human capital inputs, and publication outputs across universities. And these reveal a strong positive association between research funding, and publications.

To capture the relationship between key variables across institutions we re-estimate equations (3) and (4) dropping the institution fixed effects. Tables 5, 6 and 7 report the estimated coefficients from these pooled OLS regressions. In Table 5, the coefficient on total R&D expenditures is now an order of magnitude larger, ranging from a log of 0.093 in Economics to 0.483 in Chemistry, and all the estimated coefficients are statistically significant. It is also worth noting that the coefficients on number of faculty imply elasticities close to or slightly above 1. In other words, in most disciplines the number of publications increases approximately in proportion to the number of faculty. Table 6 repeats the estimates in Table 5

[6] We do not mean to imply that any of the estimated coefficients are causal. But it seems at least plausible that after removing most of the systematic variation in funding across institutions, the remaining short-run variations in funding would be more likely to reflect randomly assigned variations reflecting the unpredictable outcomes of particular rounds of funding competitions.

but splitting R&D expenditures into federal and non-federal components. Again, once we consider cross-university variations the relationships become stronger and more statistically significant.

Table 7 shifts the focus from R&D expenditures to labor inputs. In contrast to Table 4, we now find a relatively large and statistically significant association between postdoc numbers and publication outputs in all disciplines. The association between graduate students and publications is weaker, although it is statistically significant in four disciplines – Chemical Engineering, Chemistry, Economics and Electrical Engineering.

An interesting question given the close to unitary elasticity of publications with respect to faculty numbers is whether there are economies of scale in publication based on faculty size. As a simple test of this hypothesis, we re-estimated the relationships in Tables 5 interacting faculty size with indicator variables for quartiles of faculty numbers. Table 8 reports coefficient estimates for the case where we include total R&D expenditures. The coefficient on the log of faculty inputs measures the elasticity for the largest quartile of universities in each discipline, while the next three interaction effects reflect differences between institutions in the 2^{nd}, 3^{rd} and 4^{th} quartile relative to the largest quartile. There is some suggestion that faculty at the largest institutions are more productive than their peers at smaller institutions. The interaction effects are almost uniformly negative. However, the magnitudes of many of them are small and only a handful are statistically significant.

Discussion

As we noted at the beginning of this section, because R&D funding is allocated purposefully, with the goal of maximizing scientific outputs we cannot interpret the results of

these regressions as causal. Nonetheless, these descriptive regressions demonstrate that distribution of R&D funds aligns with productivity. In other words, funding is being translated into scientific discoveries that result in publications. Of course, there might be some other allocation of funding that would produce greater outputs, but it cannot be identified from these data.

A second implication of our analysis is that human capital inputs – faculty and postdocs – are also closely associated with scientific outputs. We do not, however, find that this association varies much across the size distribution of universities, leaving unresolved the factors that drive the inequality in faculty size and funding across universities.

Finally, we have found that even after controlling for faculty numbers and either financial or human capital inputs there is a strong upward trend in numbers of publications produced in all seven disciplines that we study. In the absence of better measures of the value of individual publications, it is not yet feasible to disentangle the effects of "technological progress" enabling the production of additional output per unit of input from publication "inflation" in which the same volume of discovery results in larger numbers of less valuable publications.

Conclusions

Although it is not possible to identify causal effects of funding on scientific productivity analysis of aggregate data on human capital, R&D inputs and publication outputs nonetheless yields a number of important insights about the research enterprise. First, and most intriguing is the evidence that, across a broad range of disciplines with different incentives and different time paths of funding and faculty growth, rates of growth of publication substantially outpaced the growth of inputs. This suggests that the scientific enterprise may indeed be becoming more

efficient. But it could also reflect the devaluation of publications as publish-or-perish imperatives encourage an ever increasing proliferation of publications without a corresponding increase in the knowledge base.

Second, we find that in each discipline studied, there is substantial positive association between research outputs and inputs across universities. This suggests that the mechanisms allocating funding across scholars and institutions are generally aligning resources with scientific outputs. We cannot, of course, say that a reallocation of funds would not further enhance productivity. Nonetheless, it is reassuring to find a positive and significant association between funding and scientific outputs.

DATA Appendix

The analysis in this paper rests on merging several different sources of data. We describe the sources and key characteristics of each briefly.

Research & Development Expenditures

These data are derived from the National Science Foundation's Survey of Research and Development Expenditures at Universities and Colleges/Higher Education Research and Development Survey. Data are available annually since 1973 for total and federally funded R&D expenditures by discipline. They are obtained from surveys responses completed by institutions of higher education, which are responsible for classifying all research expenditures by discipline. We computed non-federally funded R&D expenditures as the difference between total and federally funded R&D expenditures.

Sample institutions were selected from the university of institutions represented in this data by summing real federally funded R&D expenditures between 2009 and 2017 and then identifying the institutions that received 90% of R&D funding. Some institutions do not have departments in a given field, giving rise to different sample sizes by field.

Institutions report these data for the fiscal year corresponding most closely to the federal fiscal year. In most cases this is likely to run from July of one year to June of the following calendar year. Data are labeled with the calendar year in which the fiscal year ends. Hence data for 2009 most likely cover expenditures from July 2008 through June 2009.

In addition to the expenditures data this source also contains information on type of control (private or public) and standardized Carnegie Classifications that we use to categorize university types.

Graduate Students, Doctorates awarded and Postdoctoral Researchers

These data are derived from the National Science Foundation and National Institutes of Health Survey of Graduate Students and Postdoctorates in Science and Engineering (graduate student survey) which is conducted annually by the National Center for Science and Engineering Statistics. The survey is conducted in the fall semester of each academic year and are collected at the department level.

The level of institutional detail provided in this survey is greater than in the R&D expenditure data. In the latter survey a number of multi-campus state systems report a single aggregated number. To link the data sets we were obliged to aggregate the data in the student survey to match the level of aggregation of the R&D data.

Publications

Publication data were from Academic Analytics data matched to individual faculty members. Since many papers are coauthored, we applied fractional counts of papers to individual authors. These publications were then aggregated to the department level for analysis.

References

Adams, James D. and Zvi Griliches (1998). "Research Productivity in a System of Universities." *Annales d'Economie et de Statistique* No. 49/50: The Economics and Econometrics of Innovation, pp. 127-162.

Arora, Ashish and Alfonso Gambardella (1998). "The Impact of NSF Support for Basic Research in Economics." Photocopy, Heinz School of Public Policy, Carnegie Mellon University.

Blume-Kohout, Margaret E., Krishna B. Kumar and Neeraj Sood (2009). "Federal Life Sciences Funding and University R&D." NBER Working Paper 15146.

de Groot, Hans, Walter W. McMahon and J. Fredericks Volkwein (1991). "The Cost Structure of American Research Universities." *Review of Economics and Statistics* 73, no. 3 , pp. 424-31.

Hare, P. G. and G. J. Wyatt (1988). "Modelling the Determination of Research Output in British Universities." *Research Policy* 17, pp. 315-28.

Jacob, Brian and Lars Lefgren (2011). "The Impact of Research Grant Funding on Scientific Productivity." *Journal of Public Economics* 95: 1168-1177.

Lauer, M., Roychowdhury, D., Patel, K. Walsh, R., and Pearson, K. 2017. "Marginal Returns and Levels of Research Grant Support among Scientists Supported by the National Institutes of Health. bioRxiv (26 May 2017) https://www.biorxiv.org/content/early/2017/05/29/142554.

Lorsch, J. R. 2015. "Maximizing the Return on Taxpayers' Investments in Fundamental Biomedical Research." *Molecular Biology of the Cell* 26, 1578–82.

Rosenbloom, Joshua L., Ginther, Donna K., Juhl, Ted, and Joseph A. Heppert. 2015. "The

Effects of Research & Development Funding on Scientific Productivity: Academic Chemistry, 1990-2009." *PLoS ONE* 10(9): e0138176. doi:10.1371/journal.pone.0138176.

Payne, A. Abigail and Aloysius Siow (2003). "Does Federal Research Funding Increase University Research Output?" *Advances in Economic Analysis & Policy* 3, no. 1, Article 1.

Sattari, Reza, Bae, Jung, Berkes, Enrico, Weinberg, Bruce A. (2022). "The ripple effects of funding on researchers and output." Science Advances, 8:16: eabb7348. Online at https://www.science.org/doi/abs/10.1126/sciadv.abb7348.

Solow, Robert M. (1957). "Technical Change and the Aggregate Production Function." *The Review of Economics and Statistics*. 39(3): 312-320.

U.S., National Science and Technology Council, Subcommittee on Social Behavioral and Economic Sciences (2008). The Science of Science Policy: A Federal Research Roadmap. Online at http://www.whitehouse.gov/files/documents/ostp/NSTC Reports/39924_PDF Proof.pdf.

Whalley, Alexander and Justin Hicks. (2014) "Spending Wisely? How Resources Affect Knowledge Production in Universities." *Economic Inquiry*. 52(1): 35-55.

	Biomedical	Chemical Engineering	Chemistry	Computer Science	Economics	Electrical Engineering	Psychology
Federal R&D	48,556,043	4,478,654	7,395,026	10,783,012	893,963	14,479,751	5,184,624
Non Federal R&D	27,492,524	3,768,839	4,018,506	4,089,332	2,206,584	5,693,334	2,653,030
Total R&D	76,048,568	8,247,493	11,413,532	14,872,344	3,100,546	20,173,084	7,837,653
Regular Faculty	168	18	29	56	32	35	40
Grad Students	412	81	130	302	99	269	130
Postdocs	130	11	25	6	1	9	8
Publications	358	66	99	202	30	166	78
Citations	9452	1933	3321	2842	524	2342	1590
Observations	1,614	1,057	1,429	1,415	1,390	1,288	1,389

Table 1: Average Characteristics of the Analysis Sample by Department, 2009-2019.

	Biomedical		Chemical Engineering		Chemistry		Computer Science		Economics		Electrical Engineering		Psychology	
Lagged Log Faculty	0.507	***	0.597	***	0.481	***	0.521	***	0.412	***	0.476	***	0.352	***
	(0.070)		(0.097)		(0.090)		(0.085)		(0.120)		(0.100)		(0.058)	
Lagged Log Total R&D (Million)	0.027		0.062		0.105	***	0.035		-0.039		0.038		0.085	***
	(0.024)		(0.057)		(0.033)		(0.034)		(0.027)		(0.046)		(0.030)	
Constant	2.853	***	1.893	***	2.419	***	2.708	***	1.776	***	2.913	***	2.467	***
	(0.348)		(0.281)		(0.287)		(0.317)		(0.401)		(0.360)		(0.207)	
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Institution Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Number of observations	1357		876		1201		1174		1101		1072		1160	
R-squared	0.775		0.643		0.763		0.765		0.632		0.766		0.614	

Robust Standard Errors in Parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Table 2: OLS Fixed Effects Estimates of R&D Funding on Publications by Field, 2009-2019

	Biomedical		Chemical Engineering		Chemistry		Computer Science		Economics		Electrical Engineering		Psychology	
Lagged Log Faculty	0.506	***	0.603	***	0.479	***	0.520	***	0.409	***	0.479	***	0.343	***
	(0.069)		(0.097)		(0.092)		(0.085)		(0.119)		(0.100)		(0.056)	
Lagged Log Federal R&D (Million)	0.012		0.035		0.101	***	0.029		0.027		-0.012		0.115	***
	(0.028)		(0.039)		(0.035)		(0.036)		(0.042)		(0.043)		(0.037)	
Lagged Log Non-federal R&D (Million)	0.014		0.014		0.014		0.018		-0.065	*	0.029		-0.023	
	(0.024)		(0.048)		(0.027)		(0.030)		(0.033)		(0.035)		(0.026)	
Constant	2.880	***	1.924	***	2.450	***	2.712	***	1.783	***	2.970	***	2.495	***
	(0.346)		(0.278)		(0.287)		(0.318)		(0.401)		(0.360)		(0.206)	
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Institution Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Number of observations	1357		876		1201		1174		1101		1072		1160	
R-squared	0.774		0.648		0.710		0.761		0.629		0.755		0.587	

Robust Standard Errors in Parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Table 3: OLS Fixed Effects Estimates of R&D Funding by Source on Publications by Field, 2009-2019

	Biomedical		Chemical Engineering		Chemistry		Computer Science		Economics		Electrical Engineering		Psychology	
Lagged Log Faculty	0.590	***	0.598	***	0.498	***	0.574	***	0.438	***	0.499	***	0.347	***
	(0.085)		(0.093)		(0.077)		(0.080)		(0.108)		(0.090)		(0.053)	
Lagged Log Grad Students	-0.022		0.037		-0.047		0.014		0.001		0.067	*	0.029	**
	(0.038)		(0.037)		(0.088)		(0.016)		(0.034)		(0.036)		(0.015)	
Lagged Log Postdocs	-0.002		0.041	*	0.031	*	0.019		0.003		0.011		0.029	*
	(0.023)		(0.025)		(0.017)		(0.016)		(0.022)		(0.021)		(0.016)	
Intercept	2.680	***	1.750	***	2.686	***	2.475	***	1.660	***	2.538	***	2.431	***
	(0.450)		(0.311)		(0.439)		(0.306)		(0.409)		(0.391)		(0.210)	
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Institution Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Number of observations	1457		935		1288		1226		1174		1147		1246	
R-squared	0.766		0.666		0.626		0.788		0.636		0.782		0.567	

Robust Standard Errors in Parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Table 4: OLS Fixed Effects Estimates of Human Capital Inputs by Source on Publications by Field, 2009-2019

	Biomedical		Chemical Engineering		Chemistry		Computer Science		Economics		Electrical Engineering		Psychology	
Lagged Log Faculty	0.902	***	1.288	***	0.912	***	1.248	***	1.050	***	1.112	***	0.775	***
	(0.054)		(0.100)		(0.100)		(0.068)		(0.066)		(0.096)		(0.057)	
Lagged Log Total R&D (Million)	0.142	***	0.348	***	0.483	***	0.110	***	0.093	**	0.205	***	0.359	***
	(0.029)		(0.051)		(0.054)		(0.033)		(0.040)		(0.041)		(0.042)	
Constant	0.526	**	-0.539	**	0.193		-0.151		-0.451	**	0.399		0.511	**
	(0.219)		(0.241)		(0.269)		(0.242)		(0.204)		(0.281)		(0.201)	
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Institution Fixed Effects	No		No		No		No		No		No		No	
Number of observations	1357		876		1201		1174		1101		1072		1160	
R-squared	0.80		0.72		0.79		0.81		0.67		0.79		0.74	

Clustered Standard Errors in Parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Table 5: Pooled OLS Estimates of Effects of R&D Expenditures on Publications

	Biomedical		Chemical Engineering		Chemistry		Computer Science		Economics		Electrical Engineering		Psychology	
Lagged Log Faculty	0.897	***	1.279	***	0.913	***	1.242	***	1.059	***	1.110	***	0.808	***
	(0.054)		(0.098)		(0.096)		(0.069)		(0.066)		(0.103)		(0.058)	
Lagged Log Federal R&D (Million)	0.172	***	0.281	***	0.402	***	0.060		0.029		0.142	***	0.350	***
	(0.038)		(0.067)		(0.049)		(0.046)		(0.069)		(0.038)		(0.040)	
Lagged Log Non-federal R&D (Million)	-0.031		0.140	**	0.141	***	0.080		0.079		0.089		0.017	
	(0.039)		(0.063)		(0.046)		(0.058)		(0.055)		(0.074)		(0.042)	
Constant	0.585	**	-0.428	*	0.317		-0.097		-0.465	**	0.479		0.488	**
	(0.226)		(0.243)		(0.270)		(0.249)		(0.205)		(0.309)		(0.209)	
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Institution Fixed Effects	No		No		No		No		No		No		No	
Number of observations	1357		876		1201		1174		1101		1072		1160	
R-squared	0.80		0.72		0.79		0.81		0.67		0.79		0.73	

Clustered Standard Errors in Parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Table 6: Pooled OLS Estimates of Effects of R&D Expenditures by source on Publications

	Biomedical		Chemical Engineering		Chemistry		Computer Science		Economics		Electrical Engineering		Psychology	
Lagged Log Faculty	0.832	***	1.163	***	0.769	***	1.239	***	1.010	***	0.994	***	0.939	***
	(0.049)		(0.087)		(0.105)		(0.059)		(0.073)		(0.065)		(0.066)	
Lagged Log Grad Students	0.065		0.181	***	0.315	***	0.056		0.087	**	0.143	***	0.041	
	(0.058)		(0.050)		(0.107)		(0.044)		(0.036)		(0.051)		(0.029)	
Lagged Log Postdocs	0.139	***	0.243	***	0.242	***	0.088	***	0.104	***	0.190	***	0.241	***
	(0.035)		(0.034)		(0.056)		(0.026)		(0.034)		(0.033)		(0.023)	
Constant	0.442		-0.757	***	-0.459	**	-0.287		-0.644	***	0.196		-0.012	
	(0.281)		(0.267)		(0.224)		(0.212)		(0.200)		(0.210)		(0.238)	
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Institution Fixed Effects	No		No		No		No		No		No		No	
Number of observations	1457		935		1288		1226		1174		1147		1246	
R-squared	0.82		0.79		0.80		0.82		0.68		0.83		0.75	

Clustered Standard Errors in Parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Table 7: Pooled OLS Estimates of Effects of Human Capital Inputs on Publications, 2009-2019

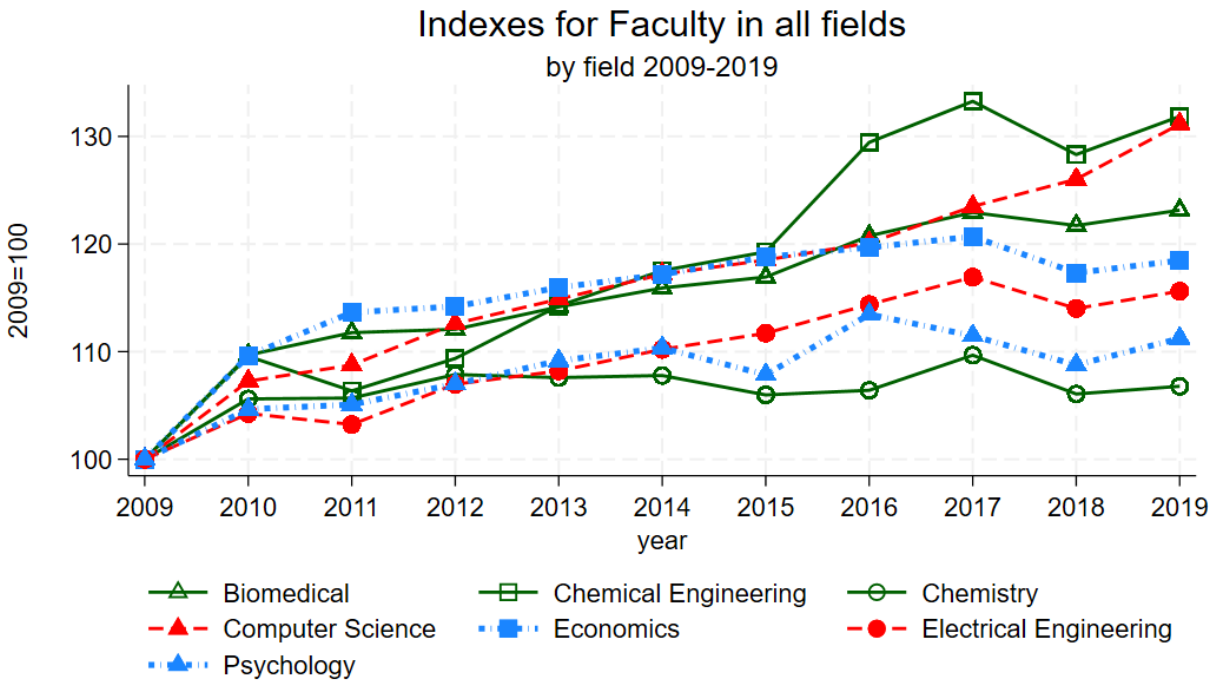
	Biomedical		Chemical Engineering		Chemistry		Computer Science		Economics		Electrical Engineering		Psychology	
Lagged Log Faculty	0.773	***	0.953	***	0.832	***	1.113	***	0.879	***	1.004	***	0.780	***
	(0.100)		(0.099)		(0.143)		(0.102)		(0.096)		(0.145)		(0.096)	
Lagged_Log_2nd_Quartile	-0.027	*	-0.008		-0.008		-0.003		-0.028		-0.003		0.017	
	(0.016)		(0.028)		(0.020)		(0.025)		(0.025)		(0.029)		(0.022)	
Lagged_Log_3rd_Quartile	-0.060	**	-0.071		-0.020		-0.010		-0.055	*	-0.013		0.078	**
	(0.024)		(0.045)		(0.026)		(0.039)		(0.031)		(0.045)		(0.032)	
Lagged_Log_4th_Quartile	-0.073	*	-0.224	***	-0.048		-0.095	*	-0.117	**	-0.078		-0.020	
	(0.043)		(0.055)		(0.048)		(0.051)		(0.050)		(0.059)		(0.046)	
Constant	1.346	**	0.689	**	0.539		0.432		0.254		0.812		0.405	
	(0.588)		(0.318)		(0.480)		(0.472)		(0.364)		(0.537)		(0.404)	
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Institution Fixed Effects	No		No		No		No		No		No		No	
Number of observations	1357		876		1201		1174		1101		1072		1160	
R-squared	0.81		0.75		0.79		0.82		0.68		0.79		0.76	

Clustered Standard Errors in Parentheses (*** p<0.01, ** p<0.05, * p<0.1)

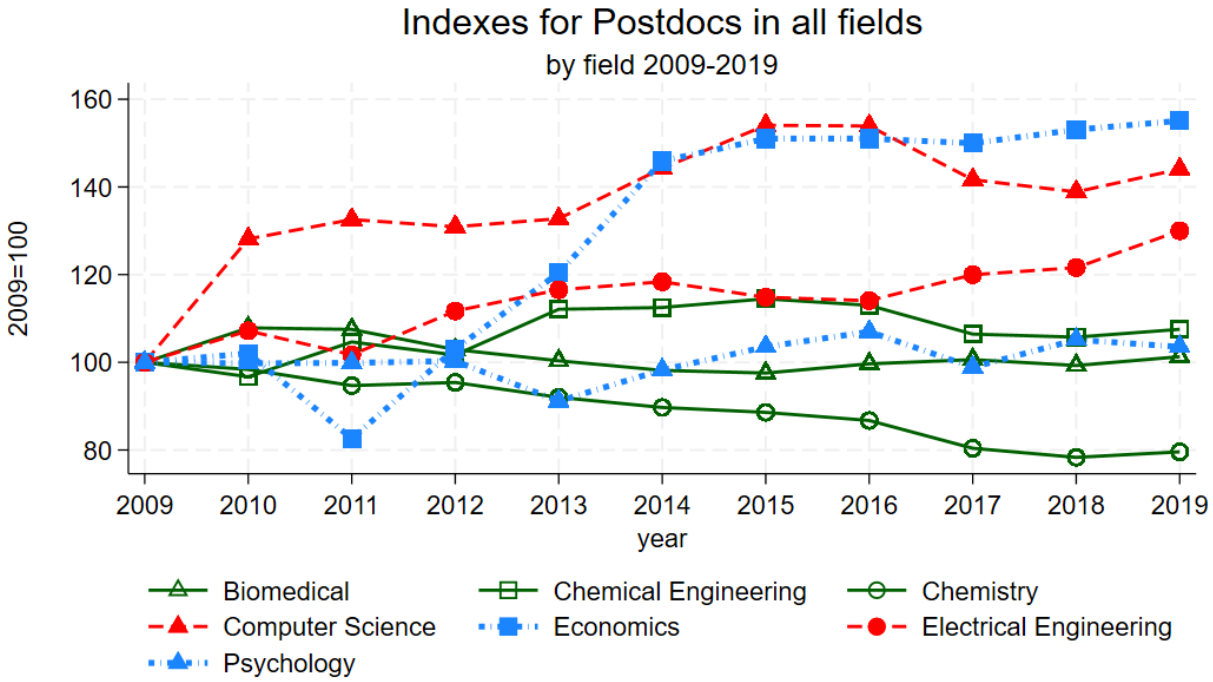
Table 8: Pooled OLS Estimates of Effects of R&D Expenditures on Publications with interactions by Faculty Size Quartile

Notes: Faculty size quartiles were established by first calculating average annual faculty size at each university over the entire sample and then ranking from largest to smallest. The first quartile contains the 25% of institutions in each discipline with the greatest number of faculty.

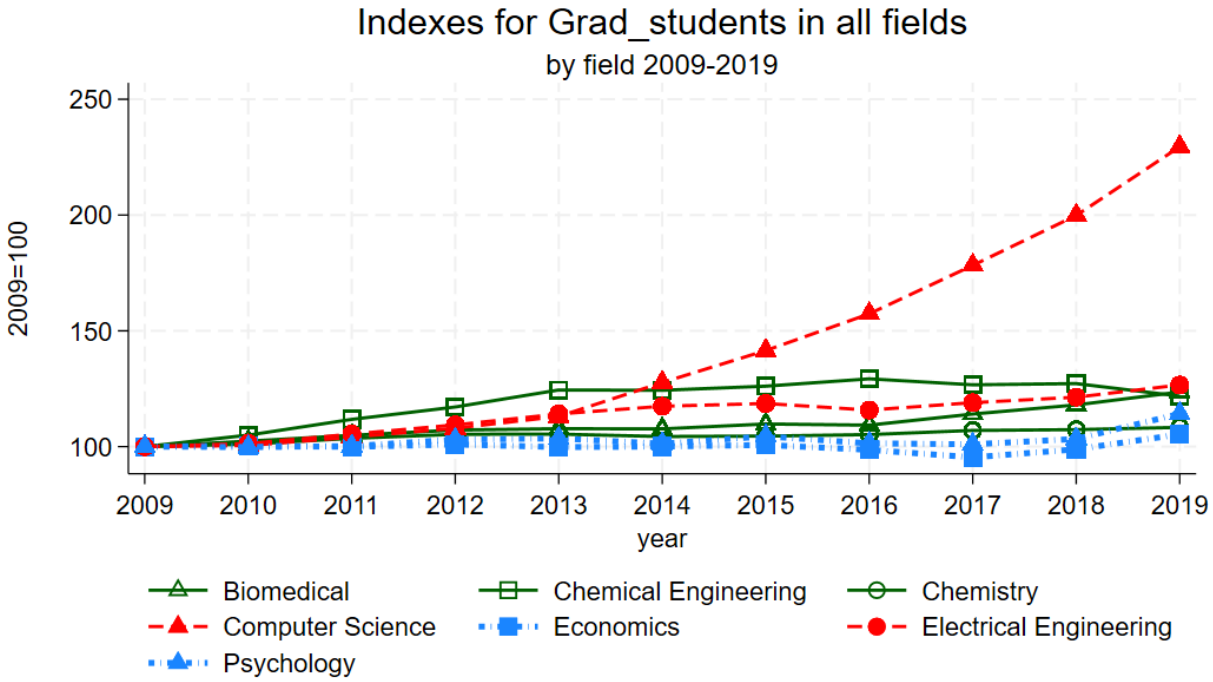
Figure 1: Indexes of Inputs and Outputs, by Field, 2009-2019



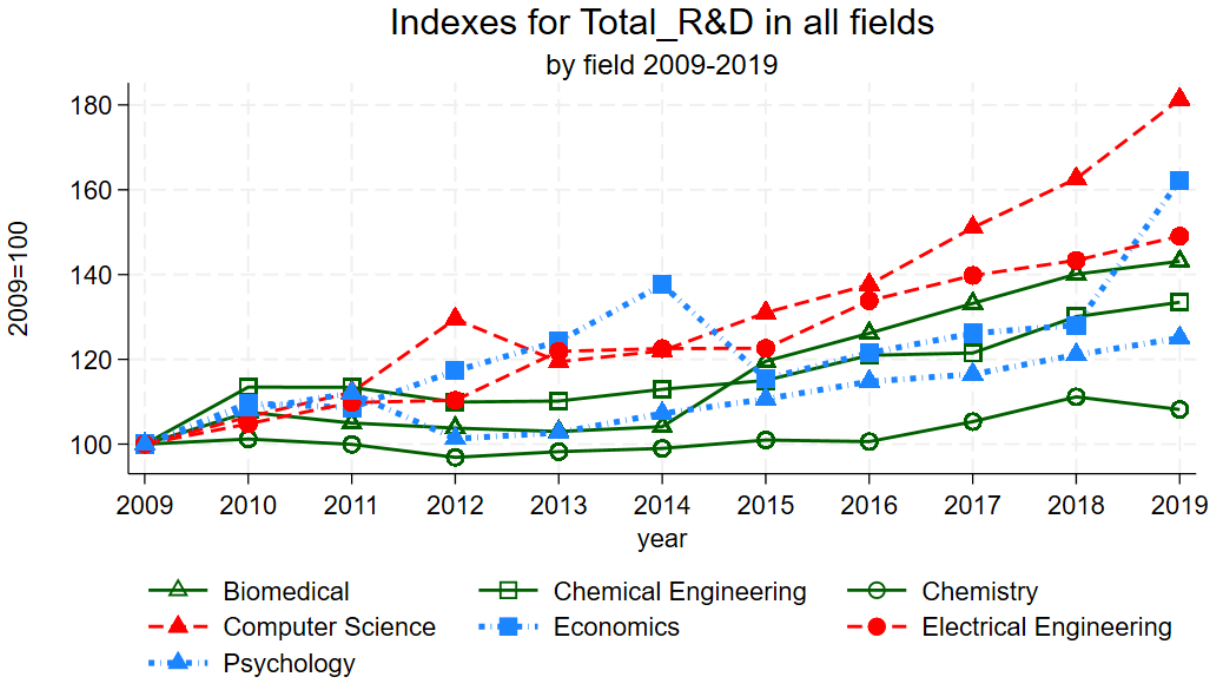
(a) Aggregate Faculty (2009=100)



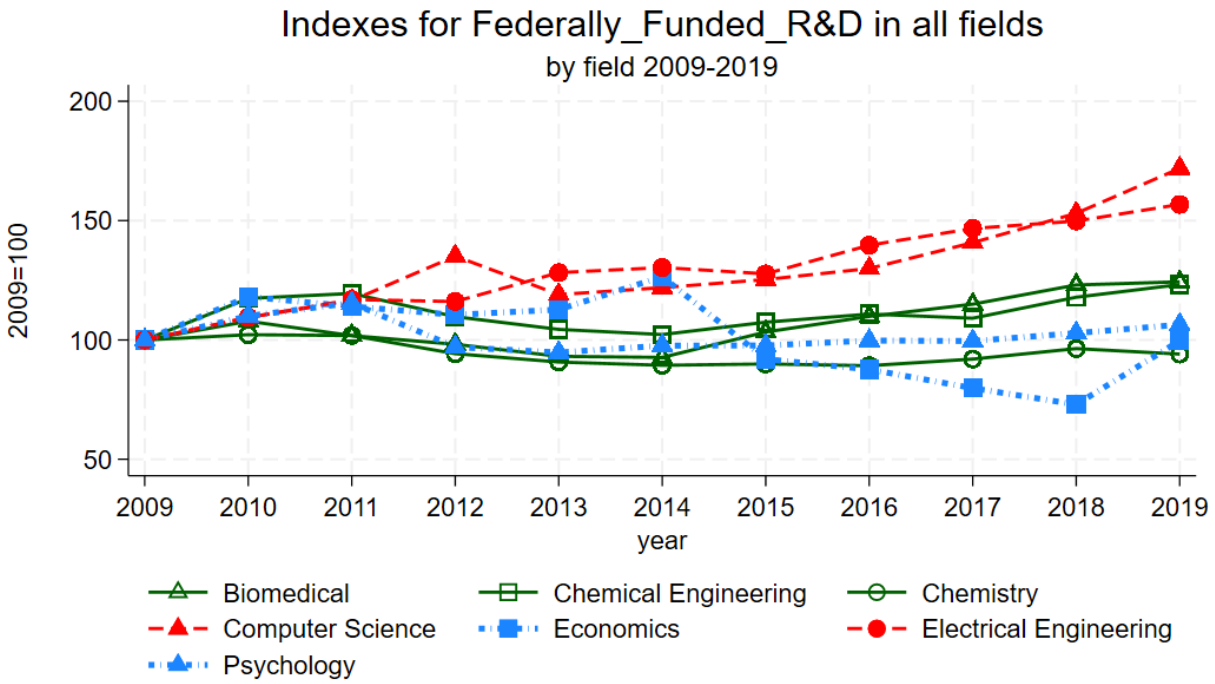
(b) Postdoctoral Researchers (2009=100)



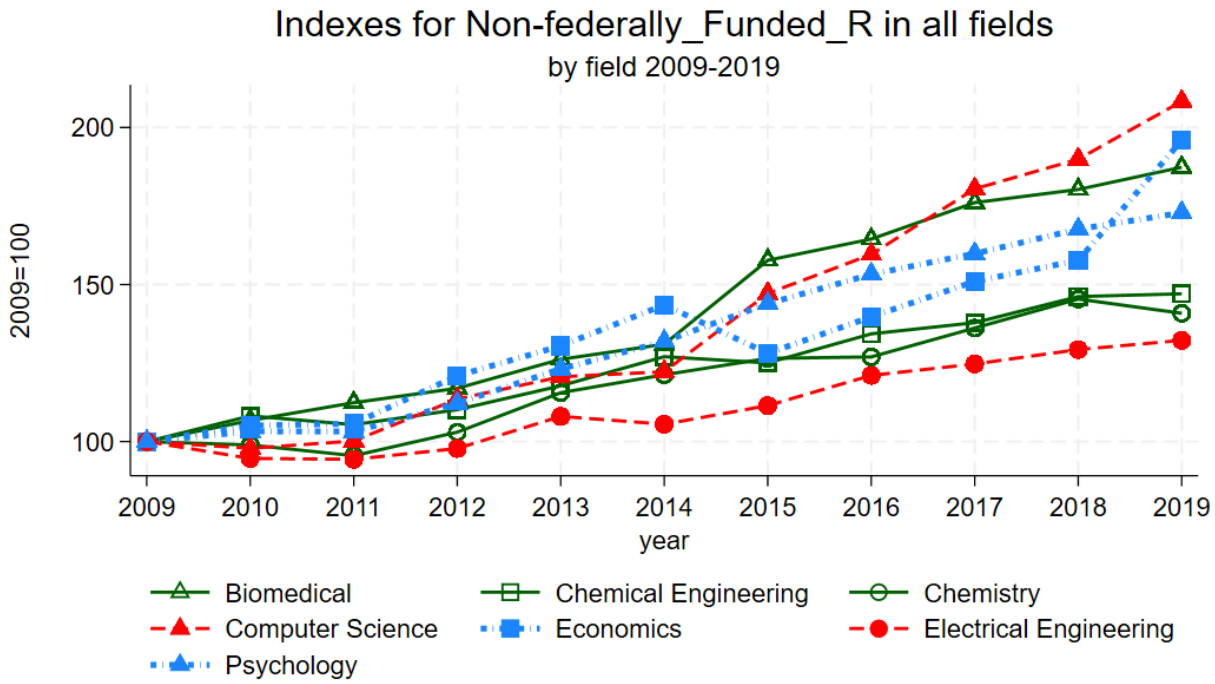
(c) Graduate Students (2009=100)



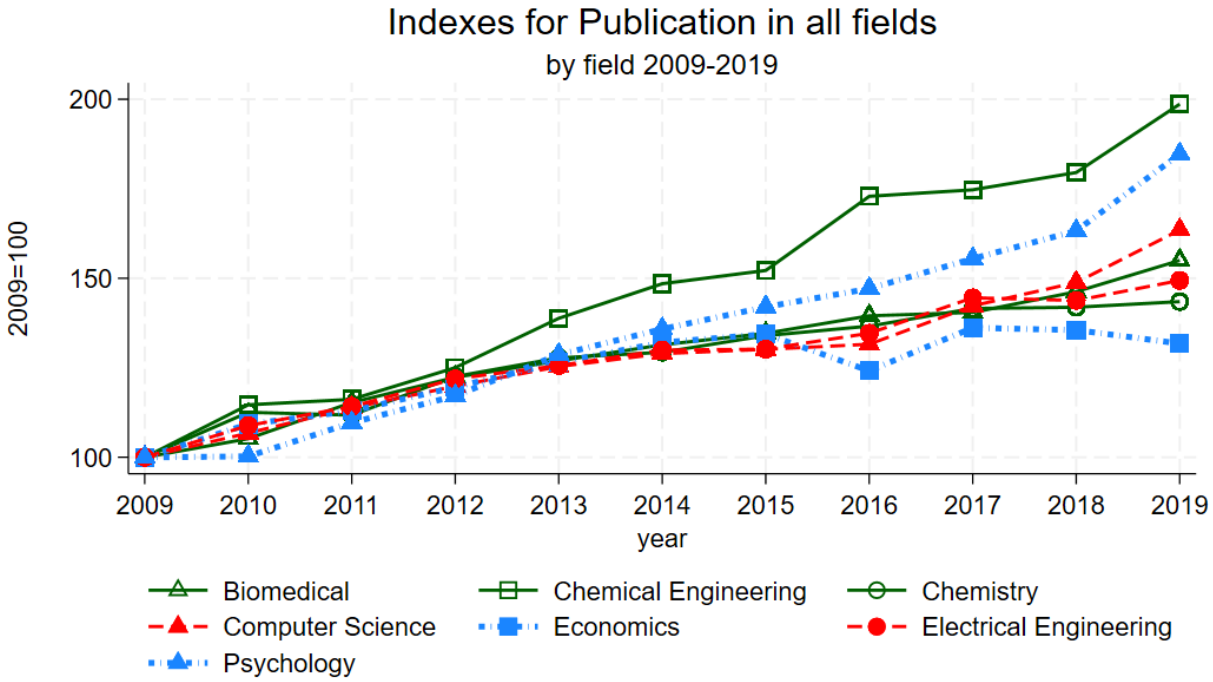
(d) Total R&D Expenditures (2009=100)



(e) Federally Funded R&D (2009=100)

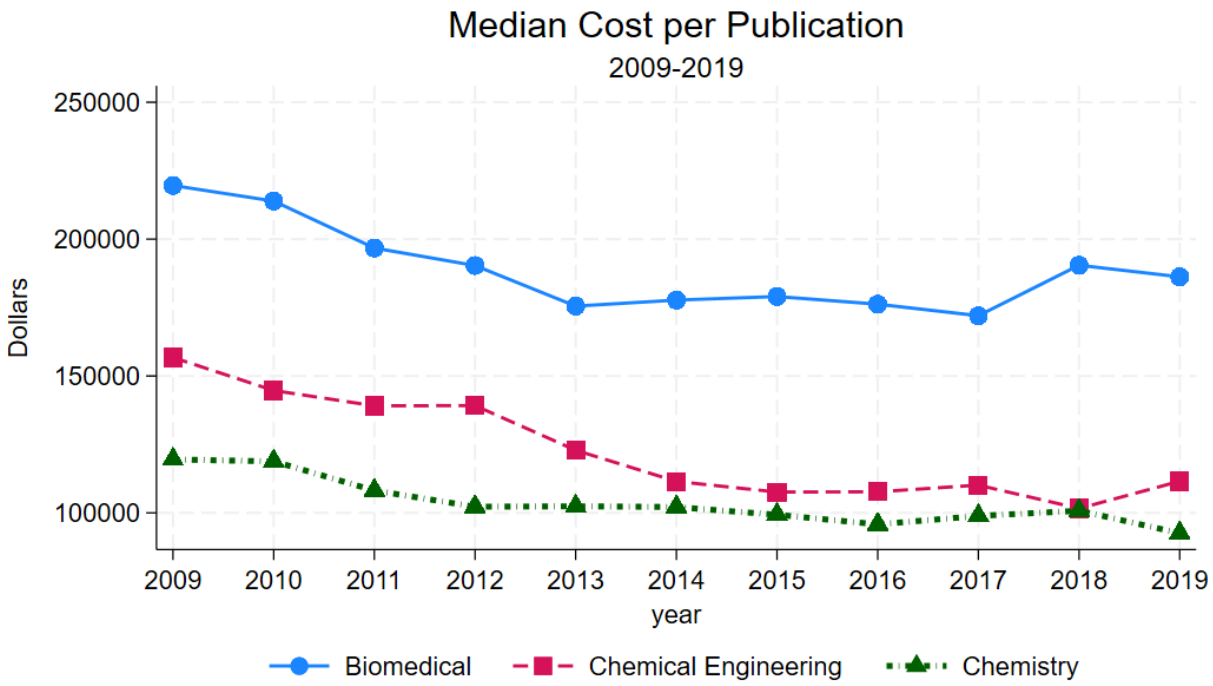


(f) Non-Federally Funded R&D

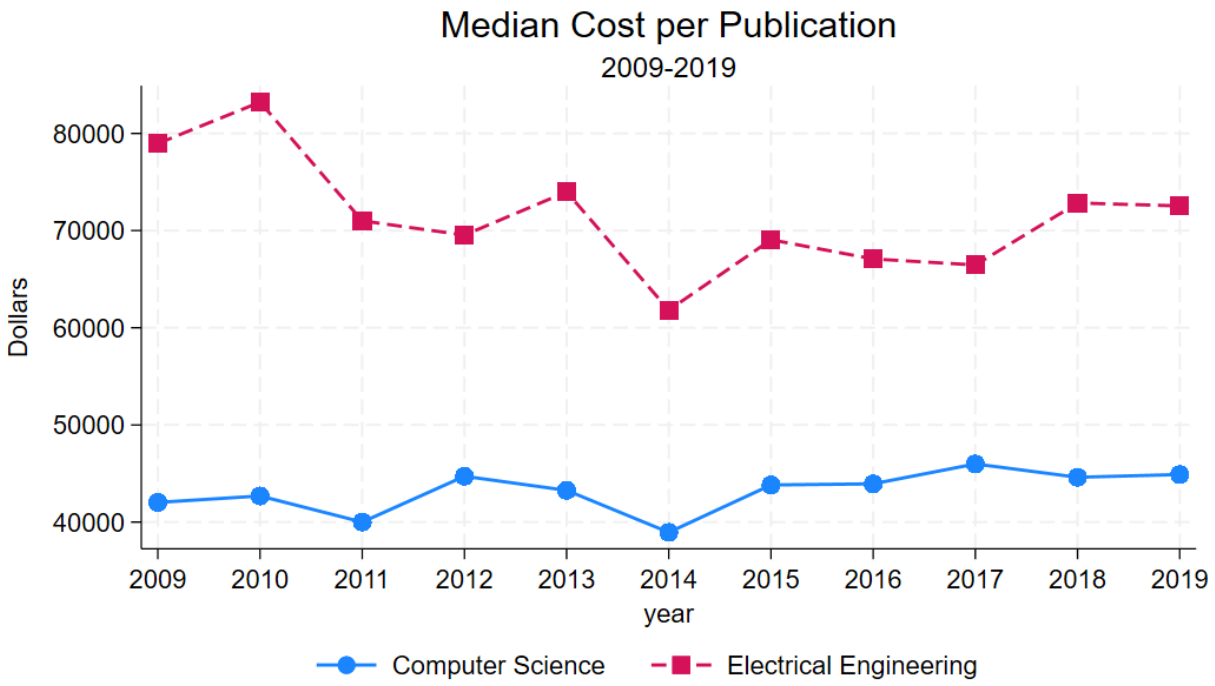


(g) Publications (2009=100)

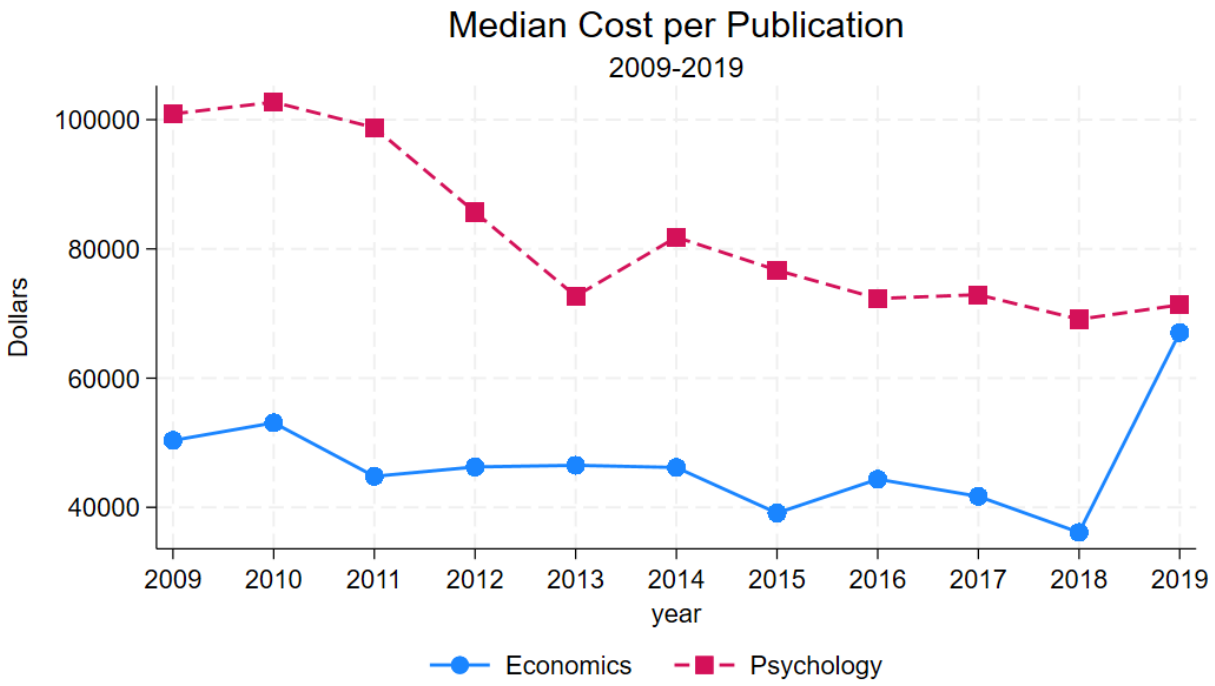
Figure 2: Median Cost per Publication



A. Biomedical Science, Chemical Engineering, Chemistry



B. Computer Science & Electrical Engineering



C. Economics & Psychology

Figure 3: Year Effects on Number of Publications by Field

