

NBER WORKING PAPER SERIES

THE EFFECTS OF RESEARCH & DEVELOPMENT FUNDING ON SCIENTIFIC PRODUCTIVITY:
ACADEMIC CHEMISTRY, 1990-2009

Joshua L. Rosenbloom
Donna K. Ginther
Ted Juhl
Joseph Heppert

Working Paper 20595
<http://www.nber.org/papers/w20595>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2014

We thank Pat Oslund, Daria Milakhina, Carlos Zambrana, Peng Chen and Nadia Kim for their research assistance. Irwin Feller offered extensive comments on an earlier draft of this article and seminar participants at George Mason University and Thomson Reuters, Research Analytics provided valuable feedback. We thank them for their contributions while absolving them of responsibility for any remaining imperfections. This material is based upon work supported by the National Science Foundation under Grant No. 1064218. 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 or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Joshua L. Rosenbloom, Donna K. Ginther, Ted Juhl, and Joseph Heppert. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Effects of Research & Development Funding On Scientific Productivity: Academic Chemistry, 1990-2009

Joshua L. Rosenbloom, Donna K. Ginther, Ted Juhl, and Joseph Heppert

NBER Working Paper No. 20595

October 2014

JEL No. H50,I23,O31,O32,O38

ABSTRACT

This article examines the relationship between Research & Development (R&D) funding and the production of knowledge by academic chemists. Using articles published, either raw counts or adjusted for quality, we find a strong, positive causal effect of funding on knowledge production. This effect is similar across subsets of universities, suggesting a relatively efficient allocation of R&D funds. Finally, we document a rapid acceleration in the rate at which chemical knowledge was produced in the late 1990s and early 2000s relative to the financial and human resources devoted to its production.

Joshua L. Rosenbloom
Department of Economics
University of Kansas
Snow Hall 436
1460 Jayhawk Blvd
Lawrence, KS 66045-2113
and NBER
jrosenbloom@ku.edu

Donna K. Ginther
Department of Economics
University of Kansas
1460 Jayhawk Boulevard
Lawrence, KS 66045
dginther@ku.edu

Ted Juhl
University of Kansas
Department of Economics
1460 Jayhawk Boulevard
Lawrence, KS 66045
juhl@ku.edu

Joseph Heppert
University of Kansas
2385 Irving Hill Rd
Lawrence, KS 66045
jheppert@ku.edu

In 2011 the federal government supported nearly \$40 billion dollars of university research and development (R&D) activities. Other sources of funding – including industry, state government and the universities themselves – provided over \$22 billion in support for university R&D. It is natural to wonder what these investments produce or, put slightly differently, what is the return on this investment?¹ In large part, public investments in R&D are motivated by the conviction that advances in scientific understanding will contribute to the nation's economic growth. Yet, we still have relatively limited knowledge about the relationship between R&D funding, the production of scientific knowledge, and the effects of this knowledge on socioeconomic outcomes. The first step in linking federal R&D funding to economic growth is to identify whether funding affects knowledge production. Thus, our study examines the effect of federal and non-federal R&D funding for chemistry research on knowledge production in these fields, measured by publications and citations. After controlling for the endogeneity of research funding, we find that that federal and non-federal chemistry research funding increases chemistry publications and citations.

We focus on academic chemistry and chemical engineering in this study because basic and applied chemistry account for a substantial share—slightly more than 4%—of all federal R&D expenditures. Moreover, the chemical sciences are large, well-established and widely represented across the spectrum of U.S. universities and they include a breadth of research topics from basic science to highly applied topics. As such, they offer an excellent opportunity to explore the

impact of federal and other sources of funding on scientific productivity. In addition, discoveries in chemistry are more likely to result in patents and other forms of commercialization. Thus identifying the effect of federal R&D chemistry funding on knowledge production is the first step in linking federal R&D expenditures to product innovations resulting in economic growth and development.

This article contributes to a greater understanding of these relationships through an examination of the relationship between R&D funding and the knowledge outputs produced by academic chemists receiving this funding over the 20 years from 1990 to 2009. Measuring knowledge outputs either by the raw number of articles published or adjusting for the quality of publications by weighting each by the number of citations it received, we document a strong, positive causal effect of funding on knowledge production. At the same time, our analysis also reveals that there was a rapid acceleration in the production of chemical science knowledge in the late 1990s and early 2000s relative to the quantities of both financial and human resources devoted to its production. In contrast to earlier research, which found that numbers of publications grew at roughly the same rate as funding inputs in the 1980s, we find that the number of articles produced grew much more quickly than either financial or human inputs into knowledge production after 1990. Put somewhat differently, we show that the cost per article written fell substantially in these years. Moreover, this trend toward declining costs occurred across a broad spectrum of institutions of higher education.

We begin, in the next section, with a discussion of previous research on the relationship between funding and knowledge production. Next we discuss our data sources and provide an overview of sample characteristics. We then examine aggregate characteristics before turning to an econometric analysis of the relationship between knowledge production and R&D funding.

The Relationship between Research Funding and Knowledge Production

We are not the first to investigate the relationship between R&D funding and knowledge production, but the research is rather limited. Previous work has taken a number of different approaches to this problem, reflecting different conceptions of the appropriate unit of measurement. One approach, exemplified by a recent study by Jacob and Lefgren (2011), looks at the impact of receiving a grant on the productivity of an individual researcher. Utilizing administrative data from the National Institutes of Health (NIH) on all training and standard grant proposals submitted between 1980 and 2000, Jacob and Lefgren estimate the effects of being awarded a grant versus not receiving the grant on a scholar's subsequent publication. To control for project quality they estimate the causal effects of a successful grant application through a regression discontinuity design that exploits the fact that funded and unfunded proposals with similar priority scores are comparable in quality. Although they do find a positive effect of funding on subsequent publication, the size of this effect appears quite modest, resulting in approximately one additional publication in the subsequent 5 years.²

These results are intriguing, but as Jacob and Lefgren acknowledge, they are subject to important limitations. Most importantly, it is not possible, given their empirical set up, to determine whether investigators placed in the control group because they did not receive funding in a particular funding round may have been successful in a later resubmission or were able to obtain resources for their research from another source. To the extent that the control-group includes investigators who subsequently obtained funding, the estimated effects of receiving a grant will be biased downward.

Another factor that will produce a downward bias in estimates at the individual level is the likely effect of leakages and spillovers in the effects of funding. Research support to individual academic researchers includes the direct costs of the research and sizeable payments to their university employers in the form of “over head” or indirect costs (Facilities & Administration) in order to support the broader scientific enterprise. These resources presumably support a broader set of researchers and research infrastructure. Further, the activity and resources of funded investigators may indirectly encourage research output by unfunded colleagues by providing a foundation of support for these investigators and supporting the institutional resources that benefit the research activities of their graduate students and postdoctoral researchers.

To better capture the effect of these spillovers, several researchers have examined how R&D funding affects outputs at the level of universities and the university system as a whole. Blume-Kohout, Kumar and Sood (2009) investigated whether federal R&D life science funding was a complement or a substitute for non-

federal R&D funding. They found that federal and non-federal funding were complements. Adams and Griliches (1998) examined the relationship between R&D funding and research publications at 109 universities in the period 1981-1993 by analyzing aggregate spending and publications in 8 broad disciplinary categories. Relying on an informal analysis of graphs showing the growth of R&D expenditures and publications or citations in each discipline, they concluded that at the aggregate, university-system level, the levels of funding and scholarly outputs grew roughly in parallel with one another in most of the disciplines they considered. They then used panel regressions for a subset of universities with complete data to estimate the elasticity of publications or citations relative to funding. In contrast to the elasticity near one suggested by the aggregate data, they found cross-section elasticities in the range of about 0.4, up to a high of 0.9. They suggested that this discrepancy might be attributable to the leakages and spillovers alluded to earlier.

More recently Payne and Siow (2003) have examined the connections between federal R&D funding and research output. Their analysis is more aggregated than Adams and Griliches, however, reporting a single aggregate estimate for all disciplines of the effect of federal R&D funds on research output using a panel of 57 universities for the period 1981-1998. Payne and Siow recognized that because R&D funding is not allocated randomly across institutions, OLS regression results cannot be interpreted as reflecting the causal effect of funding on output. To resolve this latter problem, they proposed an instrumental variables approach that relies on the effect of university alumni representation on key congressional appropriations committees. Using this instrument they found

that an additional \$1 million in federal R&D funding results in an increase of approximately 10 publications. Further, they concluded that there was no relationship between the level of funding and the number of citations per article. Payne and Siow's instrument lacks power, however, and the impact of congressional representation on the award of merit-based scientific R&D awards might be questioned a priori.

Finally, Whalley and Hicks (2014) examined the relationship between university research expenditures and knowledge production measured by publications, citations and patents for a panel of 96 research universities between 1985-1996. Like Payne and Siow, their analysis aggregated knowledge outputs and funding across different disciplines. However, they controlled for the endogeneity of funding by using changes in university endowment values as a source of exogenous variation in research expenditures. They found that increases in research expenditures increased the number of publications but had no causal impact on the average citations per paper, nor the number of patents.

The previous literature described above is dated, examining the impact of research funding on knowledge production through the mid-1990s. Given the changes in technology including the personal computer and Internet revolutions, the relationship between federal R&D funding and knowledge production has likely changed. These studies also miss important changes in federal funding priorities such as the NIH budget doubling (1998-2003) (Couzin and Miller 2007) as well as substantial increases in the NSF budget during the same period.³ Furthermore, Payne and Siow (2003) and Whalley and Hicks (2014) combine R&D funding and

aggregate publication and citation counts across all disciplines at the university level. In addition, their high level of aggregation would likely “average out” important shifts in federal funding priorities in previous decades.

Like these previous studies, we exploit variations in funding levels over time across a panel of universities. Focusing on the behavior of the university-system broadly and on the output of individual universities is more likely to capture the ways in which research funding supports the broader scientific system. Yet unlike these previous studies, we believe it is important to look at these effects at the level of individual disciplines because of differences across the disciplines in knowledge production, citation patterns, and the uses of funding. This approach is supported by the differences in the funding-output relationship documented by Adams and Griliches in their earlier study.

Data

We have collected data on the levels of research funding along with publication and citation data in chemistry and chemical engineering (for brevity we will refer to these combined fields as chemistry) for a sample of 147 universities over twenty years, from 1990 through 2009. Although there are likely to be differences in research practices and professional culture between chemistry and chemical engineering, it is not practical to cleanly distinguish between these two fields in the allocation of R&D funding or publications.⁴

Our sample was selected to include those universities that accounted for the bulk of sponsored Research & Development expenditures in Chemistry. To identify

our sample institutions, we aggregated federally funded chemistry R&D expenditures (in constant dollars) for the 20 years from 1990 through 2009 and ranked them on the basis of this total. Initially we selected the top 150 institutions, but later concluded it was necessary to drop three of these for which the data were incomplete or appeared inconsistent.⁵ After identifying the sample, we collected additional data on inputs in the knowledge production process, including degrees awarded, graduate student enrollment, postdoctoral researchers, and faculty from publicly available sources, and merged these with the R&D expenditure data. Data on publications and citations were provided by Thomson Reuters, Research Analytics group from the Web of Science citation data base. We worked closely with Thomson Reuters to identify and match publications to our sample institutions.⁶

A list of each of the institutions included in our sample is provided in Table 1, which also summarizes a number of key dimensions of real R&D expenditures and outputs. As this table illustrates, our sample exhibits a considerable degree of institutional variation. At the top of the list are institutions such as the MIT, Cal Tech, and the University of California Berkeley, with average annual R&D expenditures in the \$20-\$30 million range, employing more than 100 postdoctoral researchers, training more than 50 doctoral recipients per year and producing many hundreds of publications. At the bottom of our list are institutions such as Cleveland State University or North Carolina Agricultural & Technical State University, with total R&D expenditures of little more than \$1 million per year, with few or no postdoctoral researchers and doctoral recipients, and producing just tens of publications.

Although the 147 institutions we study do not comprise the full extent of academic research in chemistry, they account for the vast majority of the measurable research and training activities in the United States. As Figure 1 illustrates, their shares of total U.S. research expenditures, Ph.D.s awarded and postdoctoral researchers hovered around 90-95 percent, although their share of non-federally funded research expenditures was somewhat lower at about 87 percent. In comparison to R&D expenditures and graduate education, the number of publications is relatively less concentrated. Our sample institutions produced between 70 and 75 percent of chemistry publications in most years. These publications, however, received 80 to 85 percent of citations to U.S. publications over this period, suggesting that researchers affiliated with these institutions produced a greater proportion of the more important publications.

It should be noted that while the institutions in our sample accounted for a stable or slightly rising share of U.S. publications and citations, the United States share of total global publications in chemistry appears to have been declining somewhat over time. Based on data provided by Thomson Reuters from their Web of Science database, from 1990 through the early 2000s, U.S. publications accounted for about 30 percent of all chemistry publications, but after 2003, this figure began to drop, falling closer to 25 percent by 2009. U.S. publications do, however, receive a greater proportion of total citations, suggesting that they remain more important in global chemistry than the raw publication count would indicate. This share share was also declining, however, over the last decade or so.

Table 2 provides additional details about the characteristics of both the full sample and several important subsets of universities, reporting average annual values of key variable for the periods 1990-1999 and 2000-2009. Across all universities, average annual chemistry R&D expenditures increased from an average of about \$8 million in the 1990s to almost \$11 million in 2000s. Federal sources supported just under two-thirds of R&D spending in both periods, and grew at roughly the same rate as overall R&D spending. In contrast to this growth in expenditures, average numbers of graduate students enrolled, Ph.D.s awarded and employment of postdoctoral researchers held relatively steady across the two decades. On the other hand, the average number of publications produced and the number of citations to those publications both nearly doubled. The growth in the number of publications reflects both a rapid increase in the number of chemistry journals included in the Web of Science database, which increased from 244 in 1990 to 568 by 2009, and an approximately 80 percent increase in the number of publications in already established journals.

We categorized universities based on their Carnegie ranking and public or private status. Comparing across categories, important differences emerge in Table 2. As we might expect, all of the indicators of both inputs and outputs of the chemical sciences are much larger at those universities in the Carnegie Research I classification than at the other, non-Research I institutions in our sample. The Research I universities accounted, on average, for about 2.5 times the research expenditures and produced nearly three times as many doctorates as the non-Research I universities. They also employed more than three times as many

postdoctoral researchers and produced more than three times as many publications. The imbalance in citations to publications was even more striking: for 2000-2009, publications produced by the Research I institutions received almost 4.5 times as many citations as those produced by the non-Research I group, down from a ratio of more than 6 in the 1990s.

Although average research expenditures at public and private universities were similar in the 1990s, their composition was somewhat different, with non-federal funding making up almost 40 percent of total expenditures at public universities, compared to less than 30 percent at the private universities. These differences persisted over time, but funding received by the public universities grew more quickly than did funding at the private universities. Average numbers of graduate students enrolled were also higher at the public universities, but the numbers of doctorates awarded and postdoctoral researchers were comparable across the two groups. The average number of publications produced by public and private universities was quite similar, but private university publications received on average more citations than did those produced by the public universities. If citations provide a measure of the significance of publications, this result suggests that private universities were, on average, producing research of somewhat higher quality than their public counterparts.

Aggregate Trends in Chemistry Research and Funding

Before exploring the causal effect of funding on knowledge production, we consider some of the aggregate characteristics of research funding and scholarly

outputs of academic chemistry over the last two decades. We begin with the growth of funding and personnel. Figure 2 compares the growth of R&D expenditures (federally financed and total) with the number of doctorates awarded and numbers of full-time faculty and postdoctoral researchers employed.⁷ To facilitate comparison, each series is graphed as an index (set equal to 100 in 1990). Over most of the period, federal and non-federal funding grew at comparable rates, but since 2006, federal funding has stabilized while funding from non-federal sources has continued to increase.

In comparison to the nearly 70 percent increase in total R&D funding, faculty numbers were essentially flat over this 20 year period, while the number of postdoctoral researchers increased only modestly. The number of postdoctoral researchers did increase somewhat after 1996 and had grown by about 20 percent by the end of our study period in 2009. The number of doctorates awarded fluctuated with no clear trend until the early 2000s when it also began to rise slowly. Again, however, this increase was modest compared to the increase in R&D expenditures.

In Figure 3 we shift the focus to measures of research output, graphing the growth in numbers of publications and citations for our sample institutions. As in Figure 2, we have plotted each series as an index with 1990 set equal to 100, and we have included the indexes for federally-funded and total R&D expenditures for comparison. It is apparent that academic chemistry publications and citations increased much more rapidly than did funding in this period. The index of publication numbers increases consistently and reaches a value of 268 by 2009,

nearly a threefold increase. Meanwhile the index of citations to these publications grows even more quickly, achieving a value of 422 in 2008 (the last year for which we can calculate a three-year citation count).

The relationships in Figure 3 can be compared with those reported by Adams and Griliches (1998, p. 136). They found that between 1981 and 1993 publications and R&D expenditures in chemistry increased at very nearly the same rate; both increasing approximately 50 percent. Thus there appears to be a change in the knowledge production function at aggregate level for the chemistry fields in the more recent period.

The data on outputs and expenditures shown in Figure 3 make it clear that at the aggregate level the cost per publication was declining over time. In Figures 4 and 5, we use the university-level data to look more closely at the relationship between inputs and outputs over time. Dividing total chemistry R&D expenditures at each institution by the number of chemistry publications attributed to that institution yields an average cost per publication. Figure 4 plots the mean and median of the resulting distribution in each year. Both measures fell appreciably between 1990 and 1998 then leveled off. The median fell from just under \$60,000 per article to around \$30,000 per article. The divergence between the mean and the median reflects the impact of a few extremely high-cost universities, and the convergence of these two measures over time suggests that costs per publication were becoming somewhat less skewed over time.

In Figure 5 we compare the median cost of publications across subsets of universities defined by their research intensity and control (i.e., public or private).

In each group the time trends are similar and the differences in levels appear reasonable: costs are lowest at the more research-intensive universities and the public universities. Perhaps of greater interest, however, is the apparent convergence of costs across the different groups over time. Median costs fell much more sharply at less research-intensive and private universities producing a considerable convergence of costs by the late 1990s. Costs per publication appear to have risen again at the non-Research I universities during the early 2000s, but then began to fall after the early 2000s.

Knowledge Production and Research Funding at the University Level

As a rule, universities produce new scientific knowledge by combining labor (faculty, postdoctoral researchers, graduate students), with capital (buildings, laboratories, infrastructure, and specialized equipment), and other purchased inputs. While faculty salaries are supported primarily from institutional sources, support for postdoctoral researchers, graduate students, and most purchased inputs comes primarily from externally-sponsored research funding from either federal or non-federal sources. Structures and other long-lived capital equipment used in knowledge production may be treated as an institution-specific fixed effect.⁸

Formally, we can express this relationship in the following knowledge production function:

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

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 is an institution specific

effect, and A is included to capture technical changes (and other factors) that are common across all universities at a point in time. The time subscripts on the input variables are all denoted as $t-1$ to reflect the fact that there is typically a time lag between the input of resources and the publication of research results.⁹

Theory offers no guidance about the specific functional form that the relationship in equation (1) takes. In what follows, we estimate a simple linear approximation to this function:

$$(2) \quad y_{it} = \alpha_i + A_t + \gamma_f R_{it-1}^f + \gamma_n R_{it-1}^n + \delta L_{it-1} + \varepsilon_{it}$$

where ε is a stochastic error term that is assumed to be independently distributed over time for each institution. In our estimation we use faculty numbers and research funding from the prior year. It is certainly true that some publications are the result of efforts expended much longer than one year ago, but research expenditures and faculty numbers are highly serially correlated and experimentation with a number of different approaches to approximating the production lags all yield quite similar results.¹⁰

We measure knowledge outputs in two ways. The first is a simple count of the number of publications in year t that are attributed to the institution based on the affiliations recorded for each author.¹¹ The second counts the number of citations that articles published in year t receive during a three-year window beginning in the publication year.¹² To the extent that citations may be interpreted as a reflection of the significance of each publication, we can treat the citation count as a relative quality adjusted measure of knowledge production.¹³

OLS Fixed Effect Results

We begin by reporting OLS estimates of equation (2). Table 3 reports coefficient estimates for the key explanatory variables in equation (2) for three dependent variables: the number of publications, the number of citations received by those publications, and the average number of citations per publication. For both total publications and citations, the coefficients on both lagged federal and non-federal funding are positive, statistically and economically significant, and of approximately the same size. Since the funding variable is measured in \$1000s, they imply that an additional \$1 million in funding results in between 6 and 7 more publications in the following year, and 60 to 70 more citations to those publications. This implies a marginal cost per publication in the range of \$150,000. For comparison it is worth noting that in the previous section we found that the average cost per publication fell from around \$80,000 in 1990 to about \$40,000 by the early 2000s. Given the relatively small changes in the number of full-time faculty in the years being considered, it is not surprising that the coefficient on this variable is consistently small and statistically insignificant. Any effects of faculty size must be absorbed in the university fixed effects or in variations in the scale of funding. Overall, the model does a reasonably good job of accounting for the observed patterns of variation in publications and citations, explaining over half of the observed variation in the dependent variables. It should be pointed out that this is not simply a result of cross-sectional variation across institutions, as the R-squared values for within variation are also relatively high, indicating that temporal

variations in funding at a university account for a good deal of the temporal variation in research output.

The third specification, which examines the relationship between research funding and the average number of citations per publication, allows us to consider whether there is an effect of funding on the quality of publications independent of their quantity. As the estimates for this dependent variable indicate, none of the included explanatory variables contributes significantly to the average publication quality.

In Table 4 we disaggregate the sample of universities to explore whether the effects of funding on knowledge production vary by control (public vs. private) or research intensity (Carnegie Research I institutions vs. non-Research I institutions). The effects of federal R&D funding are statistically significant across all of the different subsets, and while the point estimates vary somewhat, the differences in these coefficients are not large enough to be significant. In contrast, there is marked variation in the effects of non-federal research funds. For both private universities and non-Research I institutions, non-federal funds appear to have a much smaller effect and one that is not significantly different from zero. For the average number of citations per publication, the effects of funding in the disaggregated regressions remain small in magnitude, however, there is a small, positive, and statistically significant effect on average citations for public universities and for the non-Research I universities.

In Figure 6 we have plotted the estimated year effects (A_t) which measure technological change for each subset of universities, along with the results for the

full sample. These estimates indicate that a technological shock improved chemistry productivity starting in the 1990s. The rising numbers of publications and citations to those publications was common across all of the different categories of universities we consider. However, the number of publications produced and the number of citations to those publications grew considerably more at the Research I universities than the non-Research I universities. Trends in the number of publications were quite similar between public and private universities, but the number of citations received grew more rapidly at the private universities than at the public ones. The difference was not enough, however, to cause any meaningful variation in the time trends of average citations per publication.

Because all of the explanatory variables used in the regressions reported in Tables 3 and 4 are lagged relative to the dependent variable, they are temporally exogenous. These regressions confirm that there is a positive and relatively strong association between both the quantity of federal and non-federal R&D funding an institution receives and the subsequent production of chemical knowledge. Moreover, the relative similarity of the effect sizes across different categories of universities suggests that there is no apparent inefficiency in the distribution of funding between public and private or between more and less research-intensive institutions. These results should be reassuring to funders supporting chemical research as they indicate an effective allocation of funds.

Instrumental Variable Estimation

The OLS results are not sufficient to fully identify the causal effects of funding on knowledge production. While it is not possible for future publications to directly cause past funding, these two variables may be related through a third, and unobserved variable. In particular we know that funding is allocated purposefully. Thus, investigators with more promising research programs may be more likely to be successful both in seeking funding and in publishing their research results. If this is the case, then the positive correlations revealed by the OLS regressions could simply reflect this unobserved variation in investigator quality, which is positively selected by universities in our sample.

Put somewhat differently, the OLS results are not sufficient to answer the policy-relevant question: how would the quantity of knowledge produced change if an additional \$1 million was allocated to chemistry research? Ideally, to answer this question we would want an experimental setting in which additional research funds could be randomly assigned to some universities and not to others. By comparing the impact of funds on the treatment and control groups we could identify the impact of additional research funding on scientific output. We cannot conduct such an experiment, but we can use instrumental variables to obtain estimates of the effects of such truly random variations in funding levels

The requirement for a good instrumental variable is that it be correlated with the explanatory variable of interest, but uncorrelated with the error term. There are a number of candidate instruments available that we can employ. After exploring a number of possible candidates we settled on three: federally funded R&D

expenditures for math and physics research, non-federally funded R&D expenditures for math and physics research, and the fall student enrollment numbers. The first two of these instruments should capture important variations in the larger funding environment as well as more institution specific effects such as major fundraising campaigns or faculty recruitment efforts that increase a university's overall research competitiveness. The last instrument is included to capture effects of tuition revenue as a source of faculty recruitment that should be associated with changes in levels. A priori, there is no reason to expect that any of these instruments should be correlated with random shocks in the quality of research funding proposals submitted by chemistry faculty at an institution.

Table 5 reports the first stage regression results from the two stage instrumental variable panel regressions for each of the two endogenous explanatory variables—federally-funded chemistry R&D expenditures, and non-federally-funded chemistry R&D expenditures. Overall, the model appears to do a good job of accounting for variation in the endogenous variables, and we see a strong positive effect of enrollment in both equations, while each of the math and physics expenditure variables exerts a strong positive effect in the regression for the corresponding chemistry variable. The F statistics in Table 5 indicate that we do not have a problem with weak instruments.

In Table 6 we report the second stage coefficient estimates for the knowledge production function for the full sample and for each subset of institutions. Given the poor fit of the OLS regressions for the average number of citations per article, we have dropped this variable from the IV analysis and focus only on publications and

citations received. For the full sample, we find that the effect of federally funded R&D expenditures has increased by a factor of approximately 3 for publications and 3.5 for citations. On the other hand, the effect of non-federally funded R&D expenditures has fallen substantially in magnitude and is no longer statistically significant for either publications or citations. The Sargan-Hansen statistic indicates that we cannot reject the null hypothesis of over-identified instruments.

Results are broadly consistent across the different categories of institutions we have identified, although coefficient magnitudes vary. When we split the sample between private and public institutions, the effect of federally-funded R&D expenditures is larger for private universities (for both publications and citations), and we find that public universities' non-federally funded R&D expenditures do have a positive and significant, though significantly weaker, effect on research outputs. Given the low level of non-federal R&D expenditures at private universities, it is hardly surprising to find that this category of research funding has no effect for them. When the sample is split between Research I and non-Research I universities, the effect of federally-funded R&D expenditures is much weaker for both subsets of universities than it is for the whole sample and is statistically significant only for the non-Research I universities. This suggests that much of the effect observed in the full sample is coming from variations in the level of federally funded R&D expenditures and research outputs between the two groups, rather than within either group.

The year effects, which measure the impact of technological change from the instrumental variable regressions, are summarized in Figure 7. Once again we see

that there remains a significant unexplained positive time trend for both publications and citations. This growth is most pronounced for the Research I universities, and weakest for the non-Research I institutions. As was true in the OLS regressions, the trend growth was also larger for private universities than for the publics.

Discussion

Two results emerge from the regressions described in this section. The first is confirmation of a positive relationship between research funding and knowledge production, whether measured by raw publication numbers or weighting publications by the number of citations that they receive. Universities that receive more funding produce more articles and receive more citations to those articles. While the OLS results cannot demonstrate a causal effect of funding on knowledge production, they do confirm that the purposeful allocation of funding across institutions does correspond to the presence of more productive investigators who make collectively larger contributions to chemical knowledge. The IV estimates suggest that there is a causal effect of funding on knowledge production, at least for federal funding. If true, this implies that the marginal effect of increased funding would be relatively large. Using the coefficients from the full sample, they imply that an additional \$1 million would lead to the publication of 19 more articles and an additional 260 citations to those articles. This places the marginal cost of an additional article (a bit more than \$50,000) roughly in line with the average cost per article across all institutions.

The second striking result to emerge from our analysis is the rapid growth in knowledge production in chemistry over the 20 years from 1990 through 2009. At the moment this increase appears in our analysis as a residual time effect. In other words, it is unexplained by any of the measurable input variables, but can be considered a proxy for technological change (Solow 1957). Compared to the analysis in Adams and Griliches (1998) for the 1980s, this appears to be a departure from past experience. Given the coincidence of its timing with the spread of automatic laboratory data collection and analysis using personal computers and the internet, it is possible that the growth in the number of articles published relative to measured inputs reflects an IT mediated increase in the efficiency of academic chemistry. Alternatively, given the greater emphasis on publications and bibliometric measures of influence, it is possible that rising numbers of publications and citations to those publications reflect changes in the ethos of academic chemistry that have reduced the knowledge content of individual publications. Lacking a direct measure of the knowledge contributions of individual articles, we cannot at present distinguish between these alternative mechanisms or apportion the sources of the increase between them.

Conclusion

In light of the substantial investments made by both federal and non-federal sources in supporting university R&D activities, it is quite reasonable to wonder whether these investments are productive. Only a few previous studies have sought to elucidate these relationships. Several have focused on the level of individual

investigators and grants, a perspective that we believe neglects significant spillover effects across researchers. As we have argued, we believe that it is most appropriate to consider relationships at the level of the discipline and to concentrate on aggregate and institution-level variations.

As we have shown here, for the case of academic chemistry, analysis at this level yields important insights about knowledge production and its relationship to sponsored research funding from both federal and non-federal sources. While we have documented the fact that funding goes to those institutions that are most productive, and that there appears to be a causal effect of funding on subsequent knowledge production, we have also documented a puzzling increase in scholarly output among academic chemists in the years we have selected for study. This suggests that technological change may have shifted the production function, increasing the federal government's return on investment. Resolving the sources of this increase will require further effort to refine measurements of the knowledge embodied in academic publications.

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 (<http://webcaspar.nsf.gov>). Data are available annually since 1973 for total and federally funded R&D expenditures by discipline. They are obtained from survey 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 universe of institutions represented in this data by summing real federally funded R&D expenditures (in prices of 2000) for chemistry and chemical engineering between 1990 and 2009 and then ranking institutions in descending order. We initially selected the top 150 institutions but were obliged to drop three of these from the analysis because of inconsistencies in coverage. Before adopting this sampling strategy, we examined several other rankings, using total R&D expenditures and using nominal rather than real expenditures. The lists produced in each case were quite similar.

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 data are collected at the department level.

These data are available from <http://webcaspar.nsf.gov>.

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 and Citations

Publication and citation data were computed by Thomson Reuters, Research Analytics from the data underlying the Web of Science publication and citation database. Thomson Reuters subject area experts categorize journals into subject classes based on detailed analysis of the content and focus of the journals. Our research began with the full set of journals that Thomson Reuters categorizes as Chemistry and Chemical Engineering. We also conducted an analysis of all journal titles indexed by Thomson Reuters and added a small number of additional journals that contain significant chemistry content.

We then worked closely with Thomson Reuters staff to match publications by author affiliation to universities in our sample. In addition to institution name, we considered city, state and zip code information associated with authors to verify the accuracy of article linkages.

After verifying the full list of publications, Thomson Reuters analyzed them to produce summary statistics describing the number of publications each year produced by each institution, the number of citations that those publications received in 3 and 5 year windows beginning with the publication year, and a variety of other citation related metrics.

Fall Student Enrollment

These data come from the IPEDS Enrollment Survey and were downloaded from <http://webcaspar.nsf.gov>. The original source is the Higher Education General Information Survey (HEGIS) and the Integrated Postsecondary Education Data

System (IPEDS) that is conducted by the Department of Education's National Center for Education Statistics (NCES).

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.
- Apon, Amy W., Linh B. Ngo, Michael E. Payne and Paul Wilson (2014). "Assessing the Effect of High Performance Computing Capabilities on Academic Research Output." Working Paper, Clemson University.
- 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.
- Couzin, Jennifer and Greg Miller (2007). "Boom and bust." *Science* 316, pp. 356-61.
- 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, pp. 1168-1177.
- 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.

Solow, Robert M. (1957). "Technical Change and the Aggregate Production Function."

The Review of Economics and Statistics, 39, no. 3, pp. 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.

Endnotes

¹ A recent National Science and Technology Council (2008, pp. 6, 14) report, for example, observes: “The pragmatic reality facing Federal agencies is that the resources available for investing in research are limited...” and argues that there is a need for more systematic, quantitative models relating funding inputs to a variety of significant scientific outputs.

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.

³ From 1998 to 2004, NSF R&D obligations increased from \$2.3 billion to \$3.8 billion, measured in current dollars. While not as impressive as the increase in NIH budgets, this represents a significant infusion of funding and a departure from the relatively stable funding in the years before and after these dates (NSF Survey of Federal Funds for Research and Development; <http://webcaspar.nsf.gov>; accessed 9/25/2014).

⁴ To illustrate this point, of the 150 institutions we initially examined, there were 102 chemical engineering departments that reported faculty numbers in the American Chemical Society directory and 48 that did not. However, only 22 institutions reported zero amounts of federally funded R&D expenditures for chemical engineering research in every year, and there are no institutions for which

Web of Science recorded zero chemical engineering publications in all years. The mismatch in classification across the different sources used in our analysis suggests that attempting to analyze these fields separately would likely cause more problems than it solves. Numerically, chemistry accounts for the bulk of publications, R&D expenditures and faculty, and results that are restricted to chemistry look similar to those we report below. Results for chemical engineering resemble in sign and magnitude the results for chemistry, but effects are not significant, possibly because of the smaller sample and likely discrepancies in classification across the different data sources.

⁵ The three institutions dropped from our sample included two academic medical centers—the University of California San Francisco and the University of Texas M.D. Anderson Cancer Center—which reported no chemistry faculty, graduate students or postdocs for much of the study period, and the Oregon Institute of Science and Technology, which disappears from the data after 2001.

⁶ Additional details on the construction of the dataset used in our analysis are provided in the appendix to this paper.

⁷ Aggregate faculty numbers were essentially constant over these 20 years. We are in the process of hand collecting data from biannual faculty directories produced by the American Chemical Society, but because individual departments are not represented in all years, the data are not yet of sufficient quality to be reported here. Nonetheless, we can rule out the possibility that there was any significant increase in aggregate faculty numbers.

⁸ 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.

⁹ For a somewhat different approach to modeling knowledge production, which accounts for the joint production of publications and graduate student training, see Apon et al (2014), who use a stochastic frontier production function model.

¹⁰ Initially we had hoped to directly estimate the lag-structure of the relationship by including multiple years of lagged inputs on the right hand side of the equation. It became apparent, however, that these variables were highly collinear, making it impossible to precisely estimate their separate effects. When it became clear that we could not use this strategy, we experimented with a number of ad hoc specifications, including the average levels of inputs over the previous five years, a declining sum of the previous five years' inputs, and an inverted V weighting structure. All produced essentially the same results as a one-period lag.

¹¹ The vast majority of articles in our data have authors with only one institutional affiliation. However, when the authors of an article have affiliations with more than one institution, we credit each institution with the publication. Because there was a small increase in the number of multi-affiliation articles over time, this approach introduces a small upward bias into the time trend, but it can account at most for an increase of perhaps 5-10% in the total number of articles over the period.

¹² Increasing the citation window to five years yields very similar results.

¹³ We say "relative" quality because, as noted earlier, the average number of citations per article increased substantially between 1990 and 2009, which we

interpret as a change in citation practices that may reflect the greater visibility of publications as a result of the movement toward electronic publication and distribution. Thus, the absolute number of citations is unlikely to reflect the absolute quality of an article, but it will still provide evidence about the relative quality of articles published at the same time.

Figure 1: Sample Institution Shares of Key Input and Output Variables, 1990-2009

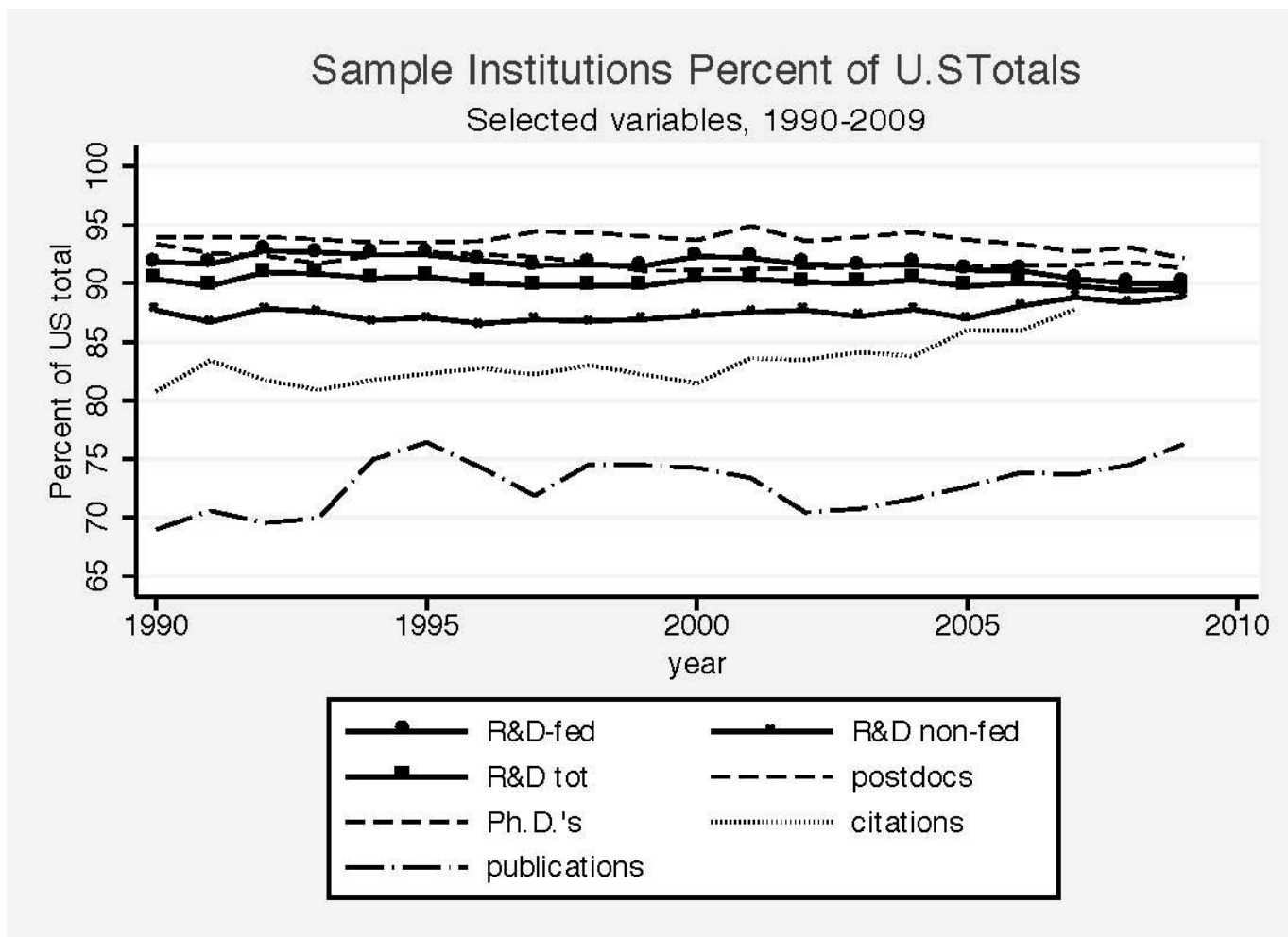


Figure 2: Indexes of R&D Expenditures, Full-time Faculty, Postdoctoral Researchers, and Ph.D.s Awarded, Sample Institutions, 1990-2009 (1990 = 100)

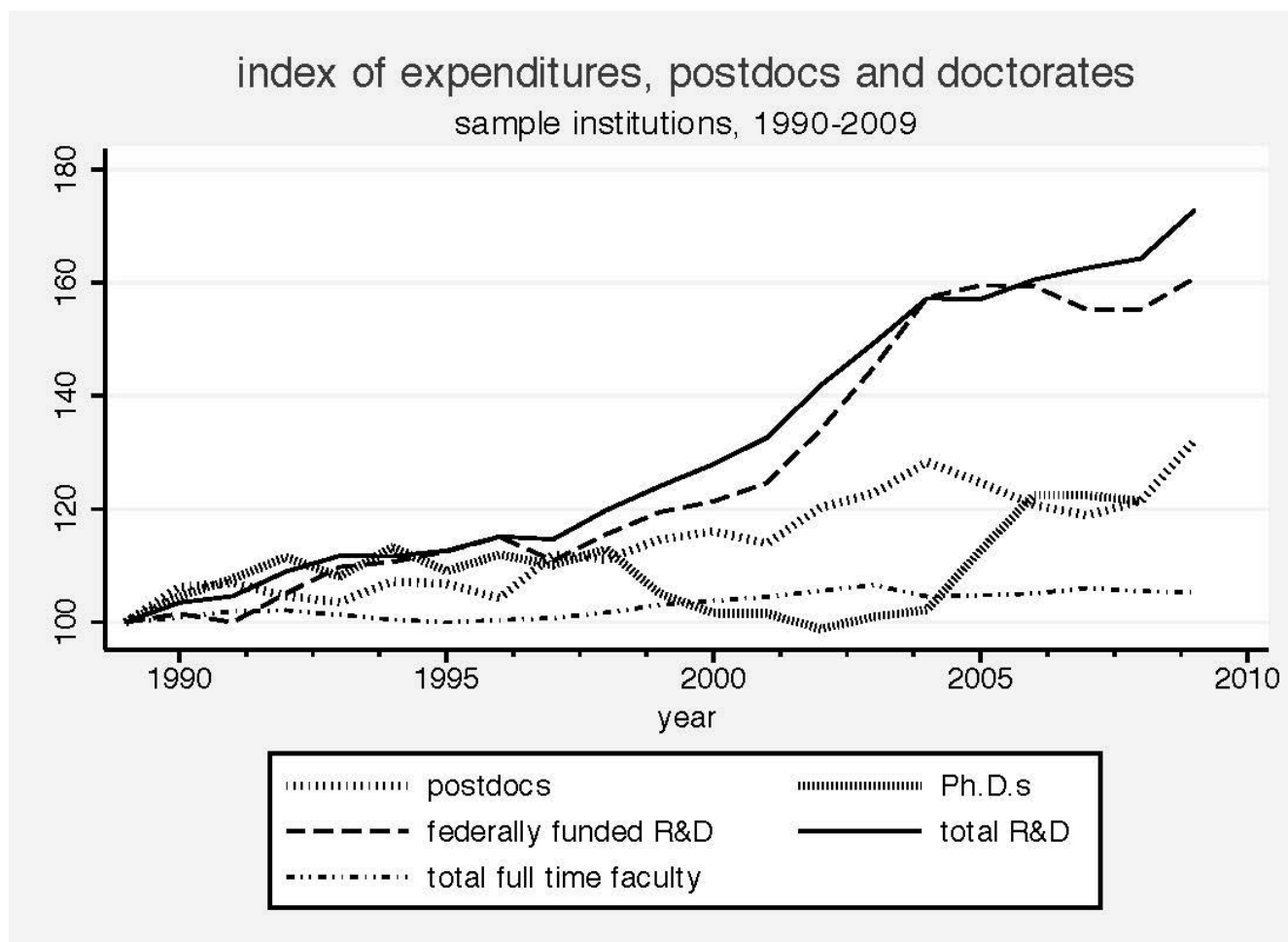


Figure 3: Indexes of R&D Expenditures, Publications, and Citations (three-year window), 1990-2009 (1990=100)

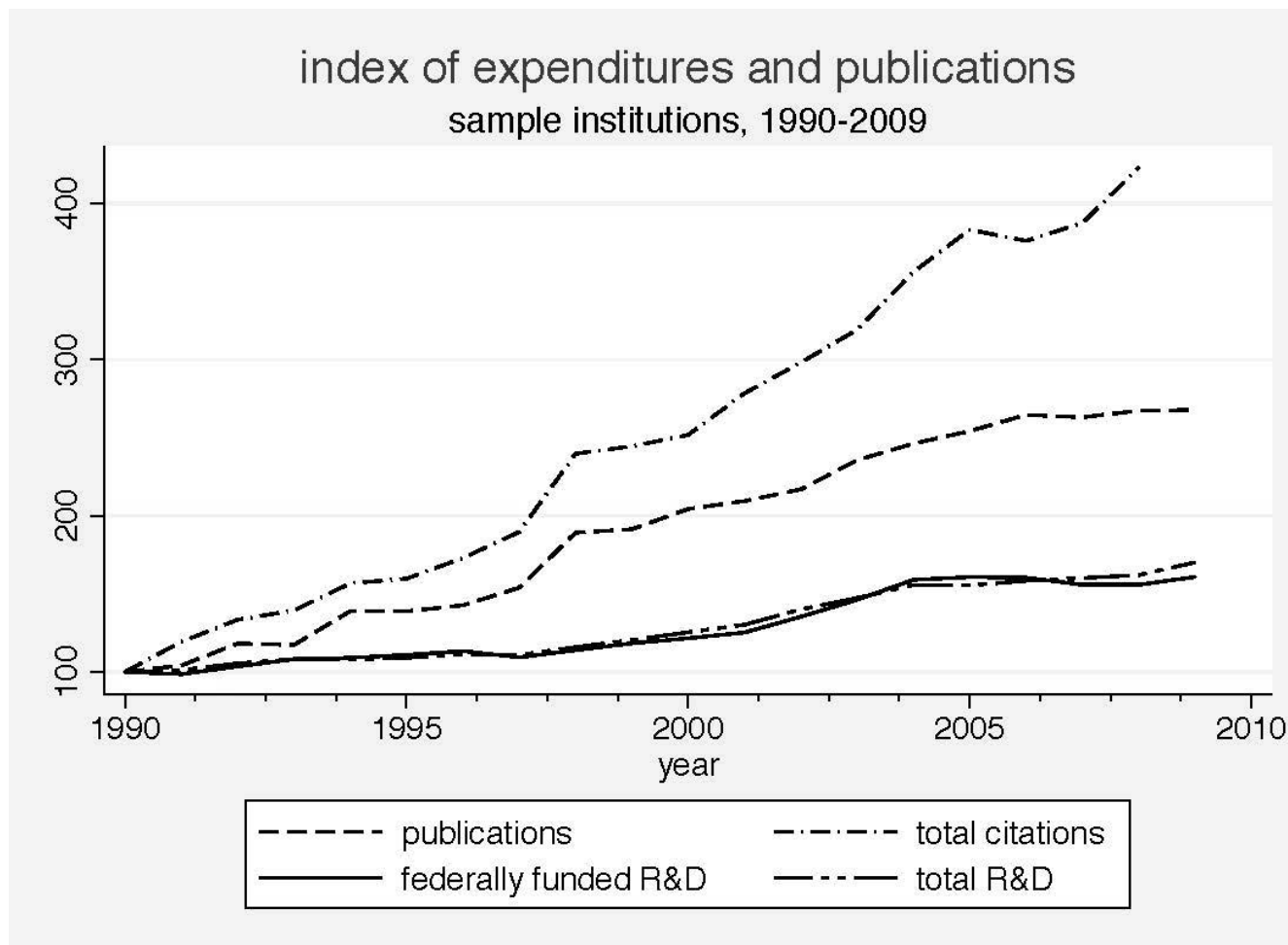


Figure 4: Average and Median Cost per Publication (Constant 2000 \$), Sample Institutions, 1990-2009

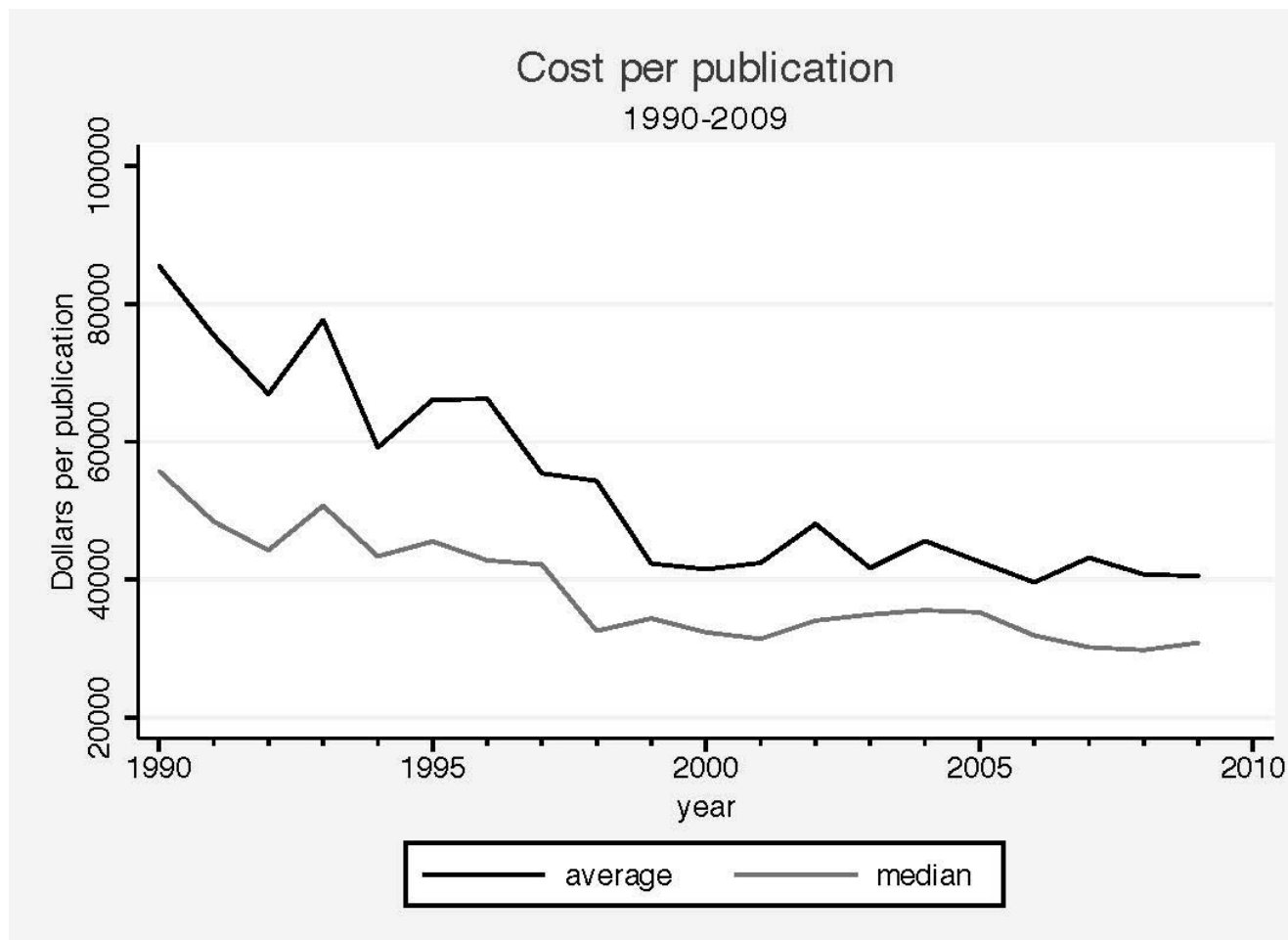


Figure 5: Median Cost per Publication (constant 2000 \$), Full Sample and by Type of University, 1990-2009

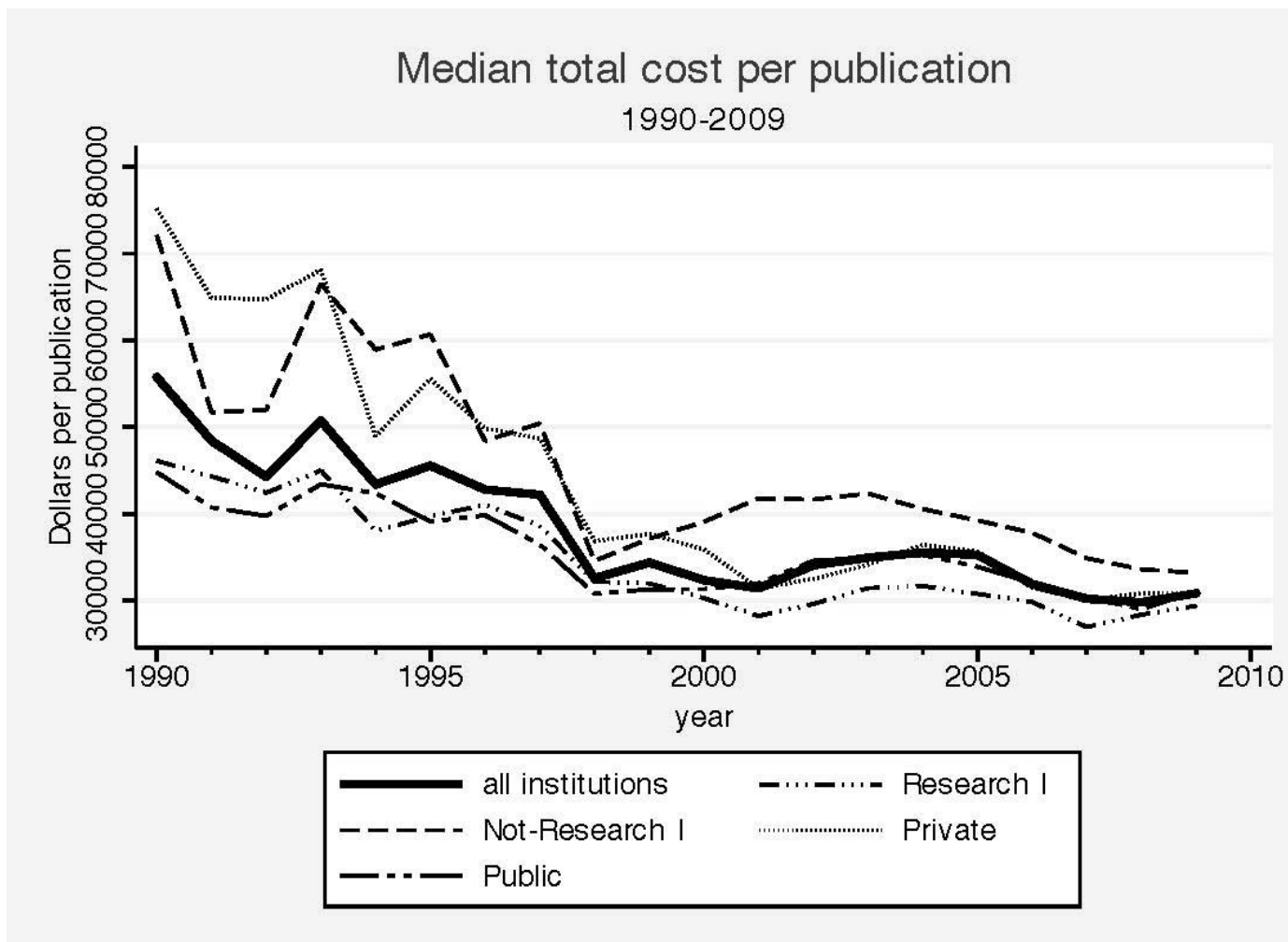


Figure 6a: Year Effects on Number of Publications from OLS Regression, 1991-2009

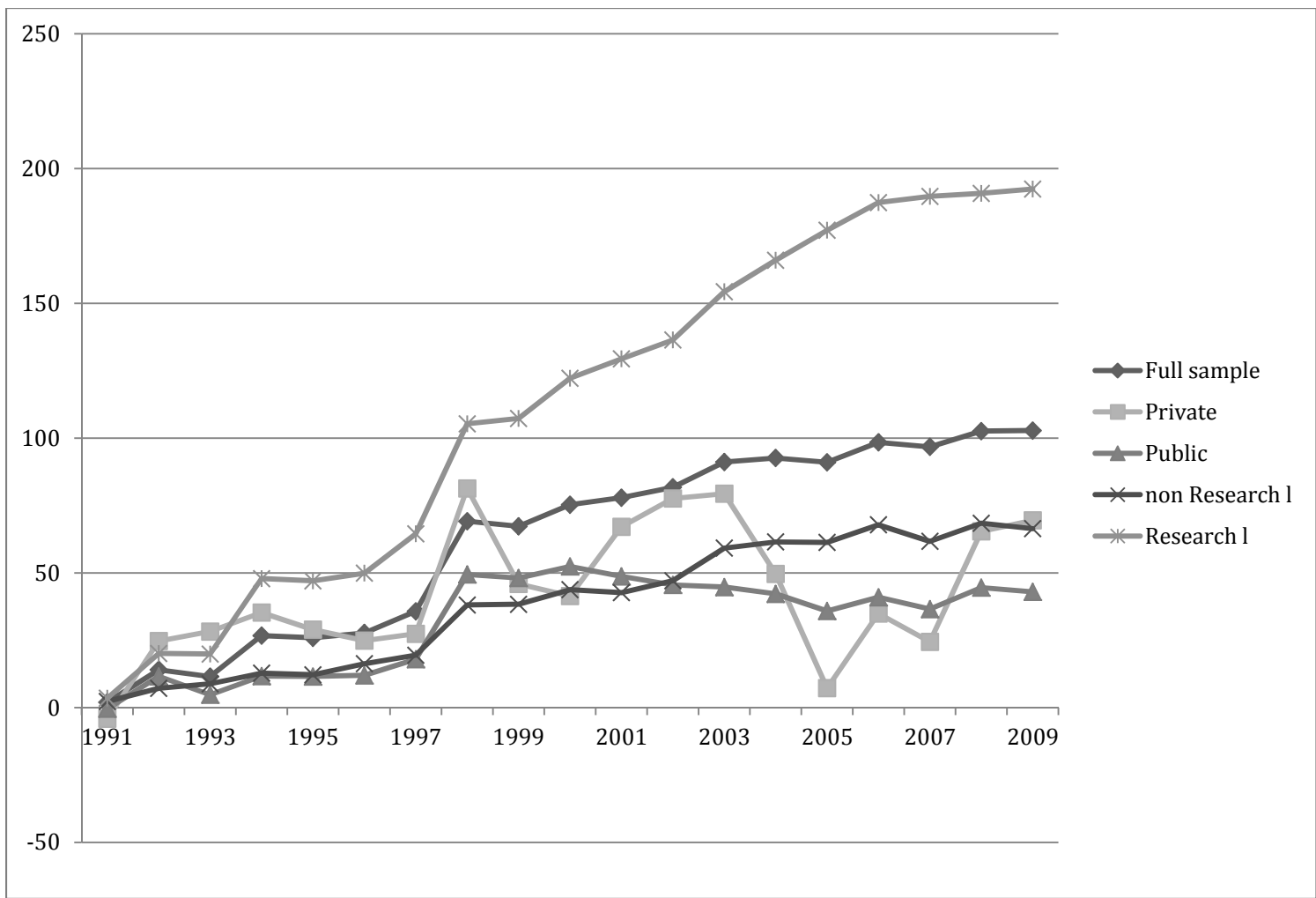


Figure 6b: Year Effects on Number of Citations from OLS Regressions, 1991-2008

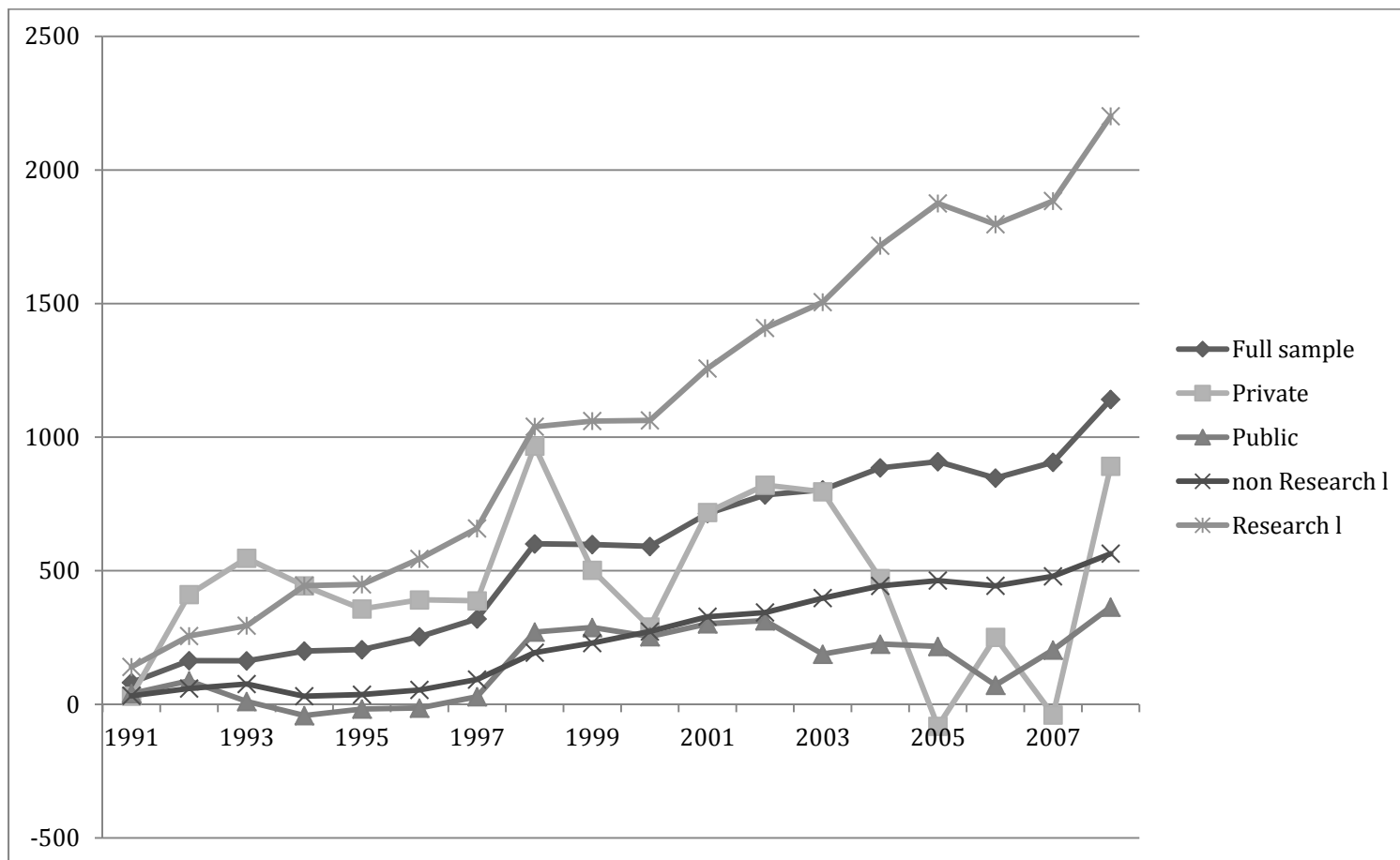


Figure 6c: Year Effects on Average Number of Citations per Article from OLS Regressions, 1991-2008

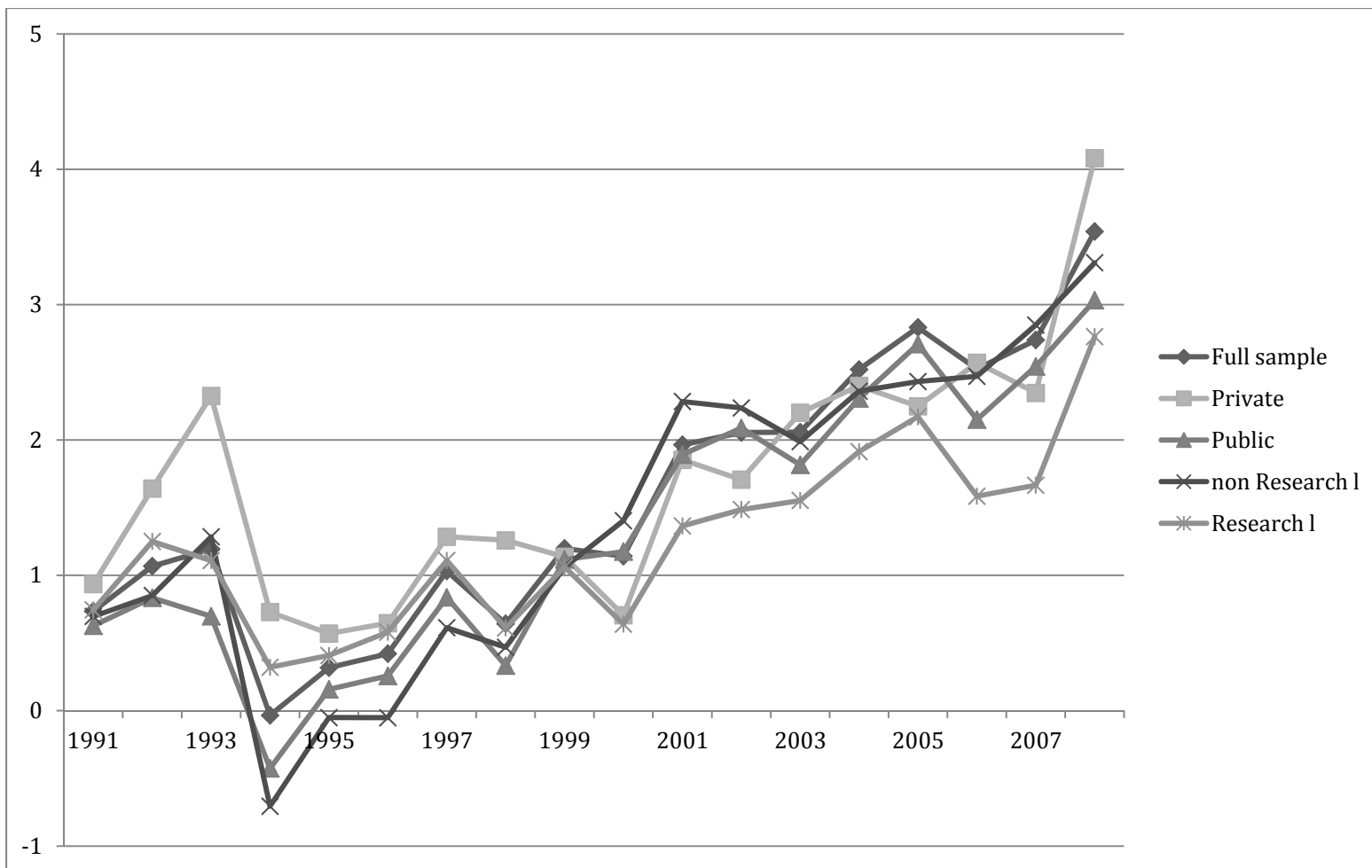


Figure 7a: Year Effects on Number of Publications from IV Regressions

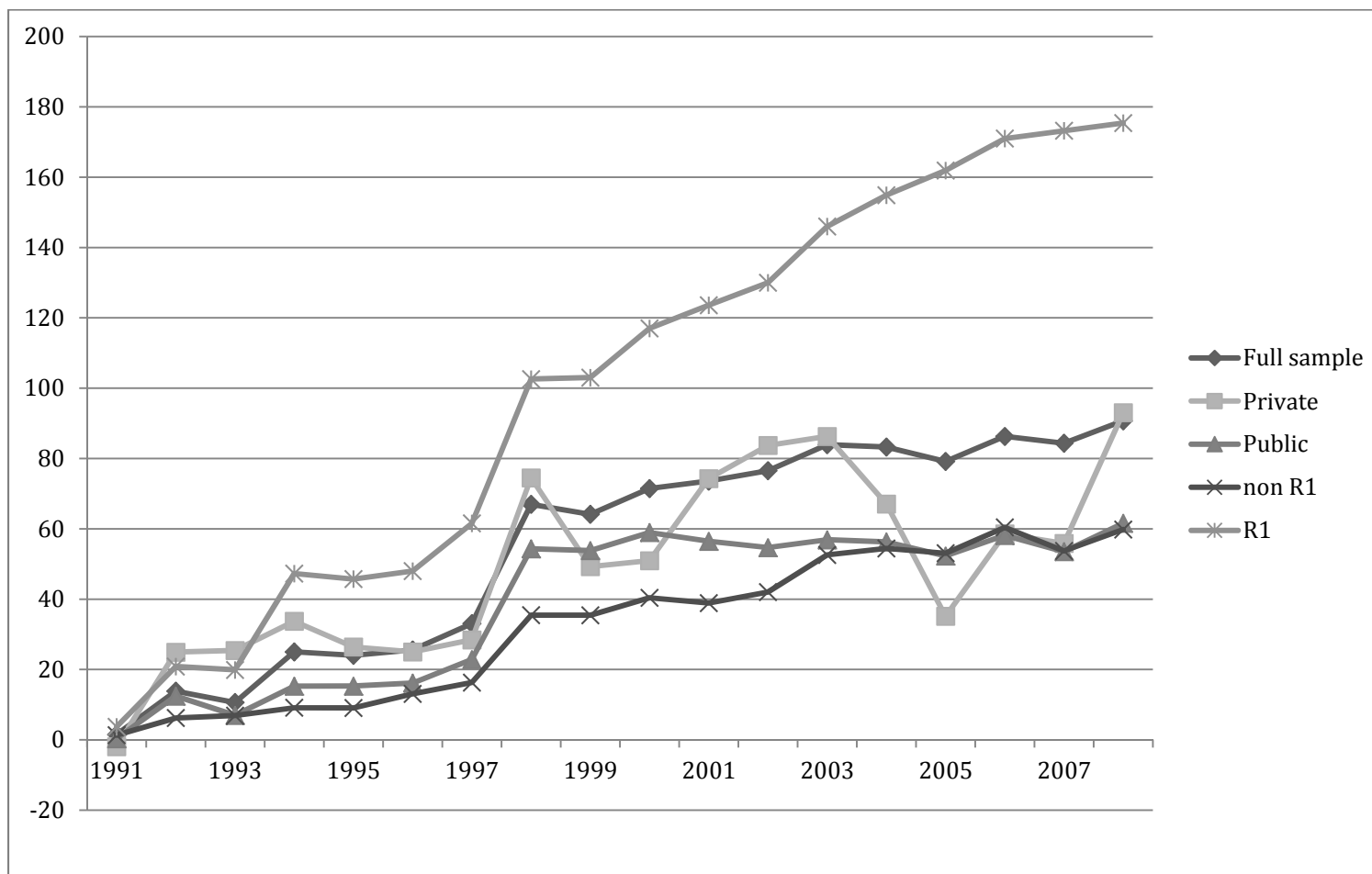


Figure 7b: Year Effects on Citations from IV Regressions

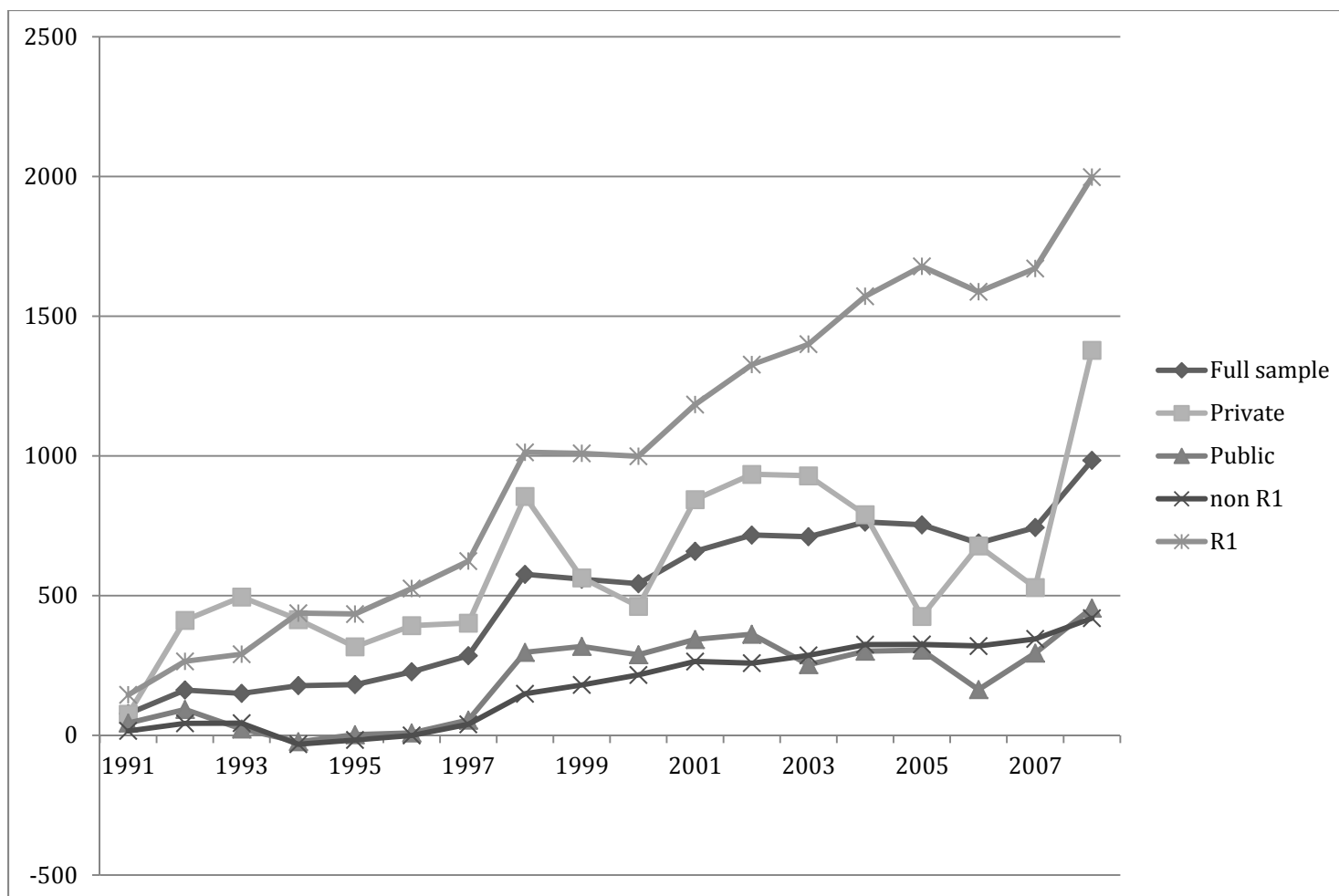


Table 1
Annual Average values of R&D Expenditures and other characteristics of Sample Institutions,
by Total Federally Funded R&D Expenditures, 1990-2009

University	R&D Expenditures by source (\$1,000s)			Postdoctoral		Citations	
	Federal	Non-Federal	Total	Ph.D.s	Researchers	Publications	over 3 years
Massachusetts Institute of Technology	\$28,262.9	\$8,574.4	\$36,837.3	66	141	522	5,126
California Institute of Technology	\$21,095.5	\$4,501.9	\$25,597.4	35	103	301	2,830
Johns Hopkins University	\$18,735.2	\$1,273.9	\$20,009.1	17	32	253	2,496
University of California-Berkeley	\$18,705.6	\$8,385.2	\$27,090.8	76	166	632	5,681
Stanford University	\$18,457.3	\$4,082.3	\$22,539.6	44	84	364	3,663
Harvard University	\$17,334.5	\$2,121.1	\$19,455.6	27	129	509	7,795
Pennsylvania State U, All Campuses	\$16,183.6	\$11,788.7	\$27,972.2	44	58	446	2,652
University of Illinois at Urbana-Champaign	\$15,787.8	\$8,835.3	\$24,623.1	52	40	424	3,216
University of Texas at Austin	\$15,527.6	\$14,590.3	\$30,117.9	64	91	349	2,714
University of California-Los Angeles	\$15,171.2	\$4,575.5	\$19,746.7	35	85	330	3,439
University of Colorado, All Campuses	\$14,894.9	\$4,523.5	\$19,418.4	29	83	256	1,855
University of Minnesota, All Campuses	\$14,398.6	\$6,406.9	\$20,805.5	55	75	472	3,429
Cornell University, All Campuses	\$13,828.9	\$6,030.0	\$19,858.9	35	55	357	3,093
University of Wisconsin-Madison	\$13,779.3	\$8,679.1	\$22,458.4	56	45	426	3,101
University of Pennsylvania	\$13,773.4	\$2,316.6	\$16,090.0	30	81	342	3,321
University of California-San Diego	\$12,701.0	\$2,962.1	\$15,663.1	21	64	310	3,213
Northwestern Univ	\$12,514.9	\$4,934.9	\$17,449.8	38	84	377	3,846
Rutgers the State Univ of NJ, All Campuses	\$12,373.2	\$5,572.0	\$17,945.2	27	43	256	1,431
University of Washington - Seattle	\$12,326.7	\$4,411.5	\$16,738.2	37	47	330	3,625
Purdue University, All Campuses	\$12,134.1	\$8,695.4	\$20,829.5	67	45	401	2,535
University of Michigan, All Campuses	\$11,832.4	\$5,095.9	\$16,928.2	48	47	452	4,162
Georgia Institute of Technology, All Campuses	\$11,249.9	\$11,690.5	\$22,940.3	35	28	257	2,321
University of Utah	\$11,060.0	\$4,505.2	\$15,565.1	30	52	297	2,315
University of Pittsburgh, All Campuses	\$10,519.2	\$2,503.7	\$13,022.9	30	59	254	2,074
University of North Carolina at Chapel Hill	\$10,319.3	\$2,725.4	\$13,044.7	35	57	328	3,039
Texas A&M University, All Campuses	\$10,292.9	\$18,344.9	\$28,637.8	58	84	403	2,632
Ohio State University, All Campuses	\$10,044.5	\$9,894.5	\$19,939.0	40	57	317	1,838
Princeton University	\$9,976.0	\$6,087.2	\$16,063.3	28	62	181	1,206
University of Notre Dame	\$9,974.5	\$1,994.2	\$11,968.7	17	44	177	1,316
University of Massachusetts at Amherst	\$9,542.5	\$8,312.0	\$17,854.5	35	35	237	1,434
Arizona State University Main	\$9,279.7	\$6,821.6	\$16,101.3	14	30	133	1,095
University of California-Irvine	\$9,191.4	\$3,516.3	\$12,707.7	25	46	211	1,671
Columbia University in the City of New York	\$8,969.1	\$2,676.5	\$11,645.6	25	59	259	2,501
University of California-Santa Barbara	\$8,871.8	\$3,185.1	\$12,056.9	23	58	241	2,406
University of Arizona	\$8,857.6	\$4,843.1	\$13,700.7	24	48	242	1,306
University of Florida	\$8,704.1	\$6,188.6	\$14,892.7	46	59	371	2,134
University of Delaware	\$8,665.0	\$5,124.4	\$13,789.4	27	39	256	1,388
University of South Carolina, All Campuses	\$8,569.4	\$6,292.3	\$14,861.6	21	30	122	908
Yale University	\$8,560.5	\$2,002.2	\$10,562.6	22	38	236	2,711
North Carolina State University at Raleigh	\$8,496.7	\$11,125.6	\$19,622.3	26	47	264	1,382
University of Chicago	\$8,460.5	\$2,389.5	\$10,850.0	21	25	185	1,650
University of California-Davis	\$8,071.5	\$3,523.2	\$11,594.7	31	45	335	2,098
Michigan State University	\$7,530.1	\$8,531.5	\$16,061.6	32	46	247	1,361
University of Virginia, All Campuses	\$7,419.8	\$1,689.0	\$9,108.8	20	40	181	1,531
Case Western Reserve University	\$7,389.1	\$4,425.7	\$11,814.8	26	15	195	1,616
Indiana University, All Campuses	\$7,274.4	\$6,195.9	\$13,470.3	20	44	216	1,631
University of Tennessee Univ-Wide Adm Cent Off	\$7,188.0	\$3,434.5	\$10,622.4	19	25	234	1,320
University of Maryland at College Park	\$7,179.9	\$5,100.2	\$12,280.1	21	18	186	1,075
New Mexico State University, All Campuses	\$7,061.8	\$1,118.6	\$8,180.4	6	4	50	362
Louisiana State Univ, All Campuses	\$6,999.2	\$7,160.4	\$14,159.6	22	31	176	889
Colorado State University	\$6,974.7	\$2,495.8	\$9,470.5	17	53	144	1,038
University of Southern California	\$6,956.0	\$3,270.7	\$10,226.7	22	51	196	1,431
SUNY at Buffalo, All Campuses	\$6,952.3	\$5,981.4	\$12,933.7	27	45	182	1,184

University	R&D Expenditures by source (\$1,000s)			Postdoctoral		Citations	
	Federal	Non-Federal	Total	Ph.D.s	Researchers	Publications	over 3 years
Virginia Polytechnic Institute and State Univ	\$5,994.4	\$7,191.8	\$13,186.1	24	27	204	899
Carnegie Mellon University	\$5,893.6	\$1,875.6	\$7,769.2	21	28	160	1,237
University of Rochester	\$5,871.4	\$942.1	\$6,813.6	17	31	145	1,242
Rice University	\$5,796.1	\$2,454.0	\$8,250.0	20	32	145	1,437
Emory University	\$5,690.6	\$2,680.7	\$8,371.2	16	39	193	1,741
SUNY at Stony Brook, All Campuses	\$5,449.7	\$3,986.2	\$9,435.9	16	19	198	1,422
University of Southern Mississippi	\$5,425.5	\$1,365.7	\$6,791.1	11	17	91	276
University of Oklahoma, All Campuses	\$5,292.0	\$9,009.5	\$14,301.4	15	29	120	670
Rensselaer Polytechnic Institute	\$5,024.4	\$2,522.7	\$7,547.1	19	22	117	553
Washington University	\$4,978.8	\$2,397.4	\$7,376.2	18	28	230	2,577
University of Kansas, All Campuses	\$4,898.4	\$4,978.9	\$9,877.3	22	29	162	1,060
University of Nebraska Central Admin Sys Off	\$4,779.8	\$3,782.4	\$8,562.2	12	17	149	846
University of Houston	\$4,707.2	\$5,677.1	\$10,384.3	22	30	156	917
Vanderbilt University	\$4,677.5	\$927.0	\$5,604.5	11	20	200	2,046
Wayne State University	\$4,665.9	\$4,238.4	\$8,904.2	22	21	153	1,107
Clemson University	\$4,541.9	\$5,466.3	\$10,008.2	14	12	147	695
University of Alabama in Huntsville	\$4,192.9	\$1,665.8	\$5,858.7	0	1	27	81
Iowa State University	\$4,150.9	\$3,794.8	\$7,945.7	35	22	335	1,906
University of California-Santa Cruz	\$4,117.7	\$1,602.1	\$5,719.8	8	21	72	476
University of Oregon	\$4,069.5	\$1,291.8	\$5,361.3	7	25	74	427
University of Illinois at Chicago	\$3,983.0	\$2,412.3	\$6,395.4	22	15	187	1,346
University of Iowa	\$3,951.2	\$2,708.4	\$6,659.6	19	21	171	1,340
Montana State University - Bozeman	\$3,931.4	\$1,260.0	\$5,191.4	6	17	50	274
University of New Mexico, All Campuses	\$3,910.9	\$1,534.7	\$5,445.7	11	15	115	768
University of California-Riverside	\$3,896.9	\$1,568.0	\$5,464.9	11	34	156	1,221
Boston College	\$3,887.5	\$1,219.3	\$5,106.8	9	19	63	611
Florida State University	\$3,886.7	\$5,761.0	\$9,647.6	11	42	134	730
University of PR Rio Piedras Campus	\$3,799.3	\$386.2	\$4,185.6	7	6	37	102
Kansas State University	\$3,753.1	\$1,303.4	\$5,056.6	9	14	89	473
CUNY City College	\$3,720.1	\$707.1	\$4,427.2	0	11	37	186
Brigham Young University, All Campuses	\$3,579.1	\$1,604.5	\$5,183.5	10	12	86	390
Mississippi State University	\$3,512.1	\$3,248.1	\$6,760.2	5	10	45	139
New York University	\$3,475.1	\$1,212.0	\$4,687.1	8	24	116	1,024
University of Alabama	\$3,451.5	\$1,620.2	\$5,071.7	10	13	85	476
Duke University	\$3,447.2	\$1,864.0	\$5,311.2	14	32	205	2,459
University of Akron, All Campuses	\$3,446.6	\$8,058.6	\$11,505.2	40	36	138	670
University of Dayton	\$3,409.9	\$594.0	\$4,004.0	0	0	34	132
Washington State University	\$3,344.5	\$1,906.0	\$5,250.6	7	12	126	665
University of Maryland Baltimore County	\$3,326.2	\$789.6	\$4,115.8	5	13	126	968
Georgetown University	\$3,275.2	\$824.1	\$4,099.4	6	6	88	709
Oregon State University	\$3,269.4	\$1,196.9	\$4,466.3	12	12	102	500
Brown University	\$3,219.0	\$2,434.5	\$5,653.5	11	11	82	541
University of Arkansas, Main Campus	\$3,156.9	\$2,457.6	\$5,614.5	8	11	68	371
Northeastern University	\$3,131.9	\$1,836.6	\$4,968.5	9	11	80	500
University of Kentucky, All Campuses	\$3,103.8	\$2,047.1	\$5,150.9	14	18	180	1,100
Rockefeller University	\$2,961.8	\$2,024.1	\$4,985.9	0	7	44	690
Auburn University, All Campuses	\$2,959.0	\$3,926.6	\$6,885.6	13	9	84	327
University of Tulsa	\$2,908.2	\$3,263.4	\$6,171.6	7	2	22	16
University of Cincinnati, All Campuses	\$2,903.3	\$3,477.4	\$6,380.7	21	14	174	1,046
Boston University	\$2,878.2	\$346.1	\$3,224.3	8	6	131	1,072
New Mexico Institute of Mining and Technology	\$2,873.2	\$3,732.8	\$6,606.1	3	0	18	72
CUNY Hunter College	\$2,839.1	\$1,175.0	\$4,014.1	0	0	41	224
Tufts University	\$2,838.6	\$934.3	\$3,772.9	7	14	89	872
North Dakota State University, All Campuses	\$2,815.6	\$1,900.4	\$4,716.0	6	12	87	349
Colorado School of Mines	\$2,804.7	\$2,410.7	\$5,215.3	9	7	48	185
Virginia Commonwealth University	\$2,727.7	\$1,659.4	\$4,387.1	7	8	110	625
Clark Atlanta University	\$2,714.2	\$398.8	\$3,112.9	1	5	18	43

University	R&D Expenditures by source (\$1,000s)			Postdoctoral		Citations	
	Federal	Non-Federal	Total	Ph.D.s	Researchers	Publications	over 3 years
Lehigh University	\$2,678.7	\$3,227.4	\$5,906.2	16	19	79	361
University of Georgia	\$2,654.3	\$5,099.0	\$7,753.4	17	17	221	1,333
University of Connecticut, All Campuses	\$2,532.5	\$2,640.9	\$5,173.4	23	16	178	928
Syracuse University, All Campuses	\$2,479.3	\$973.5	\$3,452.8	10	12	79	410
West Virginia University	\$2,440.8	\$1,210.9	\$3,651.7	7	9	67	308
Tulane University	\$2,384.1	\$1,896.6	\$4,280.6	9	12	88	567
Oklahoma State University, All Campuses	\$2,350.0	\$2,376.0	\$4,726.0	10	5	74	363
Brandeis University	\$2,233.7	\$659.7	\$2,893.4	7	14	47	394
Jackson State University	\$2,206.5	\$327.0	\$2,533.5	1	9	47	221
Illinois Institute of Technology	\$2,169.6	\$946.5	\$3,116.1	9	5	34	125
Clarkson University	\$2,144.8	\$2,596.0	\$4,740.8	10	21	57	288
New Jersey Institute Technology	\$2,136.9	\$2,441.3	\$4,578.2	5	1	36	129
Texas Tech University	\$2,104.0	\$3,913.5	\$6,017.5	11	22	85	433
University of Missouri, Columbia	\$2,031.3	\$3,841.1	\$5,872.4	13	11	124	568
University of Wyoming	\$1,804.3	\$2,737.7	\$4,542.0	8	12	44	197
University of Hawaii at Manoa	\$1,744.4	\$455.0	\$2,199.4	4	10	72	398
Dartmouth College	\$1,690.0	\$747.9	\$2,438.0	6	12	65	498
Drexel University	\$1,653.9	\$953.9	\$2,607.8	6	6	71	435
Utah State University	\$1,629.6	\$898.0	\$2,527.6	3	5	48	283
Norfolk State University	\$1,629.3	\$37.6	\$1,666.9	0	0	6	4
University of New Hampshire	\$1,583.8	\$659.8	\$2,243.6	5	4	40	152
San Francisco State University	\$1,568.0	\$99.2	\$1,667.2	0	0	18	68
Howard University	\$1,530.4	\$293.0	\$1,823.4	5	3	57	170
University of Denver	\$1,510.9	\$283.8	\$1,794.8	2	5	13	49
Polytechnic University	\$1,491.9	\$1,253.2	\$2,745.2	9	14	43	195
California State University-Los Angeles	\$1,481.8	\$76.8	\$1,558.6	0	0	29	81
University of Idaho	\$1,467.7	\$1,414.4	\$2,882.1	7	9	66	318
Georgia State University	\$1,447.5	\$1,659.7	\$3,107.3	3	21	50	263
University of Missouri, Rolla	\$1,446.9	\$1,867.3	\$3,314.3	10	4	60	232
University of Massachusetts Lowell	\$1,379.7	\$1,104.1	\$2,483.8	12	6	59	175
University of Louisville	\$1,378.5	\$898.8	\$2,277.4	7	7	51	369
University of Montana	\$1,340.3	\$448.6	\$1,788.9	3	5	21	66
University of South Florida	\$1,327.3	\$1,743.8	\$3,071.1	9	10	82	674
University of PR Mayaguez Campus	\$1,300.6	\$680.1	\$1,980.6	1	0	25	35
North Carolina Agricultural & Tech State Univ	\$1,281.5	\$89.5	\$1,371.0	0	0	13	24
Stevens Institute of Technology	\$1,252.7	\$400.2	\$1,652.9	4	4	27	118
Cleveland State University	\$1,178.6	\$518.8	\$1,697.4	5	4	24	107

Source: See Data Appendix.

Table 2
Annual Average Values of Key Variables, by Decade and University Characteristics

	Full Sample	Research Status		Control	
		Research 1	Other	Private	Public
1990-1999					
Federally Funded R&D (thousands)	\$ 5,203	\$ 7,190	\$ 2,559	\$ 6,010	\$ 4,823
Federally Funded Equipment Expenditures	\$ 515	\$ 704	\$ 263	\$ 582	\$ 483
Federally Funded Non-Equipment Expenditures	\$ 4,688	\$ 6,486	\$ 2,297	\$ 5,427	\$ 4,340
Non-Federally Funded R&D (thousands)	\$ 2,821	\$ 3,596	\$ 1,791	\$ 2,024	\$ 3,196
Non-Federally Funded Equipment Expenditures	\$ 253	\$ 311	\$ 177	\$ 184	\$ 286
Non-Federally Funded Non-Equipment Expenditures	\$ 2,568	\$ 3,285	\$ 1,614	\$ 1,840	\$ 2,910
Total R&D Expenditures	\$ 8,024	\$ 10,786	\$ 4,350	\$ 8,034	\$ 8,019
Percent R&D Federally Funded	64.6%	66.1%	62.5%	72.0%	61.1%
Full Time Faculty	33	40	24	28	35
Graduate Students Enrolled	150	190	96	126	161
Ph.D.s awarded	18	25	8	16	19
Postdoctoral Researchers	27	39	11	29	26
Number of publications	120	174	48	113	123
Total 3-year citations	832	1,295	211	1,041	734
2000-2009					
Federally Funded R&D (thousands)	\$ 6,954	\$ 9,491	\$ 3,571	\$ 7,048	\$ 6,909
Federally Funded Equipment Expenditures	\$ 524	\$ 687	\$ 307	\$ 483	\$ 543
Federally Funded Non-Equipment Expenditures	\$ 6,430	\$ 8,804	\$ 3,264	\$ 6,565	\$ 6,366
Non-Federally Funded R&D (thousands)	\$ 3,902	\$ 5,196	\$ 2,177	\$ 2,370	\$ 4,622
Non-Federally Funded Equipment Expenditures	\$ 344	\$ 450	\$ 204	\$ 192	\$ 416
Non-Federally Funded Non-Equipment Expenditures	\$ 3,558	\$ 4,746	\$ 1,974	\$ 2,179	\$ 4,206
Total R&D Expenditures	\$ 10,856	\$ 14,687	\$ 5,749	\$ 9,418	\$ 11,532
Percent R&D Federally Funded	66.2%	66.2%	66.2%	73.8%	62.6%
Full Time Faculty	34	41	26	29	36
Graduate Students Enrolled	158	205	95	135	168
Ph.D.s awarded	18	26	8	16	19
Postdoctoral Researchers	31	44	13	34	29
Number of publications	209	297	92	200	213
Total 3-year citations	1,650	2,465	562	1,857	1,552
Number of institutions	147	84	63	47	100

Table 3

OLS Panel Regressions Determinants of Publications, Citations, and Average number of Citations per Publication

Dependent Variable	Number of Publications		Number of citations(3 year horizon)		Average number of citations per publication	
	Model I	Model II	Model I	Model II	Model I	Model II
L.Federally Funded R&D		0.00609*** (0.00125)		0.0645*** (0.0177)		0.0000568 (0.0000414)
L.Non-Federally Funded R&D		0.00669*** (0.00177)		0.0693** (0.0240)		0.0000606 (0.0000476)
L.Total R&D Funding	0.00634*** (0.00107)		0.0665*** (0.0144)		0.0000584* (0.0000290)	
L.Full Time Faculty	-0.209 (0.224)	-0.213 (0.227)	-1.227 (3.171)	-1.263 (3.233)	-0.00175 (0.0119)	-0.00178 (0.0119)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	46.61*** (11.87)	47.10*** (11.86)	59.61 (152.8)	63.54 (155.0)	6.926*** (0.472)	6.929*** (0.477)
Observations	2933	2933	2933	2933	2933	2933
sigma_e	43.34	43.34	577.9	578.0	2.168	2.168
sigma_u	93.34	93.78	935.1	940.2	2.953	2.956
rho	0.823	0.824	0.724	0.726	0.650	0.650
r2_w	0.619	0.619	0.484	0.484	0.241	0.241
r2_b	0.757	0.751	0.624	0.615	0.194	0.188
r2_o	0.581	0.576	0.503	0.498	0.156	0.154

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Notes

All regressions include year and institution fixed effects. Estimated using STATA xtreg procedure with cluster robust standard errors.

Table 4
 OLS Panel Regressions. Determinants of Publications, Citations, and Average Citations per Publication, by
 University type

	<u>Full sample</u>	<u>Private</u>	<u>Public</u>	<u>not Research I</u>	<u>Research I</u>
Panel A -Dependent Variable: Number of Publications					
L.Federally Funded R&D	0.00609*** (0.00125)	0.00652*** (0.00173)	0.00611*** (0.00171)	0.00404* (0.00172)	0.00523*** (0.00131)
L.Non-Federally Funded R&D	0.00669*** (0.00177)	0.00293 (0.00347)	0.00775*** (0.00198)	0.00294 (0.00177)	0.00536** (0.00201)
L.Full Time Faculty	-0.213 (0.227)	-0.156 (0.418)	-0.282 (0.308)	0.0659 (0.203)	-0.228 (0.341)
Year Effects	Yes	Yes	Yes	Yes	Yes
Constant	47.10*** (11.86)	41.06 (23.75)	51.43*** (14.50)	17.41* (7.789)	82.24*** (18.25)
Observations	2933	940	1993	1256	1677
sigma_e	43.34	44.79	42.60	23.50	47.27
sigma_u	93.78	89.93	95.36	39.37	94.30
rho	0.824	0.801	0.834	0.737	0.799
r2_w	0.619	0.578	0.642	0.572	0.710
r2_b	0.751	0.799	0.718	0.605	0.675
r2_o	0.576	0.602	0.569	0.430	0.539

	<u>Full sample</u>	<u>Private</u>	<u>Public</u>	<u>not Research I</u>	<u>Research I</u>
Panel B -Dependent Variable: Number of Citations					
L.Federally Funded R&D	0.0645*** (0.0177)	0.0527 (0.0338)	0.0749*** (0.0201)	0.0343* (0.0131)	0.0566** (0.0202)
L.Non-Federally Funded R&D	0.0693** (0.0240)	0.0578 (0.0434)	0.0760** (0.0273)	0.0289* (0.0125)	0.0607* (0.0284)
L.Full Time Faculty	-1.263 (3.233)	-2.391 (4.238)	-0.853 (4.615)	0.272 (2.039)	-0.427 (4.865)
Year Effects	Yes	Yes	Yes	Yes	Yes
Constant	63.54 (155.0)	233.0 (327.6)	-29.95 (190.4)	-4.730 (71.68)	226.9 (251.6)
Observations	2933	940	1993	1256	1677
sigma_e	578.0	647.1	540.6	235.7	664.5
sigma_u	940.2	1252.2	733.1	284.4	1049.0
rho	0.726	0.789	0.648	0.593	0.714
r2_w	0.484	0.428	0.525	0.477	0.583
r2_b	0.615	0.632	0.673	0.466	0.478
r2_o	0.498	0.426	0.575	0.401	0.452

	<u>Full sample</u>	<u>Private</u>	<u>Public</u>	<u>not Research I</u>	<u>Research I</u>
Panel C -Dependent Variable: Average Number of Citations per Publication					
L.Federally Funded R&D	0.0000568 (0.0000414)	-0.0000453 (0.0000805)	0.000113** (0.0000428)	0.000218* (0.0000821)	0.0000349 (0.0000429)
L.Non-Federally Funded R&D	0.0000606 (0.0000476)	0.0000684 (0.0000958)	0.0000329 (0.0000492)	-0.00000232 (0.0000577)	0.000102 (0.0000598)
L.Full Time Faculty	-0.00178 (0.0119)	-0.0287* (0.0112)	0.0134 (0.0114)	0.0184 (0.0133)	-0.0124 (0.0115)
Year Effects	Yes	Yes	Yes	Yes	Yes
Constant	6.929*** (0.477)	9.662*** (0.836)	5.588*** (0.516)	4.355*** (0.475)	8.741*** (0.631)
Observations	2933	940	1993	1256	1677
sigma_e	2.168	2.620	1.906	2.396	1.953
sigma_u	2.956	3.950	2.052	2.269	2.732
rho	0.650	0.694	0.537	0.473	0.662
r2_w	0.241	0.167	0.317	0.269	0.245
r2_b	0.188	0.226	0.383	0.130	0.00450
r2_o	0.154	0.00721	0.311	0.200	0.0968

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Table 5:
First Stage IV Regressions Determinants of Number of Publications

	Endogenous Regressor	
	Federally Financed Chemistry R&D Expenditures	Non-Federally Financed Chemistry R&D Expenditures
L.Deflated federally financed R&D expenditures Math & Physics	0.0472* (0.02)	-0.0028 (0.008)
L.Deflated non-federally financed R&D expenditures Math & Physics	0.051 (0.027)	0.37*** (0.0307)
L.Full Time Faculty	-1.971 (6.363)	11.307 (6.245)
L.Fall Enrollment	0.178*** (0.025)	0.058** (0.019)
Year effects	Yes	Yes
Constant	915.25 (515.127)	197.97 (388.321)
Observations	2933	2933
F-stat (p-value)	38.07(0.000)	41.96(0.000)
sigma_e	1937.38	1460.47
sigma_u	4045.66	2042.82
rho	0.813	0.662
r2_w	0.2407	0.2589
r2_b	0.295	0.5267
r2_o	0.2848	0.4577

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Note: Estimates derived from first stage of STATA XTIVREG2 estimation with cluster robust standard errors.

Table 6

IV Regressions Regressions. Determinants of Publications, Citations, by University type. Instrumets: Math & Physics Federal R&D,
Math and Physics non-federal R& D, Fall enrollment

	Full sample	Private	Public	non Research I	Research I
Panel A -Dependent Variable: Number of Publications					
L.Federally Funded R&D	0.0190*** (0.00422)	0.0587* (0.0258)	0.0223*** (0.00466)	0.00884** (0.00298)	0.00510 (0.00312)
L.Non-Federally Funded R&D	0.00327 (0.00247)	-0.0510 (0.0272)	0.00781** (0.00258)	-0.00632 (0.00340)	0.00348 (0.00218)
L.Full Time Faculty	-0.181 (0.152)	-0.463 (0.532)	-0.315 (0.231)	0.342* (0.161)	-0.208 (0.173)
Constant	-7.493 (13.83)	-148.3 (131.8)	-15.78 (13.85)	14.78* (6.444)	88.22*** (18.58)
Observations	2933	940	1993	1256	1677
Sargan-Hansen statistics (p-value)	0.675(0.4114)	0.161(0.6882)	0.085(0.771)	30.015(0.000)	0.927(0.3357)
sigma_e	49.82	116.7	52.86	25.70	47.40
sigma_u	63.76	196.9	59.25	42.62	97.35
rho	0.621	0.740	0.557	0.733	0.808
r2_w	0.497	.	0.449	0.488	0.709
r2_b	0.763	0.749	0.790	0.203	0.685
r2_o	0.693	0.612	0.712	0.295	0.511

	Full sample	Private	Public	non Research I	Research I
Panel B -Dependent Variable: Number of Citations					
L.Full Time Faculty	-0.620 (2.233)	-7.373 (7.181)	-1.167 (3.400)	3.924 (2.046)	-0.661 (2.500)
L.Federally Funded R&D	0.260*** (0.0565)	0.700* (0.330)	0.312*** (0.0562)	0.150*** (0.0403)	0.132** (0.0410)
L.Non-Federally Funded R&D	0.00614 (0.0358)	-0.747 (0.387)	0.0703* (0.0355)	-0.0700 (0.0397)	0.00460 (0.0338)
Constant	-740.0*** (190.7)	-1791.8 (1731.2)	-999.9*** (184.9)	-179.1* (71.94)	-107.7 (269.5)
Observations	2933	940	1993	1256	1677
Sargan-Hansen statistics (p-value)	0.009(0.924)	0.008(0.9295)	0.731(0.393)	19.157(0.000)	1.685(0.194)
sigma_e	686.7	1532.5	705.7	286.9	687.7
sigma_u	674.3	2327.9	705.1	272.8	854.1
rho	0.491	0.698	0.500	0.475	0.607
r2_w	0.271	.	0.191	0.226	0.554
r2_b	0.730	0.658	0.778	0.380	0.633
r2_o	0.621	0.502	0.645	0.350	0.572

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001