Funding the Ivory Tower: The Effects of NSF Institutional Grants on Universities and Local Innovation*

Gabriele Cristelli†

June 2025

Abstract

This paper studies the effects of the NSF Science Development Program on universities and local innovation, combining historical data from scientific publications, doctoral dissertations, and patents. Introduced in 1965, the program awarded large institutional grants to natural science and engineering departments at U.S. research universities. I exploit top-ranked universities excluded from the program as a comparison group in a difference-in-differences research design. First, I find that Science Development awards increased faculty size, the number of PhDs awarded, and publications at funded universities. Second, I find a patenting increase in commuting zones hosting funded universities, primarily attributable to incumbent firms located near those institutions, and driven by commuting zones with established R&D-intensive sectors. I find a larger effect in technology fields with high exposure to local universities' research. I provide evidence indicating two main mechanisms behind the patenting increase: greater reliance on scientific knowledge in patenting and the employment of local PhD graduates in industrial R&D.

*I thank Ran Abramitzky, Hoyt Bleakley, Nick Bloom, Thomas Chaney, Mihai Codreanu, Dylan Connor, José-Antonio Espín-Sánchez, Christian Fons-Rosen, Dominique Foray, Michela Giorcelli, Daniel Gross, Florencia Hnilo, Jennifer Hunt, Francesco Lissoni, Tamri Matiashvili, Nicholas Monath, David Mowery, Stephen Redding, Paola Sapienza, Carlo Schwarz, Francesca Truffa, Hans-Joachim Voth, Fabian Waldinger, Laura Weiwu, and Gavin Wright for insightful comments. All errors are my own.

†Stanford University, Department of Economics. gcristel@stanford.edu

1. Introduction

Universities are a crucial component of modern economies: they advance the knowledge frontier and foster human capital, both key elements for the development of new technologies (Jaffe 1989; Rosenberg and Nelson 1994; Mansfield 1995; Zucker et al. 1998; Ahmadpoor and Jones 2017). Previous studies have shown that investments in basic research conducted by academic scientists or individual laboratories can increase innovation in the private sector (Azoulay et al. 2019b; Bergeaud et al. 2022). Other works have documented that the establishment of modern research universities positively influenced innovation and economic development in their host locations (Dittmar and Meisenzahl 2022; Andrews 2023).

However, little is known about the effects of institutional funding to established universities on local innovation, particularly in the long term. While institutional grants involve larger investments compared to typical principal-investigator programs, they are less complex interventions than establishing new research institutions. Nevertheless, the impact of such funding programs on local innovation is ex-ante ambiguous, due to the substantial institutional and geographic variation in how universities influence technology development (Zucker et al. 2002; Bikard and Marx 2020; Lerner et al. 2024).

On the one hand, institutional grants may positively influence local innovation by enhancing universities' research capacity, thereby increasing the supply of scientific human capital and facilitating knowledge diffusion from universities to private sector R&D. On the other hand, even if institutional funds expand universities' research capacity, it may not necessarily stimulate local innovation. This could happen, for instance, if university outputs lack commercial impact, or if firms lack the capital to absorb the increased supply of scientists, networks to establish partnerships with academic scientists, or, more broadly, the absorptive capacity to benefit from university knowledge spillovers (Cohen and Levinthal 1989, 1990).

This paper studies the effects of institutional university funding on local innovation, focusing on one of the largest programs of this kind in U.S. history, the National Science Foundation's (NSF) Science Development Program (SDP). Introduced in 1965, the SDP awarded sizeable grants to natural science and engineering departments at 31 research universities across the U.S., aiming to expand the number of academic centers of research excellence in the country. Each institution received a grant supporting up to six departments, totaling between $30 and $90 million (in 2024 USD) and lasting five years, representing an expansion in the resources of a typical chemistry department between 20% and 50% (NSF 1977a).

My analysis is based on a large-scale data collection, combining information from scientific publications, PhD dissertations, and patents between 1960 and 1990. First, I track the publications histories of all scientists affiliated with U.S. research universities and companies. Second, I collect data on all PhD graduates from U.S. research universities. Third, I combine historical and modern patent records from the United States Patent and Trademark Office (USPTO) and disambiguate inventors listed in both sources. Fourth, relying on citation data from Marx and Fuegi (2022), I measure the links between patented inventions and scientific articles. Fifth, I link PhD graduates and the authors of scientific publications to inventor records, enabling me to assess their direct contribution to patenting and to observe publications co-authored by academic and industrial scientists.

Because the program aimed to increase the number of institutions conducting first-rate research, the NSF dismissed grant proposals from institutions already considered elite (NSF 1964; Page 1968; Lomask 1976). I exploit these excluded top-ranked universities as a comparison group in a difference-in-differences research design. To study university outcomes, I directly compare funded and excluded universities. For local innovation outcomes, I compare commuting zones hosting funded institutions with those hosting excluded top-ranked ones.

Top-ranked universities are a suitable comparison group for three main reasons. First, their exclusion from the program was based solely on their pre-existing elite status. Second, they were precluded from the SDP selection process and were thus not evaluated alongside institutions receiving funds. Third, historical evidence suggests they would have applied for—and likely received—SDP funding if permitted. An analogous rationale supports commuting zones hosting excluded top-ranked universities as a comparison group for those hosting the funded ones. In addition, the selection of universities for the SDP was based on evaluations of their scientific capabilities rather than anticipated future trajectories of industrial R&D within their regions, further mitigating concerns about selection bias between SDP-funded and top-ranked universities' commuting zones. Importantly, I provide evidence about the absence of differential trends between SDP-funded and top-ranked universities and between the commuting zones hosting them prior to the introduction of the program.

I start my analysis by focusing on university outcomes. I show that funded universities rapidly increased the number of scientists, PhD dissertations, and publications in the natural sciences and engineering after SDP grants were awarded, with no difference in the publications' quality. I find similar results comparing a narrower set of funded and top universities ranked around the cutoff that the NSF used to define excluded elite institutions; using a comparison group including all research universities in my sample; and testing a specification at the university–scientific domain–year level, where I can control for domain–by–year fixed effects. The results of two permutation tests indicate that my estimates are not driven by individual SDP-funded or excluded top-ranked universities and that they are unlikely due to random chance.

Reassuringly, I do not detect any difference between funded and comparison universities when I study the evolution of publications in the social sciences and humanities, disciplines not receiving any SDP fund. Overall, my estimates indicate that SDP grants increased the research capacity of funded universities.

The study of local innovation outcomes constitutes the main part of my investigation. My analysis is based on a panel dataset of commuting zone–technology field pairs, allowing me to account for year, commuting zone–by–technology field, and technology field–by–year fixed effects. I find an increase in patenting for commuting zones hosting SDP-funded universities relative to those hosting elite institutions. The effect becomes detectable approximately four years after the SDP grants and lasts for the subsequent ten years, indicating a patenting increase between 18% and 32% per year.

I test the robustness of my estimates in several ways. First, using newly released data from Gross and Sampat (2024), I control for federal funding to industrial R&D in each commuting zone-technology field-year cell. I find results highly comparable to the baseline, suggesting that the large investments in the private sector initiated by the U.S. government during WWII and sustained during the Cold War (Gross and Sampat 2023; Kantor and Whalley 2023) are not a confounding driver my estimates. Second, I introduce commuting zone–specific time trends, controlling for unobserved time-varying factors unique to each commuting zone and evolving systematically over time. I obtain estimates similar to the baseline, ruling out the possibility that inventive activities in commuting zones hosting SDP-funded universities were already on a virtuous trajectory prior to the SDP awards. Third, I test a specification using a control group of commuting zones constructed with Mahalanobis matching, with results once again comparable to those of my baseline specification. Lastly, I conduct two permutation tests, suggesting that my estimates are not driven by any particular commuting zone in the sample—either hosting SDP-funded or elite universities—and are unlikely driven by random chance.

I continue my analysis by studying the heterogeneity of these effects based on commuting zone, assignee, and technological characteristics. First, I find that the patenting increase is driven by commuting zones displaying above-median patenting per capita prior to the introduction of the SDP, as well as by commuting zones with an above-median share of patents citing the scientific literature. Second, I show that the entire effect is attributable to inventors located within five miles of the local university. Third, I find that the effect is due to incumbent firms, excluding any contribution from newly created companies or firms relocating to SDP-funded commuting zones. Fourth, I find that patenting increase is driven by the electrical & electronics and chemical fields, the largest patenting groups in SDP-funded commuting zones and those most closely linked to the disciplines receiving the bulk of SDP funds (physics and chemistry, NSF 1977a). I also find a decline in drugs & medical patents, suggesting a slight reallocation of inventive efforts away from a smaller field tied to disciplines that received less SDP support (biological sciences, NSF 1977a).

I deepen my analysis by testing how the patenting effect varies across technology fields based on their exposure to local university research. I do so by measuring the intellectual proximity between patents in a given commuting zone–technology field pair and the publications of the local university prior to the SDP—a procedure akin to Bergeaud et al. (2022) and Bergeaud and Guillouzouic (2024). Introducing this exposure measure as a continuous term in my difference-in-differences specification, I find that the positive effect of the SDP on local patenting is larger in technology fields with higher exposure to local university research.

Next, I investigate the mechanisms underlying the patenting increase following the Science Development Program. The results of my analysis at the university level indicate three main channels. First, the SDP may have increased the local supply of scientific human capital, with new PhD graduates directly contributing to private-sector R&D activities. Second, funded universities' larger departments may have increased local firms' opportunities to establish formal collaborations with academic scientists. Third, funded universities' expanded research capacity may have increased knowledge spillovers from universities to companies through informal channels, such as local conferences or higher chances of interactions between academic and industrial scientists.

Since all mechanisms may involve the diffusion of scientific knowledge from universities to private-sector R&D, I first test whether local patenting's reliance on science increased following the SDP introduction. Following Ahmadpoor and Jones (2017), I construct a citation network linking patents and scientific publications. I find that both patents directly citing scientific publications and patents close but indirectly connected to the scientific literature increase in commuting zones hosting an SDP-funded university. In contrast, I find a smaller and shorter-lived positive effect for patents more distant in the network or fully unconnected from a scientific publication. Importantly, I estimate an increase in the overall proportion of patents directly citing the scientific literature in funded commuting zones after the SDP awards.

I then assess the role of scientific human capital. I find a sizable increase in patents co-filed by PhD graduates from local universities and a smaller positive effect of the SDP introduction on patents not filed by any local PhD graduate. Nevertheless, the relatively small proportion of PhD graduates' patents makes it unlikely that their contribution was the only mechanism behind the patenting increase. To test the importance of formal university-industry collaborations, I track patents and publications co-authored by academic and industrial scientists. I find very few such patents and publications in my sample, and my estimates rule out any positive effect of the SDP on either group.

Overall, my estimates indicate that scientific knowledge diffusion from universities to private sector R&D and the increased availability of scientific human capital are the two drivers of the increase in patenting following the introduction of the SDP. While these results exclude any role played by formal university-industry collaborations, they do not exclude informal channels contributing to scientific knowledge spillovers.

Finally, I test whether the patenting increase in SDP-funded commuting zones came at the expense of nearby areas. Using a difference-in-differences framework akin to my main specification, I compare commuting zones lying within successive distance bands from those hosting an SDP-funded university to commuting zones within the same distance from areas hosting top-ranked institutions. Across all distance groups, I find no meaningful drop in patenting or evidence indicating inventor outflows.

The results of this study contribute to several strands of literature. First, they add to the literature on the relationship between academic research and industrial innovation (Jaffe 1989; Mansfield 1991, 1995; Rosenberg and Nelson 1994; Cockburn and Henderson 1996, 2001; Henderson et al. 1998; Zucker et al. 1998; Mowery and Sampat 2006; Furman and MacGarvie 2007; Foray and Lissoni 2010). Recent studies in this literature have focused on the effects of funding programs targeting principal investigators or individual research laboratories (Azoulay et al. 2019b; Bergeaud et al. 2022), declines in federal funding to individual academic scientists (Babina et al. 2023), shocks to university revenues (Tabakovic and Wollmann 2019), changes in university patenting legislation (Hvide and Jones 2018; Hausman 2022), and open access mandates on publicly funded research (Bryan and Ozcan 2021). This paper, instead, studies one of the largest and few institutional funding programs in U.S. history, investigating its effects on university research capacity and on local innovation.[1]

Second, this study contributes to the broader literature on the economic effects of universities (Cantoni and Yuchtman 2014; Kantor and Whalley 2014; Dittmar and Meisenzahl 2022;

[1] A broader and long-standing literature focuses on the relationship between basic science and innovation (Bush 1945; Maclaurin 1953; Nelson 1959, 1962; Rosenberg 1982; Kline and Rosenberg 1986; Mowery 1997; Stokes 1997; Mokyr 2002; Ahmadpoor and Jones 2017; Poege et al. 2019). Closely related research streams investigate university patenting and technology licensing, particularly in relation to the Bayh-Dole Act introduction in 1980 (e.g., Mowery et al. 2001, 2002, 2004; Jensen and Thursby 2001; Agrawal and Henderson 2002; Thursby and Thursby 2002; Sampat 2006; Lissoni et al. 2008; Azoulay et al. 2009; Lissoni 2010), the dynamics of private sector's investments in basic science research (e.g., Cohen and Levinthal 1989, 1990; Zucker et al. 2002; Arora et al. 2021a,b), and the role of public- and private-sector research in the U.S. innovation system (e.g., Arora et al. 2019, 2020; Fleming et al. 2019).

Andrews 2021b, 2023; Andrews and Smith 2023; Russell et al. 2024; Russell and Andrews 2024). Prior research has mostly focused on the effects of establishing new universities. One exception is Kantor and Whalley (2014), who study the local productivity effects of changes in university expenditures driven by stock market returns shocks. My paper adds to this literature by studying a policy aimed at increasing the research capacity of already established institutions, providing novel evidence on its effects on private sector's innovation and the mechanisms linking them to university funding.

Third, this paper contributes to the literature on the economic effects of R&D funding shocks during and after World War II. Previous studies have focused on the effects of U.S. federal funding to applied R&D performed by both academic and private sector organizations (Gross and Sampat 2023, 2025; Gross and Roche 2023), U.S. federal funds targeting mostly industrial contractors (Kantor and Whalley 2023), or mission-oriented programs in the Soviet Union (Schweiger et al. 2022). This paper adds to this literature by studying a funding initiative directed exclusively at universities and investigating its effects on university research capacity and its spillovers on local private sector innovation.[2]

The remainder of the paper is organized as follows. Section 2 provides a historical overview of the Science Development Program. Section 3 describes my data sources and my disambiguation and record linkage procedures. Section 4 describes my main empirical strategy. Section 5 presents the results on university outcomes. Section 6 presents the results on local innovation outcomes. Section 7 concludes.

2. Historical Context: The NSF Science Development Program

In March 1964, the National Science Foundation (NSF) announced the Science Development Program, also known as the "Centers of Excellence" initiative. The program aimed to strengthen the research capacity of non-elite universities, particularly in regions that lacked an institution regarded as top-tier (NSF 1964). Running between 1965 and 1971, the Science Development Program allocated approximately $179 million (equivalent to $1.78 billion in 2024 USD) through institutional grants to 31 universities, almost exclusively to their departments in the natural sciences and engineering (NSF 1977a).

The program emerged from a public debate on the need to increase the number of top research universities in the U.S., both for the country's welfare and to sustain the Cold War science and technology race, made particularly salient by the Soviet Union's 1957 launch of the first artificial Earth satellite Sputnik (Geiger 1997). Such debate culminated in a 1960 report

[2] A related body of work investigates the effects of management training and technology transfer programs introduced during and after World War II on firm productivity (Giorcelli 2019; Bianchi and Giorcelli 2022; Giorcelli and Li 2024; Giorcelli 2024a,b).

from the President's Science Advisory Committee, chaired by Glenn T. Seaborg, Nobel Prize-winning chemist and UC Berkeley's Chancellor. The "Seaborg Report," as it became known, called for an expansion of "first-rate academic centers" from "fifteen or twenty today [to] thirty or forty in another fifteen years", arguing that "[e]xisting strong institutions cannot fully meet the nation's future needs" and that "support to the rising centers will be repaid many times over in service to society" (President's Science Advisory Committee 1960, pp. 14-15).

The Science Development Program was also influenced by a closely related debate on the geographic and institutional distribution of federal research funds. Several members of Congress and the Kennedy's and Johnson's Administration viewed the allocation of federal research funds as too concentrated in few locations and few universities. That contrasted with the view held by most members of the U.S. scientific elite, who believed that research excellence should be the main criterion to allocate research funds, regardless of geographic or institutional equity considerations (Graham and Diamond 1997).[3]

The Science Development Program addressed both the Seaborg Report's plea and the government's pressure to provide more support to academic research outside the geographic or institutional elite. In 1963, the NSF approached Congress to ask for the necessary financial appropriation for the Science Development Program. The program was initially turned down and funds for its first three years of activity were approved by Congress only in January 1964, after a period of uncertainty (Lomask 1976).

The NSF opened the Science Development Program to any higher education institution with departments in science or engineering, except those recognized as part of the elite of the country. In its March 1964 announcement, the NSF stated that "[s]ince the goal is to increase the number of strong academic centers in science, institutions already recognized as being outstanding in science should continue to depend on existing programs for assistance" (NSF 1964, p. 4). Howard E. Page, Head of the NSF Office of Institutional Programs in the 1960s, noted that "[t]he program did exclude proposals from the unnamed and unnamable institutions in the magic circle of the top twenty" (Page 1968, p. 115). According to Lomask (1976, p. 133) "it was understood that none of the "fifteen or twenty" top-rated institutions need apply. When one of them [Caltech] did, its president received a polite note from Director Haworth,

[3] For instance, based on a series of hearings held in 1963, the House Committee on Science generated a report which critically noted the concentration of federal research funds in institutions from the Northeast and the Pacific Coast. In 1963, President Kennedy stated that there should be an outstanding university in every major region of the country (Page 1968). In a 1965 exchange with his Cabinet, President Johnson complained that research funds were "still concentrated in too few institutions in too few areas of the country." On the other hand, the NSF director Leland J. Haworth defended the allocation of his agency's research funds, arguing that the government should not turn to institutions "which would first have to build up a capability." The chief scientific advisor to both President Kennedy and President Johnson, Donald F. Horning, stated that "the first criterion for funding an R&D program by Government agencies is the excellence of the institution." The Seaborg Report itself warned against the allocation of research funds to institutions deemed as not qualified (Lomask 1976).

reminding him that the development grants were strictly for the second-stringers."

The NSF hosted an application round in each year from 1964 to 1968 (NSF 1977a). Universities were evaluated based on a proposal detailing their plans to use NSF funds and after an on-site visit from the evaluation body, which consisted of both NSF staff and external scientists and science administrators. The evaluation of universities' proposals and the reports from on-site visits contributed to a recommendation report submitted by the evaluation body to a panel of experts advising the NSF and to the NSF leadership, which ultimately decided whether a grant could be awarded (Drew 1975; NSF 1977a).

Table 1 lists the universities that received a Science Development Program grant. Of the 31 institutions, 20 are located in Southern or Midwestern states, including flagship universities such as the University of Texas at Austin or the University of Florida, public ones such as the University of Virginia or Michigan State University, and private ones such as Duke University or Washington University in St. Louis. The first grant was awarded to the University of Oregon in 1965, while the last one to the University of Pittsburgh in 1969.

Each institution received a grant supporting up to six departments, for a total amount between \$3 and \$9 million (equivalent to \$30 and \$90 million in 2024 USD) and lasting for five years. Although matching funds were not formally required, receiving institutions were expected to sustain the increased financial resources for each department after the program ended (Drew 1975; NSF 1977a). The grants were allocated for hiring new faculty members, enlarging PhD programs, acquiring new research equipment, and improving or expanding research facilities (Drew 1975; NSF 1977b).

The Science Development Program grants became some of the biggest sources of university research funding in the 1960s, especially for the physical sciences. The NSF estimated that for a typical chemistry department of the period, a Science Development Program grant would represent an increase of research resources of about 20% to 50% (NSF 1977a). The last column of Table 1 benchmarks SDP awards against each university's total federal research funding of both 1963 and 1964, revealing they averaged roughly 29% of those resources. These grants were larger than similar contemporary institutional grants awarded by agencies such as the Department of Defense (DoD) and the National Institutes of Health (NIH), larger than most institutional grants awarded by the National Aeronautics and Space Administration (NASA), and surpassed only by the biggest institutional grants from the Ford Foundation (NSF 1977a).[4]

In 1966, the NSF introduced two smaller subprograms extending institutional funding to

[4]Nevertheless, it is important to note that the NSF, DoD, NASA, and NIH sponsored academic research through various means beyond institutional programs during the 1960s. Considering all funding sources, the NSF was the third-largest supporter of university research in the mid-1960s, following the NIH and the DoD, and ahead of the Atomic Energy Commission (AEC) and NASA (Geiger 1993).

a broader set of universities: the Special Science Development Program and the Departmental Science Development Program. The first funded up to two departments at institutions which applied to the Science Development Program, but were rejected because judged to lack the strength to maintain a larger program, and awarded grants of approximately $1 million (equivalent to roughly $9.6 million in 2024 USD) to 11 universities. The second was an even more focused program, typically supporting only one department, and awarded grants averaging $600,000 (equivalent to $5.8 million in 2024 USD) to 62 institutions (NSF 1977a).

Funding restrictions prevented the NSF to accept new applications after 1969, and the program was officially terminated in 1971 within broader budgetary cuts introduced by the Nixon Administration.

In the mid 1970s, the NSF performed two main evaluations of the program. The first was carried out by the National Board of Graduate Education and involved mostly a qualitative analysis based on interviews and on-site visits (Drew 1975). The second was based on reports the NSF requested to each funded university's president and faculty deans (NSF 1977b). According to these accounts, the Science Development Program enabled funded departments to increase faculty hiring, to enlarge the size of PhD programs, to improve research facilities and experimental equipment, to expand their libraries' collections, and to increase the likelihood to obtain future external funds due to departments' improvement.

In summary, the Science Development Program substantially increased the research funds of numerous universities across the United States, excluding top-ranked institutions—even though historical evidence suggests that elite universities would have applied for and likely received grants if they had been eligible.

3. Data

3.1. Sources

My analysis is based on five main data sources: scientific publications, PhD dissertations, patent documents, citations between patents, and citations from patents to scientific publications.

I obtain scientific publication data from the OpenAlex database (Priem et al. 2022), the successor of the Microsoft Academic Graph database. I extract bibliographic information for all articles published between 1960 and 1990 and listing at least one author affiliated with a research university surveyed and ranked by Cartter (1966), roughly corresponding to institutions ranked as "R1: Doctoral Universities" by the modern *Carnegie Classification of Institutions of Higher Education.* For the same time period, I also extract publications from scientists affiliated with firms based in the U.S. The resulting dataset includes the authors' full name and affiliation, scientific field, and citations received, for 2,008,779 publications.[5]

I track PhD dissertations using the ProQuest Dissertations & Theses Global database, a collection of dissertations published since 1861 and the official dissertations repository for the Library of Congress. I focus on graduates from the group of research universities evaluated by Cartter (1966), extracting all dissertations published between 1960 and 1990, including the author's full name, scientific field, and title of 619,862 dissertations.

I measure U.S. innovation activities using patent documents. Despite their recognized limitations ("not all inventions are patentable, not all inventions are patented, and the inventions that are patented differ greatly [...] in the magnitude of the inventive output associated with them," Griliches 1990, p. 1669), patents are a key measure of innovation in advanced economies since at least the early twentieth century (Mansfield 1986; Cohen et al. 2000; Moser 2016). For patents granted before 1975 ("historical" patents, Andrews 2021a), I combine data from the Patstat database (inventor and assignee name), the PatentCity database (inventor and assignee location; Bergeaud and Verluise 2024), the USPTO historical masterfile (patents' grant date and technology class; Marco et al. 2015), and Google Patents (patents' filing date). For patents granted since 1975 ("modern" patents), I obtain the same information from the USPTO PatentsView database (USPTO 2024). I focus on all patents filed between 1960 and 1990 and listing at least one inventor with a U.S. address, obtaining a dataset of 1,102,861 patents.

Lastly, I track links between patents using citations extracted from the patent text by Verluise et al. (2020) and I measure connections between patented technologies and knowledge embodied in scientific articles using the database on patent citations to publications constructed by Marx and Fuegi (2022).

3.2. Author and Inventor Disambiguation

OpenAlex assigns unique identifiers to the same author across different publications using a disambiguation algorithm. Regarding inventors, only those listed on modern patents report unique identifiers, which are assigned by the USPTO using a supervised classification algorithm combined with hierarchical agglomerative clustering (Monath et al. 2021).[6]

I apply the same machine learning technique to disambiguate also inventors listed on historical patents, assigning unique identifiers to all inventors filing a patent at the USPTO and residing in the U.S. between 1920 and 2015. A detailed description of my algorithm can be found in Appendix B. In summary, I first train a logistic classifier on 2 million disambiguated inventor-patent instances from modern patents, followed by fine-tuning on a development set

[5] Appendix Table A1 lists the research universities surveyed by Cartter (1966) and included in my sample.

[6] More information about OpenAlex's disambiguation algorithm can be found at following link: `https://github.com/ourresearch/openalex-name-disambiguation/tree/main/V3` (last access: December 2024).

of additional 100,000 disambiguated records. The features include the inventors' name, residential location, co-inventors' and assignees' identity, and technology classes. Next, I use the classifier's predictions to create a distance matrix for each group of inventors sharing the same last name and first name two initials. Finally, I apply hierarchical agglomerative clustering to these matrices, assigning the same identifier to inventors whose distance scores fall within the optimal threshold maximizing precision and recall determined during fine-tuning.[7]

I evaluate the performance of the algorithm in three steps. First, the optimal threshold set during fine-tuning achieves precision and recall of approximately 0.99 and 0.95, respectively. Second, focusing only on inventors from modern patents, my identifiers match the USPTO's with 96% and 99.5% similarity, indicating a nearly identical disambiguation. Third, following Akcigit et al. (2022), I search for the top 50 most prolific inventors in my dataset in a crowdsourced list of the most prolific inventors known, maintained on Wikipedia. I find 45 of these inventors in the list and verify the identities of the remaining five through their biographical profiles on company or university websites. Additionally, for inventors with careers fully covered in both my dataset and Wikipedia, I observe minimal differences in their total patent stock.

3.3. Record Linkage: PhD graduates, Scientists, and Inventors

I link the authors of doctoral dissertations and of scientific publications to inventor records, enabling me to study the contribution of PhD graduates and of academic scientists to patenting, and to analyze collaborations on scientific projects between academic and industrial scientists. I adopt the class of ABE algorithms (Abramitzky et al. 2012, 2014, 2021), building on their linkage routine to leverage the information provided by doctoral dissertations, scientific publications, and patents.

I link PhD graduates' records from ProQuest to inventor records from USPTO patents based on the following routine. First, I create a set of candidate links by matching PhD graduates and inventor records on first name, last name, and middle name initials. Second, I keep only records with a difference between the PhD graduation and first patent within a $[-5; +30]$ years interval. Third, I discard records where the difference between the PhD graduation and the last patent is beyond 40 years. Lastly, I retain only unique PhD graduate-inventor pairs (that is, I exclude PhD graduate records with multiple inventor candidates or multiple PhD graduates linked to the same inventor). I link around 15% of PhD graduates to an inventor.[8]

[7] This disambiguation strategy is similar to the method used by Akcigit et al. (2022), who disambiguated inventors from historical patents using a training set of inventors from modern patents disambiguated by Li et al. (2014).

[8] The lower limit of -5 years accounts for inventions that may have been patented during the PhD program. The upper limit of $+30$ years is based on data from Kaltenberg et al. (2023), which show that most inventors file their first patent by age 55.

I follow a similar procedure to link authors from scientific publications to inventors. First, I create a set of candidate links by matching scientific authors and inventor records on first name, last name, and middle name initials. Second, I keep only records with a difference between the first publication and first patent within a $[-5; +30]$ years interval. Third, I discard records where the difference between the first publication and the last patent is beyond 40 years. Lastly, I retain only unique author-inventor pairs. I link around 5% of scientific authors to an inventor.

4. Empirical Strategy

The main empirical strategy of this paper is based on the exclusion of top-ranked universities from the Science Development Program. Since the program aimed to expand the number of research centers of excellence across the country, the NSF dismissed grant proposals from institutions already considered elite. I exploit these excluded top-ranked universities and their local economies as a comparison group in a difference-in-differences research design. Specifically, I compare funded and excluded universities to examine university-level outcomes and compare commuting zones hosting funded institutions versus those hosting excluded top-ranked ones to study local innovation outcomes.[9]

Excluded top-ranked universities are a suitable comparison group for three main reasons. First, their exclusion was based solely on their pre-existing elite status, not on any anticipated differences in future research trajectories compared to funded universities. Second, these universities were excluded from the SDP selection process and were not evaluated alongside institutions that ultimately received SDP funds, unlike universities that applied but were not selected. Third, as detailed in Section 2, top-ranked institutions would have applied to the program if allowed—as evidenced by inquiries from their presidents to the NSF—and would likely have received SDP funds had their applications been considered.

An analogous rationale supports using commuting zones hosting excluded top-ranked universities as a comparison group for those hosting SDP-funded institutions. First, the selection of universities for the SDP was based solely on evaluations of their scientific capabilities, not on pre-existing trends or anticipated future trajectories of industrial R&D or other economic outcomes within their commuting zones. Second, even if commuting zones' economic trajectories might have been correlated with the research performance of their universities, the exclusion of top-ranked universities—and thus their commuting zones—from the SDP selection process mitigates concerns about selection bias favoring regions with SDP-funded institutions.

Below, I detail how I identify the group of excluded top-ranked universities and I test

[9] Commuting zones are clusters of counties exhibiting strong commuting ties and approximate the local economies hosting each university (Tolbert and Sizer 1996; Autor and Dorn 2013).

the validity of this research design by investigating the trends of SDP-funded and top-ranked universities and their commuting zones prior to the Science Development Program.

4.1. Identifying Top-Ranked Universities and Commuting Zones

While the exclusion of top-ranked universities was made public and communicated by the NSF to those institutions' administrators, a formal list with excluded universities was never compiled. As described in Section 2, the NSF excluded the top-ranked institutions based on their perceived research excellence, a group generally referred to as the "top fifteen" or "top twenty" universities in the country (President's Science Advisory Committee 1960; Page 1968; Lomask 1976). According to Howard E. Page, Head of the NSF Office of Institutional Programs in the 1960s, "[n]o list of institutions disqualified because of their distinction was ever prepared. One way of preparing such a list would be to take the twenty institutions presently receiving the largest amount of federal funds for science" (Page 1968, p. 115).

I follow a similar data-guided procedure to identify top-ranked universities excluded from the NSF Science Development Program. First, I identify the top-ranked universities based on a measure of their perceived research excellence in the early 1960s. I rely on the evaluation of U.S. research universities by discipline produced by Cartter (1966), based on an extensive survey of senior and junior U.S. scholars administered in 1964, one year prior to the SDP introduction. Using Cartter's (1966) scores, I rank universities in the biological sciences, physical sciences, and engineering (Appendix Table A7, Table A8, and Table A9). Then, I retain all universities ranked in the top twenty in at least one of those domains and categorized as "distinguished" or "strong" across all three domains—the highest-rated categories by Cartter (1966) and the only ones receiving a score. Second, I validate this group by comparing it with the top twenty research universities by total federal research funds received in fiscal year 1964. I use two rankings, one produced by the National Science Foundation (NSF 1967) and one from the U.S. General Accounting Office (Comptroller General of the U.S. 1976).[10,11]

Table 2 reports my results. In panels B and C, I list universities highly-ranked in natural science and engineering disciplines, but not considered for my comparison group because funded by either the Science Development Program or by one of its subprograms. Panel A shows the 18 top-ranked universities constituting my comparison group. All of them are part of the

[10] Appendix Figure A1 lists the scientific disciplines ranked by Cartter (1966) grouped by domain. Appendix Figure A2 provides examples of the rankings and scores found in Cartter (1966).

[11] Although the University of Chicago is not ranked in any engineering domain, I include it in the group of top-ranked universities for two key reasons. First, it lacked an engineering school in the 1960s and therefore could not be evaluated in that domain. Second, and more importantly, it was explicitly mentioned by Page (1968) as an elite institution excluded from the SDP, as follows: "For example, a proposal to establish a doctoral program in anthropology at the Massachusetts Institute of Technology or an engineering school at the University of Chicago would not be entertained, on the grounds that these institutions could request assistance from other public and private sources" (Page 1968, p. 115).

top twenty institutions in terms of total federal research funds received in the 1964 based on the NSF (1967) ranking, except three universities: Brown University, the California Institute of Technology, and Princeton University. The latter is reported in the top twenty by the U.S. General Accounting Office's ranking; the California Institute of Technology is directly mentioned as one of the excluded institutions from the SDP in Lomask (1976); Brown University is ranked in the top fifteen in one discipline group by Cartter (1966).

Table 3 lists the commuting zones in my sample, distinguishing between those hosting SDP-funded universities and those hosting excluded top-ranked universities. Most commuting zones host only one institution, with the exceptions of the areas of Boston, Pittsburgh, and of Chapel Hill-Durham-Raleigh. I do not consider the commuting zones of Los Angeles, Newark, and New York City because they host both SDP-funded and top-ranked universities.

In both the analyses of university and commuting zone outcomes, I test the sensitivity of my results to the exclusion of each top-ranked university and each commuting zones from the sample.

4.2. Trends Prior to the Science Development Program

The validity of my difference-in-differences research design hinges on the assumption that, in the absence of the Science Development Program, the outcomes of SDP-funded universities and excluded top-ranked universities—as well as their respective commuting zones—would have followed similar trajectories. Although this parallel trends assumption cannot be directly tested, it appears reasonable if SDP-funded and excluded units do not exhibit differential trends prior to the program's introduction.

To assess the plausibility of this assumption for universities, I examine pre-SDP trends by regressing each university outcome on an interaction between an indicator for SDP-funded universities and year dummies from 1960 to 1964, using 1964 as the baseline year. I include year and university fixed effects, and cluster standard errors at the university level. I analyze four outcomes within the natural sciences and engineering: the number of new scientists, the total number of scientists, the number of PhD dissertations, and the number of publications. Panel (a) of Figure 1 presents the results. For each dependent variable, the coefficients are statistically insignificant, often estimated very close to zero, and exhibit flat trends.

I conduct an equivalent test for commuting zones. I regress each patenting outcome on an interaction between an indicator for commuting zones hosting SDP-funded universities and year dummies from 1960 to 1964, again using 1964 as the baseline year. I include year and commuting zone fixed effects and cluster standard errors at the commuting zone level. I examine the number of patents filed in each commuting zone and year for four categories: all patents, patents filed by firms, patents citing the scientific literature, and patents listing a PhD graduate as an inventor. Panel (b) of Figure 1 presents the results. For each patenting outcome, the estimated coefficients are close to zero and display a flat trend. The estimates for patents by PhD graduates are less precise due to the relatively sparse nature of this outcome compared to other patent counts.

I further inspect commuting zones trends prior to the Science Development Program using information on total employment and the number of establishments between 1951 and 1964, based on County Business Patterns data digitized by Eckert et al. (2022). Appendix Figure D1 shows that, both for employment and establishment, the coefficients are all statistically indistinguishable from zero and display a flat trend.

Overall, these findings exclude differential trends between SDP-funded and top-ranked universities, as well as their commuting zones, prior to the initiation of the Science Development Program. This evidence supports excluded universities and their commuting zones as an appropriate comparison group in my difference-in-differences research design.

5. The Science Development Program and University Outcomes

To study the effects of the NSF Science Development Program on universities, I construct a panel dataset of university outcomes observed yearly between 1960 and 1990. I adopt a dynamic difference-in-differences approach, comparing universities which received SDP funds to the comparison group of top-ranked universities excluded from the program. Formally, I estimate the following equation:

$$\mathbb{E}[y_{ut}|X_{ut}] = exp(\alpha + \sum_{\substack{\tau=-5 \\ \tau\neq-1}}^{20} \beta_\tau \cdot I_\tau \times SDP_u + \phi_t + \gamma_u + \xi_u \cdot t + \delta X_{ut}) \tag{1}$$

where y_{ut} denotes an outcome for university u in year t; ϕ_t are year fixed effects, accounting for time-variant shocks common to all universities; γ_u are university fixed effects, capturing time-invariant characteristics of each university; and $\xi_u \cdot t$ are university-specific time trends. X_{ut} is a vector of controls for other institutional funding programs varying by university and year. I control for each of the major institutional funding programs contemporary to the SDP, from NASA, the DoD, the NIH, and the Ford Foundation. For each funded university, I allocate the total grant amount evenly across the years corresponding to the duration of the respective program. To account for longer-term effects, I calculate a linearly decaying amount starting from the first year after the program ends, extending it for an additional period equal to the program's original duration.[12]

[12]I obtain identical results by controlling only for the years when the program was running. Appendix Table A3, Table A4, Table A5, and Table A6 list the universities funded by each program and specify the grant amounts

SDP_u is an indicator equal to 1 for universities which received an SDP grant, while I_τ is an indicator equal to 1 in period τ. For each SDP-funded university, I set $\tau = 0$ to the calendar year when they received the SDP grant, which ranges between 1965 and 1969. For control universities, I set $\tau = 0$ in 1965. I consider all periods between $\tau \geq -5$ and $\tau \leq 20$, with $\tau = -1$ as reference period. Since all my outcomes are count variables, I follow other econometric studies of innovation and science (e.g., Henderson and Cockburn 1994, Blundell et al. 1995, Azoulay et al. 2019a, Catalini et al. 2020) and produce pseudo-maximum-likelihood (PML) estimates based on Hausman et al.'s (1984) Poisson fixed effects model. I cluster standard errors at the university level.

The main identifying assumption for this difference-in-differences model is the parallel evolution of outcomes for SDP-funded and top-ranked universities had the Science Development Program not been introduced. Section 4.2 provides evidence in support of this assumption. I address additional concerns about the selection of SDP-funded universities by including university-specific time trends in my specification, capturing differential trends across universities over time.

5.1. Results

Figure 2 reports the estimated β_τ from Equation 1 for four dependent variables. In panel (a), y_{ut} equals the number of new scientists working in a natural science or engineering field and affiliated with university u and publishing in year t. I define new scientists as authors never observed publishing with an affiliation to university u prior to year t. The estimates for $\tau < 0$ are all close to zero and not statistically significant, displaying no particular trend. For $\tau > 0$, the coefficients progressively gain statistical significance increasing in size until the ninth year after the SDP grants were awarded, after which they decrease, reaching values around zero for the final years in the sample. In panel (b), I report estimates for the total number of scientists working in the natural sciences and engineering, finding results highly comparable to panel (a).

The positive and statistically significant coefficients in the post-SDP period in panels (a) and (b) correspond to an increase in natural science and engineering scientists approximately between 5% and 38%. This range reflects the dynamic nature of the SDP effects, with smaller effects observed immediately after the grants were awarded and larger ones occurring during the mid-period of my study. For the average SDP-funded university in $\tau = -1$, this corresponds to approximately 7 to 53 additional scientists.[13]

awarded.

[13]The estimated Poisson pseudo-maximum-likelihood coefficients can be interpreted as log-relative changes in the outcome variable. To express these effects as percentage increases, I exponentiate the coefficients and subtract one. For example, the estimate for period $\tau = 9$ in panel (a) indicates an increase in new scientists by approximately $[exp(0.319) - 1] \cdot 100 = 38\%$.

In panel (c), the dependent variable equals the number of PhD dissertations in the natural sciences or engineering published by university u in year t. The estimates are similar to those in panel (a) and (b), albeit with slightly larger confidence intervals. The coefficients for the period preceding an SDP grant award display a moderately flat trend and the estimates are consistent with a positive effect of the program peaking around seven years after the grant and progressively decreasing to zero for the last years in my sample. In panel (d), I set y_{ut} equal to the number of publications in the natural sciences and engineering from university u and year t, again finding results close to those in panels (a), (b), and (c).

In Figure 3, I present the results of several robustness checks, using scientific publications as my dependent variable. In a panel (a), I compare institutions clustered around the federal funding top twenty "cutoff." Based on the 1964 ranking produced by the NSF (1967), I restrict the sample of SDP-funded universities to the first five ranked immediately below the top twenty (notably, all ranked 21 through 25) and the group of comparison universities to the first five ranked immediately above it. The estimates for $\tau < 0$ exhibit no systematic trend and are statistically insignificant. For $\tau > 0$, the point estimates mirror the baseline pattern, though with wider confidence intervals, showing a fast increase in publications within the first ten years after an SDP grant award, progressively fading to zero during the final years of my sample. I find comparable estimates when including all research universities in my sample in the control group (panel (b)).[14]

In panel (c), I report the results of a placebo test, using as a dependent variable the number of publications in the social sciences and humanities, all fields not funded by Science Development Program. Unless the award of an SDP grant triggered substantial re-allocation of pre-existing resources within funded universities, the publication outcomes in these fields should have not been influenced by the program. Reassuringly, the estimated coefficients are all close to zero, displaying a flat trend across the entire time frame.

In panel (d), I report the estimates from a difference-in-differences model equivalent to Equation 1, where I count publications in university, scientific domain, and year cells. This enables me to account for scientific domain-by-year fixed effects, absorbing unobserved confounders varying by domain and year across universities. The estimates are almost identical to my baseline results in terms of statistical significance, magnitude, and temporal dynamics.

In panel (e), I test the sensitivity of my results to individual SDP-funded universities and to the definition of the comparison group of top-ranked universities, by re-estimating Equation

[14] The SDP-funded universities in the restricted sample are Maryland (21), Rochester (22), Pittsburgh (23), Colorado Boulder (24), and Washington University in St. Louis (25). The comparison universities are Johns Hopkins (16), Pennsylvania (17), Yale (18), and Ohio State (19). Because Duke (20) and the two institutions ranked 14th and 15th also received SDP grants, I add Cornell (13) to balance the comparison group at five units.

1 and iteratively excluding each university in my sample. The estimates are always comparable to my baseline results, ruling out the possibility that any single university is the driver of my results.

Panel (f) reports the results of an additional permutation test. Following Abadie et al. (2010), I randomly assign a placebo SDP-funded status and timing (between 1965 and 1969) to all universities in my sample which were not funded by the program. I then estimate a two-period differences-in-differences version of Equation 1, repeating this process 1,000 times. The coefficient from my true specification can be considered significant if it is extreme relative to the distribution of placebo estimates. The plot in panel (f) shows that the true estimate, denoted by a vertical blue line, lies in the right tail of the placebo estimates' distribution, suggesting that it is unlikely driven by random chance.

I conclude this section by studying any change in the quality of scientific publications in the natural sciences and engineering after the SDP grants were awarded. I consider two publication quality measures: the number of citation-weighted publications and the average citations received by publications from university u and year t. Panel (e) in Figure 2 shows estimates similar to those for publications not weighted for quality, while panel (f) shows estimates close to zero, jointly excluding any positive or negative change in SDP-universities publications after the awards.

Taken together, the results in this section indicate that universities funded by the NSF Science Development Program expanded their natural science and engineering faculties and PhD programs, leading to an increase in publications within these fields. These effects lasted approximately for the first fifteen years following the SDP grant awards.

6. The Science Development Program and Local Innovation

To investigate the effects of the Science Development Program on local innovation, I construct a panel dataset of commuting zone-technology field pairs, observed yearly between 1960 and 1990. I adopt the technology field categorization of Hall et al. (2001), which groups USPTO technology classes into 37 broader fields.[15] I employ a dynamic difference-in-differences specification, comparing commuting zone-technology field pairs hosting a university that received SDP funds to the comparison group of pairs hosting top-ranked universities excluded from the program. Formally, I estimate the following equation:

[15] Hall et al.'s (2001) categorization is relatively narrow. For example, the broader category of electrical and electronic engineering includes the following fields: electrical devices; electrical lightning; measuring & testing; nuclear & X-rays; power systems; semiconductor devices; and miscellaneous-elec.

$$\mathbb{E}[y_{ict}|X_{ict}] = exp(\alpha + \sum_{\substack{\tau=-5 \\ \tau\neq-1}}^{20} \beta_\tau \cdot I_\tau \times SDP_c + \phi_t + \gamma_{ic} + \lambda_{it} + \delta X_{ct}) \tag{2}$$

where y_{ict} is an outcome for technology domain i, commuting zone c, and year t; ϕ_t are year fixed effects, accounting for time-variant shocks common to all commuting zone-technology field pairs; γ_{ic} are commuting zone-by-technology field fixed effects, capturing time-invariant characteristics of each commuting zone-technology field pair; λ_{it} are technology field-by-year fixed effects, absorbing time-varying shocks specific to each technology field.

X_{ct} is a vector of controls for R&D funding programs varying by commuting zone and year. I control for each of the major university funding programs contemporary to the SDP, introduced by NASA, the DoD, the NIH, and the Ford Foundation, with the same variables defined for Equation 1.

SDP_c is an indicator equal to 1 for commuting zones hosting a university which received an SDP award, while I_τ is an indicator equal to 1 in period τ. For each commuting zone hosting a SDP-funded university, I set $\tau = 0$ to the calendar year when the SDP grant was received, while I set $\tau = 0$ in 1965 for commuting zones in the comparison group. I consider all periods between $\tau \geq -5$ and $\tau \leq 20$, with $\tau = -1$ as reference period. I still rely on Hausman et al.'s (1984) Poisson fixed effects model due to the count nature of most of my dependent variables. I cluster standard errors at the commuting zone level.

Akin to Equation 1, the main identifying assumption for this specification is the parallel evolution of outcomes for commuting zones where SDP-funded universities were located and for those hosting top-ranked universities, had the Science Development Program not been introduced. The absence of differential trends between the two commuting zone groups prior to the Science Development Program across patenting and other economic outcomes reported in section 4.2 provides evidence in support of this assumption.

6.1. Main Results

Figure 4 shows the estimated β_τ from Equation 2, where the dependent variable equals the number of patents filed in technology field i, commuting zone c, and year t. For the period preceding the award of SDP grants, all estimated coefficients are close to zero, statistically insignificant, and display a flat trend. The coefficients remain statistically insignificant for the first four years after the SPD grants awards, progressively increasing and gaining statistical significance, displaying similar point estimates until the 14th year post-SDP. For the remaining years, the coefficients remain positive, although they slightly decrease and lose statistical significance.

These results indicate that commuting zones hosting universities which received an SDP grant increased their patenting output relative to the comparison group after the higher education institutions received the grants. The effect appears after around four years and lasts for the following ten years. The positive and statistically significant coefficients throughout that period indicate that patenting in SDP-funded commuting zone increased by 18% to 32% on average. For the mean commuting zone-technology field pair hosting SDP-funded universities, this amounts to an increase of 1.3 to 2.3 additional patents per year.[16,17]

I test the robustness of this baseline finding in several ways. First, I control for additional potential confounding factors varying by commuting zone and year. The period under examination witnessed substantial investments by the U.S. federal government in industrial R&D, starting during WWII and continuing throughout the Cold War, with salient shocks such as the Soviet Union's launch of Sputnik (Mowery and Rosenberg 1991; Geiger 1997; Mowery 2010; Kantor and Whalley 2023; Gross and Sampat 2023). A legitimate concern is that the estimates in Figure 4 may be biased by federal funding to industrial R&D flowing into commuting zones when SDP grants were awarded.

I address this concern by re-estimating Equation 2 and adding as a further control the share of patents funded by the Department of Defense, the Department of Energy, the Department of Health and Human Services, and NASA in each technology field, commuting zone, and year cell. I leverage newly-released data by Gross and Sampat (2024) on the universe of patents funded by government agencies to construct a measure of federal funding to industrial R&D akin to that employed by Gross and Sampat (2023). Conditional on patenting, the average proportion of federally-funded patents in each technology field, commuting zone, and year cell in my sample is 6%, resulting from a skewed distribution with over two thirds of those observations not linked to any federal funding, and cells in the top quintile displaying proportions over 20% (Appendix Figure C1).[18]

Panel (a) of Figure 5 shows that the baseline results are virtually unchanged with the addition of this control. In panel (b), I re-estimate Equation 2 and directly exclude federally-funded patents from my dependent variable. This constitutes a rather restrictive robustness check, as any innovation effect of the Science Development Program may have interacted with

[16] I also estimate a two-period difference-in-differences specification, substituting the set of I_τ in Equation 2 with a single indicator $Post_\tau$, equal to 1 from period $\tau = 0$ onwards. The estimate is reported in column 1 of Table 4 and indicates a patenting increase of about 13% following the SDP awards.

[17] Appendix Figure D2 presents alternative estimates of Equation 2 that use citation-weighted patent counts as the dependent variable; the results closely mirror those in Figure 4.

[18] Gross and Sampat (2024) show that the DoE, the HHS, NASA, and in particular the DoD, accounted for almost all federally-funded patents between 1920 and 2015. While it would be preferable to control for federal funding to R&D using the actual funds rather than based on patenting output, to the best of my knowledge, granular data at the commuting zone-year level and covering the entire United States during this historical period is not readily available.

government-sponsored industrial research. I still obtain estimates very similar to my baseline results, excluding the effect of federal funding to private-sector R&D as a confounding driver of the estimates shown in Figure 4.

Second, I introduce commuting zone-specific time trends in my baseline specification, controlling for unobserved time-varying factors unique to each commuting zone and evolving systematically over time. I aim to further address the concern that commuting zones where SDP-funded universities were located and those hosting universities excluded from the program may have already been on different patenting trajectories, regardless of SDP funding. Panel (c) of Figure 5 shows comparable estimates to the baseline, still indicating a positive effect of the SDP on local patenting. I detect two main differences. First, the positive and statistically coefficients for the post-SDP period are slightly smaller than the baseline ones. Second, coefficients for the last years in the sample are estimated close to zero.

My third robustness check involves the construction of a control group with Mahalanobis matching. For each commuting zone hosting SDP-funded universities, I select a control unit which minimizes the Mahalanobis distance between a set of features observed prior to SDP funding. As potential controls I consider all commuting zones hosting a research university not funded by the SDP. I adopt the following matching features: average patenting between 1960-1964; total employment, employment shares by industry (1-digit SIC codes), and total establishments, all in 1964 (County Business Patterns historical data digitized by Eckert et al. 2022); and total population in 1960. I find a matched unit for 17 commuting zones out of the total of 24 hosting an SDP-funded university.

Panel (d) of Figure 5 reports β_{τ} estimates from Equation 2 using the Mahalanobis-matched control group. The results are still similar to the baseline specification, excluding differential trends in the period prior to SDP awards and indicating an increase in local patenting for commuting zones hosting SDP-funded universities. The positive effect starts three years after SDP grants were awarded and lasts for the next ten years, with three positive coefficients in the post-SDP period imprecisely estimated compared to baseline estimates.

Lastly, similar to section 5.1, I conduct two permutation tests. First, I re-estimate Equation 2 iteratively excluding each commuting zone in the sample. Panel (e) of Figure 5 reports results broadly similar to the baseline, indicating that no single commuting zone—whether hosting an SDP-funded university or an institution excluded from the program—is driving the results. Second, I randomly assign a placebo SDP-funded status and timing to all commuting zones not hosting any true funded university, estimating a two-period difference-in-differences version of my baseline specification 1,000 times. Panel (f) in Figure 5 shows that the true estimate (vertical blue line) lies in the right tail of the placebo estimates' distribution, ruling out the possibility of that it is driven by random chance.

Taken together, the results presented in this section support a causal interpretation of the positive effect of the Science Development Program on local patenting. In the next sections, I first explore how the patenting effects vary across and within commuting zones, assignees, and technology fields. I then test effects based on a measure of exposure to local universities' research, I investigate the mechanisms underlying the patenting increase and, lastly, I assess any effect on commuting zones geographically close to those hosting SDP-funded institutions.

6.2. Heterogeneity

Commuting zone characteristics – I begin by investigating how the effects of the SDP varied across commuting zones with different pre-existing industrial R&D characteristics. First, I focus on the size of local R&D-intensive sectors by calculating the number of patents per capita filed during the period immediately preceding the SDP introduction (1960-1964). I then estimate Equation 2 separately for commuting zones with either above-median or below-median patenting prior to the SDP-introduction. Panel (a) in Figure 6 reports the estimation results. Commuting zones in the above-median group display estimates similar to the baseline, while those in the below-median group report estimates clustered around zero and not statistically significant. I obtain comparable results when I consider also the quality of pre-SDP patenting, by categorizing commuting zones based on their total citation-weighted patents per capita filed between 1960 and 1964 (Panel (b)).

Next, I consider a measure of industrial R&D's capacity to productively use scientific knowledge for the development of new technologies. Specifically, I calculate the share of each commuting zone's total pre-SDP patents that cite the scientific literature. Panel (c) in Figure 6 shows that the increase in patenting observed after the SDP grants were awarded is entirely concentrated in commuting zones with an above-median share of pre-SDP patents citing the scientific literature.[19]

Distance to the local university – In Figure 7, I test the SDP effects by patents' distance to the nearest local university. For each patent in the sample, I geocode the inventor's city and compute its distance to the closest SDP-funded or top-ranked university in the same commuting zone. I then re-estimate Equation 2 four times, restricting the dependent variable to patents whose inventors lie at up to 5 miles, 5–10, 10–20, and beyond 20 miles.[20]

[19] Appendix Figure Figure C2 displays the kernel density distributions of each pre-SDP characteristic. Appendix Table C1 reports weak correlations between commuting zones' pre-SDP patents per capita and the share of patents citing the scientific literature, indicating that these two measures capture different characteristics of the local economies in my sample.

[20] I choose distance cut-offs aligning with the quartiles of the distance distribution (Appendix Figure C3).

Panel (a) shows that patents originating within 5 miles of the local university account for the entire SDP effect: the estimated coefficients are flat leading to the SDP awards and, beginning six years after the grants, rise to roughly twice the magnitude of my baseline estimates, retaining the same size and statistical significance through the end of the sample. Panels (b)–(d) reveal no discernible impact beyond the 5-mile radius, with only few positive and statistically significant coefficients in the 10–20-mile band.

Assignee type – Next, I investigate the type of assignee driving the patenting effect. I differentiate among three categories: firms, universities, and independent inventors, that is, inventors patenting in their own name and not affiliated with any particular organization. The majority of patents in my sample are filed by firms, accounting for approximately 84% of filings in each technology field, commuting zone, and year cell, while patents from independent inventors account for around 15% of filings. University patents are rare, constituting less than 2% of filings in each cell.[21]

Panel (a) of Figure 8 reports estimates where I restrict the dependent variable to patents filed by firms. The estimated coefficients' size, statistical significance, and temporal evolution are almost identical to my baseline estimates. In panel (b), the dependent variable includes only patents filed by independent inventors. I detect few positive and statistically significant coefficients between the fifth and tenth year following the SDP introduction, although the point estimate is smaller than panel (a) and the rest of other coefficients are estimated close to zero and mostly statistically insignificant. In panel (c), the dependent variable includes only university patents. Most estimated coefficients are statistically indistinguishable from zero, with few positive and statistically significant coefficients in the final years of the sample. These results, combined with the relatively small proportion of patents filed from independent inventors and (especially) universities, suggest that the patenting increase following the SDP introduction was mainly driven by technologies developed by firms.

Next, I distinguish between patents from incumbent assignees—that is, those observed patenting prior to the introduction of the Science Development Program—and patents from new assignees, those filing their first patent in a given year t. The second group includes both organizations established after the SDP introduction and pre-existing entities which never patented. In panel (d) of Figure 8, the dependent variable includes only patents from incum-

[21]I define patents from firms as those from the patent assignee name string contains the following keywords: "co", "co.", "company", "corp", "corporation", "industries", "limited", "incorporated", "inc", "ltd", "llc", "plc". Similarly, I identify university patents by looking for terms in the assignee name strings such as "university", "institute of technology", "regents of the", or university acronymis such as "mit", "caltech", "uc berkeley", "ucla", "nyu", "ut austin", "umass", or "ucsf". Lastly, I define patents filed by independent inventors as those filed by single inventors and where the name of the inventor and the assignee coincide and where the assignee string does not contain any of the previous keywords for firms or university patents.

bent assignees. The estimates are similar to those considering all patents, with slightly larger coefficients, remaining positive and statistically significant for the entire post-SDP period. In panel (e), I focus only on new assignees' patents. I still estimate positive coefficients for the majority of years in the period following the SDP grants, although most of them are statistically indistinguishable from zero. In Appendix Figure D3, I further distinguish between local incumbent assignees and incumbents from a different location, which never patented in the focal commuting zone prior to the SDP introduction. The results indicate that the patenting increase effect is driven entirely by local incumbent assignees.

Technology category – In Figure 9, I estimate Equation 2 for each of six categories grouping the 37 technology fields in my dataset. I find estimates comparable to the baseline for patents in the electrical & electronics category and, with fewer precisely estimated coefficients, in the chemicals one. The computers & communications and the mechanical category report only limited positive effects, while all coefficients for the residual "others" category are statistically indistinguishable from zero. I estimate a decline for drugs & medical patents, with coefficients for the period following the SDP awards consistently negative and statistically significant through the last ten years in the sample. Appendix Figure D4 replicates the analysis after removing the Bay Area and Boston commuting zones from the comparison group; the results are unchanged, ruling out the emerging biotechnology clusters in those locations as the source of my negative estimates (Zucker et al. 1998; Hughes 2011).

These findings indicate that the overall patenting increase in commuting zones hosting SDP-funded was driven by fields closely connected with the disciplines that received the bulk of SDP funds (physics and chemistry, NSF 1977a). Before the grants, these fields dominated local inventive output, averaging roughly 8 electrical & electronic and 14 chemical patents per year, compared with only about 1.5 drug-related patents. The decline in drugs & medical may thus reflect a slight reallocation of inventive effort away from technologies related to disciplines which attracted less SDP support (biological sciences, NSF 1977a) and which constituted smaller local clusters.[22]

Overall, the results in this section show that the increase in patenting was driven by firms already located in the commuting zones hosting SDP-funded universities, at close proximity to those institutions, and in areas with both high pre-existing levels of industrial R&D and high absorptive capacity for scientific knowledge. In other words, the increased funding of the

[22] A NSF report on the Science Development Program notes: "the question of the seemingly neglected biological sciences has been discussed by Page (1968), who indicates that the institutions had not included them in the [SDP] proposals. Page speculates that the universities might have looked to NIH for support in this area, or that the biological sciences on most campuses were in a transition stage not yet ready for development" (p. 24, NSF 1977a).

Science Development Program appears to have influenced innovation outcomes in the private economy only in locations able to absorb any spillover from higher education institutions, and it did not foster the creation of new organizations developing and patenting new technologies, nor did it attract any of them from other areas of the country.

6.3. Effects by Exposure to Local Universities' Research

I deepen my investigation by studying how the patenting effect estimated in Section 6.1 varies across technology fields, based on their exposure to local universities' research prior to the SDP introduction. To do so, I measure the intellectual proximity between patenting in a given technology field-commuting zone pair ic and the research published by the university located in the same commuting zone c. I construct a measure akin to that developed and validated by Bergeaud et al. (2022) and Bergeaud and Guillouzouic (2024) by calculating:

$$Exposure_{ic} = \sum_{j} s_{icj} \cdot s_{cj} \tag{3}$$

where s_{icj} is the share of citations from patents in technology field-commuting zone pair ic to scientific papers published in journal j, between 1960 and 1964, and s_{cj} is the share of scientific papers from the university hosted in commuting zone c published in journal j, also between 1960 and 1964.[23,24]

In practice, $Exposure_{ic}$ measures the overlap between the scientific journals where a given university published and where R&D-intensive firms in the local technology field source their knowledge. The measure exhibits substantial variation across technology field-commuting zone pairs and, crucially, between the same technology fields across different locations. It displays a skewed distribution, with over 60% of pairs with $Exposure_{ic}$ equal to 0, a maximum value of 0.093, and mean equal to 0.002 (Appendix Figure C4 and Table C2).

I introduce this exposure measure in my analysis by estimating Equation 2 and adding $Exposure_{ic}$ as a third term to the interaction $I_\tau \times SDP_c$. Figure 10 reports the estimated β_τ for the triple interaction $I_\tau \times SDP_c \times Exposure_{ic}$. In the period prior to the SDP introduction all coefficients are estimated close to zero and display an overall flat trend. The coefficients become positive and statistically significant starting in the second year after the SDP introduction and remain positive and—except few coefficients—statistically significant for the next twelve years. The estimates for the final years in the sample are all estimated close to zero and not statistically

[23] Most commuting zones host only one SDP-funded or top-ranked university. For those hosting more than one, I consider the joint research output of all of such institutions in the commuting zone (e.g., the commuting zone of Pittsburgh hosting both Carnegie Mellon University and the University of Pittsburgh).

[24] Virtually all citations to scientific articles used to construct s_{icj} are found within the main text of the patents. Although the USPTO introduced citations to prior art on patent documents' front page in 1947, front-page references to the scientific literature remained almost non-existent until the early 1970s.

significant.

This result indicates that the positive effect of the SDP introduction on local patenting was larger in technology fields with higher exposure to local universities' research. In particular for a one-standard-deviation increase in $Exposure_{ic}$, the positive and statistically significant coefficients estimated for the post-SDP period imply a patenting increase between 4.1% and 8.2%.

In order to verify that this result is not driven by differential patenting trajectories between higher- and lower-exposed technology field-commuting zone pairs, I re-estimate this specification introducing technology field-commuting zone-specific time trends. The estimates are reported in Appendix Figure D5 and show results comparable to Figure 10.

6.4. Mechanisms

Based on the analysis of university outcomes presented in Section 5, there are three potential (non-mutually exclusive) mechanisms underlying the increase in local patenting after the Science Development Program grants were awarded. First, the increased availability of scientific human capital in local economies. New PhD graduates remaining in the labor market hosting their university upon graduation, taking positions in industrial R&D, may have directly contributed to patenting and also diffused scientific knowledge useful for the development of new technologies. Second, larger departments at local universities may have increased the opportunity of local firms to establish formal collaborations with local academic scientists. This may have resulted on joint projects giving rise to patents (and also scientific publications) co-filed by academic and industrial scientists, all the while increasing the diffusion of scientific knowledge to industrial R&D. Third, funded universities' increased research capacity may have intensified scientific knowledge spillovers from universities to firms through informal channels.

Since all mechanisms may involve the diffusion of scientific knowledge, I start this section by investigating whether the increase in patenting is accompanied by a rise in local R&D's reliance on the scientific literature. Following Ahmadpoor and Jones (2017), I construct a citation network connecting patents to scientific publications and compute the shortest path from each patent to any scientific publication (distance metric D). I categorize patents into three groups based on their distance D. The first group comprises patents that directly cite a scientific publication, with $D = 1$. The second group includes patents indirectly linked to a scientific publication through citations to other patents, with $D \in \{2, 3, 4\}$. The third group consists of patents that are more remotely connected to the scientific literature via citations to other patents or are fully unconnected from it ($D \geq 5$).[25]

[25] Ahmadpoor and Jones (2017) construct a citation network using only front-page citations, whether between

In Figure 11, panel (a), I re-estimate Equation 2 by considering only patents that directly cite a scientific publication ($D = 1$). The estimates still show no evidence of pre-trends and indicate a sharp surge in patents directly relying on science in commuting zones hosting SDP-funded universities, indicating a quantitatively larger effects than my estimates for the full sample (Figure 4). In panel (b), I focus only on patents indirectly linked to the scientific literature ($D \in \{2, 3, 4\}$). The coefficients for the pre-SDP period show no distinct trend, while those for the post-SDP period progressively become positive and increase in magnitude, although most estimates are imprecise, with p-values ranging between 0.05 and 0.1 for periods 5 to 14. In panel (c), I consider only patents remotely connected or unconnected to the scientific literature ($D \geq 5$), still finding a patenting increase after the SDP awards, although the effect is quantitatively smaller and shorter-lived than the estimates for the full sample and, especially, for patents at $D = 1$.

Table 4 (columns 4 to 6), presents similar findings from an equivalent two-period difference-in-difference estimation. Panel A reports results for the entire study period, while panel B restricts the analysis to the first 14 years post-SDP awards, where most patenting effects are concentrated. The size of the coefficients is similar across panels A and B, although those in panel A are often less precisely estimated. Specifically, in panel A, columns 4 and 5 indicate that patents at $D = 1$ and at $D \in \{2, 3, 4\}$ increased by 16% and 35%, respectively, in commuting zones hosting SDP-funded universities following the SDP awards. The estimate for patents at $D \geq 5$ in column 6 is also positive but smaller and not statistically significant.

In columns 7 to 9, I focus on the share of patents in each group relative to total patenting in each commuting zone-technology field-year cell, estimating a two-period difference-in-differences specification equivalent to Equation 2 by OLS. The estimates indicate an increase in the proportion of patented inventions relying on the scientific literature, mostly driven by patents directly citing a scientific publication. In particular, the coefficient in column 7 (panel A) implies that the proportion of patents at $D = 1$ increased by 1.3 percentage points following the SDP awards, an economically significant effect representing approximately 6.5% of the dependent variable's mean.

These results indicate that the observed increase in patenting after the SDP introduction was primarily driven by a growth in patents directly reliant on science or closely connected to the

patents or between patents and scientific articles. I use only in-text citations between patents (data from Verluise et al. 2020) and both in-text and front-page citations between patents and scientific articles (data from Marx and Fuegi 2022). Since in-text citations among patents are rarer than front-page ones (Verluise et al. 2020), the groups of patents indirectly connected to the scientific literature is smaller than in Ahmadpoor and Jones (2017). Marx and Fuegi (2022) replicate Ahmadpoor and Jones's (2017) analysis using both in-text and front-page citations to the scientific literature, finding overall comparable results, with a higher proportion of patents at $D = 1$. I obtain almost identical results to those presented below when I use only in-text citations to scientific publications (Appendix Figure D6 and Table D1).

scientific publication-patent frontier. The rise in the proportion of patents closely linked to the scientific literature suggests that commuting zones hosting SDP-funded universities intensified their reliance on scientific research for their inventions following the SDP awards. Besides highlighting the diffusion of scientific knowledge from universities to firms as a mechanism for the post-SDP increase in local patenting, these results provide additional evidence on the linkage between universities and innovation outcomes in the local economies hosting them after the SDP grants were awarded.

I continue my investigation by focusing on the role of scientific human capital. I estimate a two-period difference-in-differences specification based on Equation 2 restricting y_{ict} to the number of patents co-filed by a PhD graduate from a university located in commuting zone c. Column 1 in Table 5 reports a positive and statistically significant coefficient, similar in size across panels A and B (full and restricted post-SDP period), implying an increase in the number of patents from local PhD graduates of about 30% for commuting zones hosting SDP-funded universities. In column 2, I focus only on patents not listing any local PhD graduate, still finding positive estimates although smaller in size and less precisely estimated. These results suggest that the direct contribution of local PhD graduates to technology development was one of the mechanisms driving the patenting increase observed after the SDP introduction.

Two points are important for interpreting this result. First, the average number of patents co-filed by local PhD graduates is relatively small compared to all patents in a commuting zone-technology field-year cell, implying that the direct contribution of local PhD graduates cannot fully explain the overall patenting increase post-SDP awards. Second, my measure of PhD graduates' patenting does not capture all patents filed by this group, as it relies on a non-deterministic linkage procedure between inventors and PhD dissertations' authors, matching only a fraction of the entire population of interest. Consequently, while still remaining small relative to the full set, the true proportion of patents co-filed by local PhD graduates is likely higher than measured.[26]

Next, I test the whether direct collaborations between academic and industrial scientists may also have contributed to the post-SDP patenting increase. Column 3 in Table 5 reports two-period difference-in-differences estimates where y_{ict} equals the number of patents co-filed by an academic scientist affiliated with a university located in commuting zone c. The estimates are statistically indistinguishable from zero, and the lower number of observations contributing

[26]The type of linking algorithm I adopted typically matches between 25% and 30% of candidate links (Abramitzky et al. 2021). Due to the absence of a gold-standard dataset for this specific inventor demographic, I cannot assess the exact match rate for patenting PhD graduates. Based on linkage rates from similar algorithms, the average number of patents co-filed by local PhD graduates in my sample could be three to four times higher than observed.

to the estimates relative to columns 1 and 2 suggests that academic-industry co-patenting events are relatively rare in my sample. In column 4, the dependent variable is the number of scientific articles co-published by academic scientists and industrial inventors co-located in commuting zone c. The coefficient is statistically insignificant, and the small number of such publications in my sample further excludes direct collaborations between academic and industrial scientists as a mechanism behind the patenting increase.

Taken together, these results indicate the diffusion of scientific knowledge from universities to private sector R&D and the increased availability of scientific human capital as the mechanisms underlying the patenting increase observed after the SDP introduction. While my estimates rule out any influence of direct university-industry collaboration, they do not exclude informal channels connecting the two. The increase in patents not involving local PhD graduates may reflect harder-to-measure knowledge diffusion mechanisms, such as more frequent local scientific conferences open to industrial scientists or overall higher chances of contact between scientists working for firms and a larger number of their academic peers.

6.5. Effects on Nearby Commuting Zones

I conclude the analysis by asking whether the patenting gains observed for SDP-funded commuting zones came at the expense of nearby areas. I re-estimate Equation 2 after replacing the treatment indicator SDP_c with $NearSDP_c$, a dummy equal to 1 for commuting zones lying within a given radius of the closest SDP-funded university and 0 for commuting zones located within the same distance from the nearest top-ranked university. I estimate the model separately for four distance bands, corresponding to the quartiles of the SDP-university-distance distribution: up to 100 miles, 100–200, 200–300, and beyond 300 miles. When constructing $NearSDP_c$, I alternately consider distance to the full set of SDP-funded universities and to only those hosted by above-median-patenting commuting zones, the subset that drives most of the patenting effect. The sample includes every U.S. commuting zone not hosting SDP-funded or top-ranked universities.[27]

Figure 12 presents my estimates. For commuting zones within 100 miles and for those 100–200 miles away, the coefficients are close to zero and are broadly statistically insignificant (panels (a) and (b)). In the 200–300-mile band (panel (c)) late-period coefficients turn negative when distance is measured from SDP-funded universities in above-median-patenting commuting zones, although most are not statistically significant. Beyond 300 miles (panel (d)), all coefficients are statistically indistinguishable from zero when distance is defined relative to above-median-patenting commuting zones. Using the full set of SDP-funded universities yields

[27] For each commuting zone, I measure distance between the centroid of its most populous county in 1960 and the coordinates of the nearest university, either SDP-funded or top-ranked.

some negative and significant estimates in the final sample years.

I further investigate effects on neighboring commuting zones in two ways. First, I repeat the exercise with 50-mile distance bins, again yielding coefficients mostly statistically indistinguishable from zero (Appendix Figure D7). Second, I test as dependent variable the number of inventors permanently leaving a commuting zone in any given year, finding results broadly consistent with a null effect (Appendix Figure D8).

Overall, the results in this section exclude a large reallocation of inventive activity from areas near commuting zones hosting SDP-funded universities.

7. Conclusions

This paper studies the effects of the NSF Science Development Program on universities and on local innovation. Leveraging excluded elite universities from the program as a comparison group, I show that the large institutional grants awarded by the Science Development Program enabled funded universities to enlarge their departments and PhD programs, and to increase their publications. I then show that such an increase in universities' research capacity positively influenced innovation in their local economies. My estimates indicate a sizeable increase in patenting, mostly due to incumbent firms with inventors located near funded universities. This effect is driven by commuting zones with established R&D-intensive sectors and is larger in technology fields with high exposure to local universities' research. I provide evidence indicating that the diffusion of scientific knowledge from universities to industrial R&D and the increased availability of scientific human capital are the two main mechanisms behind the patenting increase.

These findings provide new evidence on the role of universities in their local economies and, more broadly, the effects of public R&D funding on innovation (Bryan and Williams 2021). Prior research has highlighted the positive effects of establishing new higher education institutions on innovation and economic development (Dittmar and Meisenzahl 2022; Andrews 2023). It has also shown a positive link between funding basic research and innovation via principal investigator programs or initiatives targeting individual laboratories (Azoulay et al. 2019b; Bergeaud et al. 2022). This study demonstrates that increasing the research capacity of already-established universities can generate significant positive spillovers to technology development in the private sector.

The fact that the increase in patenting is mostly due to incumbent firms in locations with high pre-existing levels of R&D and stronger reliance on scientific knowledge for technology development underscores the critical role of the private sector's absorptive capacity (Cohen and Levinthal 1989, 1990) in capturing the spillovers of public funding to universities. This emphasizes the challenges and complexities faced by interventions aimed at stimulating private sector R&D and entrepreneurship (Lerner 2013).

There are two main areas not examined by this paper, which I intend to explore in future research. First, my analysis focuses only on the effects of the Science Development Program on local economies. It is possible that the increase in PhD graduates and scientific publications influenced science and innovation outcomes in the broader national economy, beyond the locations hosting funded universities. Second, although I examine the effects of the SDP over a relatively long horizon, some of its effects may require an even longer time frame to become evident. This may include scientific advancements that enable the development of new technologies several decades after their initial discovery, or the contributions of scientists—whose graduate studies were supported by the Science Development Program—to the training of new generations of researchers.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010): "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*, 105, 493–505.

Abramitzky, R., L. Boustan, K. Eriksson, J. Feigenbaum, and S. Pérez (2021): "Automated linking of historical data," *Journal of Economic Literature*, 59, 865–918.

Abramitzky, R., L. P. Boustan, and K. Eriksson (2012): "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration," *American Economic Review*, 102, 1832–56.

——— (2014): "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, 122, 467–506.

Agrawal, A. and R. Henderson (2002): "Putting patents in context: Exploring knowledge transfer from MIT," *Management Science*, 48, 44–60.

Ahmadpoor, M. and B. F. Jones (2017): "The dual frontier: Patented inventions and prior scientific advance," *Science*, 357, 583–587.

Akcigit, U., J. Grigsby, T. Nicholas, and S. Stantcheva (2022): "Taxation and Innovation in the Twentieth Century," *The Quarterly Journal of Economics*, 137, 329–385.

Andrews, M. J. (2021a): "Historical patent data: A practitioner's guide," *Journal of Economics & Management Strategy*, 30, 368–397.

——— (2021b): "Local Effects of Land Grant Colleges on Agricultural Innovation and Output," in *Economics of Research and Innovation in Agriculture*, ed. by P. Moser, Univ. of Chicago Press, 139–175.

——— (2023): "How do institutions of higher education affect local invention? Evidence from the establishment of US colleges," *American Economic Journal: Economic Policy*, 15, 1–41.

Andrews, M. J. and A. Smith (2023): "Do Local Conditions Determine the Direction of Science? Evidence from Land Grant Colleges," *Mimeo.*

Arora, A., S. Belenzon, K. Kosenko, J. Suh, and Y. Yafeh (2021a): "The rise of scientific research in corporate america," *NBER WP 29260, National Bureau of Economic Research.*

Arora, A., S. Belenzon, and A. Patacconi (2019): "A theory of the US innovation ecosystem: evolution and the social value of diversity," *Industrial and Corporate Change*, 28, 289–307.

Arora, A., S. Belenzon, A. Patacconi, and J. Suh (2020): "The Changing Structure of American Innovation: Some Cautionary Remarks for Economic Growth," in *Innovation Policy and the Economy*, ed. by J. Lerner and S. Stern, Univ. of Chicago Press, vol. 20, 39–93.

Arora, A., S. Belenzon, and L. Sheer (2021b): "Knowledge spillovers and corporate investment in scientific research," *American Economic Review*, 111, 871–898.

Autor, D. H. and D. Dorn (2013): "The growth of low-skill service jobs and the polarization of the US labor market," *American Economic Review*, 103, 1553–1597.

Azoulay, P., W. Ding, and T. Stuart (2009): "The impact of academic patenting on the rate, quality and direction of (public) research output," *The Journal of Industrial Economics*, 57, 637–676.

Azoulay, P., C. Fons-Rosen, and J. S. Graff Zivin (2019a): "Does Science Advance One Funeral at a Time?" *American Economic Review*, 109, 2889–2920.

Azoulay, P., J. S. Graff Zivin, D. Li, and B. N. Sampat (2019b): "Public R&D investments and private-sector patenting: evidence from NIH funding rules," *The Review of Economic Studies*, 86, 117–152.

Babina, T., A. X. He, S. T. Howell, E. R. Perlman, and J. Staudt (2023): "Cutting the innovation engine: how federal funding shocks affect university patenting, entrepreneurship, and publications," *The Quarterly Journal of Economics*, 138, 895–954.

Bergeaud, A. and A. Guillouzouic (2024): "Proximity of firms to scientific production," *Annals of Economics and Statistics*, 105–134.

Bergeaud, A., A. Guillouzouic, E. Henry, and C. Malgouyres (2022): "From public labs to private firms: magnitude and channels of R&D Spillovers," *Centre for Economic Performance Discussion Paper No. 1882*.

Bergeaud, A. and C. Verluise (2024): "A new dataset to study a century of innovation in Europe and in the US," *Research Policy*, 53, 104903.

Bianchi, N. and M. Giorcelli (2022): "The dynamics and spillovers of management interventions: Evidence from the training within industry program," *Journal of Political Economy*, 130, 1630–1675.

Bikard, M. and M. Marx (2020): "Bridging academia and industry: How geographic hubs connect university science and corporate technology," *Management Science*, 66, 3425–3443.

Blundell, R., R. Griffith, and J. V. Reenen (1995): "Dynamic Count Data Models of Technological Innovation," *The Economic Journal*, 105, 333–344.

Bryan, K. A. and Y. Ozcan (2021): "The impact of open access mandates on invention," *The Review of Economics and Statistics*, 103, 954–967.

Bryan, K. A. and H. L. Williams (2021): "Innovation: market failures and public policies," in *Handbook of Industrial Organization*, ed. by K. Ho, A. Hortaçsu, and A. Lizzeri, Elsevier, vol. 5, 281–388.

Bush, V. (1945): *Science, the endless frontier: A report to the President*, Washington: Government Printing Office.

Cantoni, D. and N. Yuchtman (2014): "Medieval universities, legal institutions, and the commercial revolution," *The Quarterly Journal of Economics*, 129, 823–887.

Cartter, A. M. (1966): *An assessment of quality in graduate education.*, Washington, D.C.: American Council of Education.

Catalini, C., C. Fons-Rosen, and P. Gaulé (2020): "How Do Travel Costs Shape Collaboration?" *Management Science*, 66, 3340–3360.

Cockburn, I. and R. Henderson (1996): "Public–private interaction in pharmaceutical research," *Proceedings of the National Academy of Sciences*, 93, 12725–12730.

Cockburn, I. M. and R. M. Henderson (2001): "Publicly funded science and the productivity of the pharmaceutical industry," in *Innovation Policy and the Economy*, ed. by A. B. Jaffe, J. Lerner, and S. Stern, MIT Press, vol. 1, 1–34.

Cohen, W. M. and D. A. Levinthal (1989): "Innovation and learning: the two faces of R & D," *The economic journal*, 99, 569–596.

——— (1990): "Absorptive capacity: A new perspective on learning and innovation," *Administrative science quarterly*, 128–152.

Cohen, W. M., R. Nelson, and J. P. Walsh (2000): "Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)," *NBER WP 7552, National Bureau of Economic Research.*

Comptroller General of the U.S. (1976): "Geographical Distribution of Federal Science Funds To Colleges and Universities," *General Accounting Office, Washington.*

Dittmar, J. and R. R. Meisenzahl (2022): "The research university, invention and industry: evidence from German history," *CEPR Discussion Paper No. 1856.*

Drew, D. E. (1975): *Science development: An evaluation study*, National Academy of Sciences.

Eckert, F., K.-l. Lam, A. Mian, K. Muller, R. Schwalb, and A. Sufi (2022): "The Early County Business Pattern Files: 1946-1974," *NBER WP 30578, National Bureau of Economic Research.*

Fleming, L., H. Greene, G. Li, M. Marx, and D. Yao (2019): "Government-funded research increasingly fuels innovation," *Science*, 364, 1139–1141.

Foray, D. and F. Lissoni (2010): "University research and public–private interaction," in *Handbook of the Economics of Innovation*, Elsevier, vol. 1, 275–314.

Furman, J. L. and M. J. MacGarvie (2007): "Academic science and the birth of industrial research laboratories in the US pharmaceutical industry," *Journal of Economic Behavior & Organization*, 63, 756–776.

Geiger, R. L. (1993): *Research and Relevant Knowledge: American Research Universities Since World War II*, Oxford University Press.

——— (1997): "What happened after Sputnik? Shaping university research in the United States," *Minerva*, 349–367.

Giorcelli, M. (2019): "The long-term effects of management and technology transfers," *American Economic Review*, 109, 121–152.

——— (2024a): "Creating the "American Way" of Business: Evidence from WWII in the United States," in *AEA Papers and Proceedings*, vol. 114, 582–585.

——— (2024b): "The Effects of Business School Education on Manager Career Outcomes," *Mimeo.*

Giorcelli, M. and B. Li (2024): "Technology Transfer and Early Industrial Development: Evidence From the Sino-Soviet Alliance," *Mimeo.*

Graham, H. D. and N. Diamond (1997): *The Rise of American Research Universities. Elites and Challenges in the Postwar Era*, The Johns Hopkins University Press.

Griliches, Z. (1990): "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 28, 1661–1707.

Gross, D. P. and M. P. Roche (2023): "Coordinated R&D programs and the creation of new industries," *HBS working paper 124-027.*

Gross, D. P. and B. N. Sampat (2023): "America, jump-started: World War II R&D and the takeoff of the US innovation system," *American Economic Review*, 113, 3323–3356.

——— (2024): "The Government Patent Register: A New Resource for Measuring US Government-Funded Patenting," *NBER WP 32136, National Bureau of Economic Research.*

——— (2025): "The Therapeutic Consequences of the War: World War II and the 20th-Century Expansion of Biomedicine," *NBER WP 33457, National Bureau of Economic Research.*

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001): "The NBER patent citation data file: Lessons, insights and methodological tools," *NBER WP 8498, National Bureau of Economic Research.*

Hausman, J., B. H. Hall, Z. Griliches, et al. (1984): "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," *Econometrica*, 52, 909–938.

Hausman, N. (2022): "University innovation and local economic growth," *The Review of Economics and Statistics*, 104, 718–735.

Henderson, R. and I. Cockburn (1994): "Measuring Competence? Exploring Firm Effects in Pharmaceutical Research," *Strategic Management Journal*, 15, 63–84.

Henderson, R., A. B. Jaffe, and M. Trajtenberg (1998): "Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988," *The Review of Economics and Statistics*, 80, 119–127.

Hughes, S. S. (2011): *Genentech: the beginnings of biotech*, University of Chicago Press.

Hvide, H. K. and B. F. Jones (2018): "University innovation and the professor's privilege," *American Economic Review*, 108, 1860–1898.

Jaffe, A. B. (1989): "Real effects of academic research," *American Economic Review*, 957–970.

Jensen, R. and M. Thursby (2001): "Proofs and prototypes for sale: The licensing of university inventions," *American Economic Review*, 91, 240–259.

Kaltenberg, M., A. B. Jaffe, and M. E. Lachman (2023): "Invention and the life course: Age differences in patenting," *Research Policy*, 52, 104629.

Kantor, S. and A. Whalley (2014): "Knowledge spillovers from research universities: evidence from endowment value shocks," *Review of Economics and Statistics*, 96, 171–188.

Kantor, S. and A. T. Whalley (2023): "Moonshot: Public R&D and growth," *NBER WP 31471, National Bureau of Economic Research.*

Kline, S. J. and N. Rosenberg (1986): "An overview of innovation," in *The Positive Sum Strategy*, ed. by R. Landau and N. Rosenberg, National Academy Press, Washington, D.C.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017): "Technological innovation, resource allocation, and growth," *The Quarterly Journal of Economics*, 132, 665–712.

Lerner, J. (2013): "The Boulevard of Broken Dreams: Innovation Policy and Entrepreneurship," in *Innovation Policy and the Economy*, ed. by J. Lerner and S. Stern, Univ. of Chicago Press, vol. 13, 61–82.

Lerner, J., H. Manley, C. Stein, and H. Williams (2024): "The Wandering Scholars: Understanding the Heterogeneity of University Commercialization," *NBER WP 32069, National Bureau of Economic Research.*

Li, G.-C., R. Lai, A. D'Amour, D. M. Doolin, Y. Sun, V. I. Torvik, Z. Y. Amy, and L. Fleming (2014): "Disambiguation and Co-Authorship Networks of the US Patent Inventor Database (1975–2010)," *Research Policy*, 43, 941–955.

Lissoni, F. (2010): "Academic inventors as brokers," *Research Policy*, 39, 843–857.

Lissoni, F., P. Llerena, M. McKelvey, and B. Sanditov (2008): "Academic patenting in Europe: new evidence from the KEINS database," *Research Evaluation*, 17, 87–102.

Lomask, M. (1976): *A minor miracle: An informal history of the National Science Foundation*, National Science Foundation, Washington, D.C.

Maclaurin, W. R. (1953): "The sequence from invention to innovation and its relation to economic growth," *The Quarterly Journal of Economics*, 67, 97–111.

Mansfield, E. (1986): "Patents and innovation: an empirical study," *Management Science*, 32, 173–181.

——— (1991): "Academic research and industrial innovation," *Research Policy*, 20, 1–12.

——— (1995): "Academic research underlying industrial innovations: sources, characteristics, and financing," *The Review of Economics and Statistics*, 55–65.

Marco, A. C., M. Carley, S. Jackson, and A. Myers (2015): "The USPTO historical patent data files. Two centuries of innovation," *USPTO Economic Working Paper No. 2015-1.*

Marx, M. and A. Fuegi (2022): "Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations," *Journal of Economics & Management Strategy*, 31, 369–392.

McCallum, A., K. Nigam, and L. H. Ungar (2000): "Efficient clustering of high-dimensional data sets with application to reference matching," in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, 169–178.

Mokyr, J. (2002): *The Gifts of Athena: Historical Origins of the Knowledge Economy*, Princeton University Press.

Monath, N., S. Madhavan, C. DiPietro, A. McCallum, and C. Jones (2021): "Disambiguating patent inventors, assignees, and their locations in patentsview," *Tech. Rep.*

Moser, P. (2016): "Patents and innovation in economic history," *Annual Review of Economics*, 8, 241–258.

Mowery, D. C. (1997): "The Bush report after 50 years: Blueprint or relic?" in *Science for the Twenty-first Century. The Bush Report Revisited*, ed. by C. E. Barfield, The AEI Press, Publisherfor the American Enterprise Institute, Washington, D.C.

——— (2010): "Military R&D and innovation," in *Handbook of the Economics of Innovation*, ed. by B. Hall and N. Rosenberg, Elsevier, vol. 2, 1219–1256.

Mowery, D. C., R. R. Nelson, B. N. Sampat, and A. A. Ziedonis (2001): "The growth of patenting and licensing by US universities: an assessment of the effects of the Bayh–Dole act of 1980," *Research Policy*, 30, 99–119.

——— (2004): *Ivory tower and industrial innovation: University-industry technology transfer before and after the Bayh-Dole Act*, Stanford University Press.

Mowery, D. C. and N. Rosenberg (1991): *Technology and the pursuit of economic growth*, Cambridge University Press.

Mowery, D. C. and B. N. Sampat (2006): "Universities in National Innovation Systems," in *The Oxofrd Handbook of Innovation*, ed. by J. Fagerberg, D. C. Mowery, and R. R. Nelson, Oxford University Press, 209–239.

Mowery, D. C., B. N. Sampat, and A. A. Ziedonis (2002): "Learning to patent: Institutional experience, learning, and the characteristics of US university patents after the Bayh-Dole Act, 1981-1992," *Management Science*, 48, 73–89.

Nelson, R. R. (1959): "The simple economics of basic scientific research," *Journal of Political Economy*, 67, 297–306.

——— (1962): "The Link Between Science and Invention: The Case of the Transistor," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, ed. by R. R. Nelson, Princeton University Press.

NSF (1964): "National Science Foundation. Fourteenth Annual Report for the Fiscal Year Ended June 30, 1964," Tech. rep., National Science Foundation.

——— (1967): "Federal Support for Academic Science and Other Educational Activities in Universities and Colleges. Fiscal Years 1963-66," Tech. rep., National Science Foundation.

——— (1977a): "The NSF Science Development Programs, Volume I: A Documentary Report," Tech. Rep. NSF-77-17, National Science Foundation.

——— (1977b): "The NSF Science Development Programs, Volume II: Budgets, Statements of Goals, and Grantees' Summaries of Grant Impact," Tech. Rep. NSF-77-17, National Science Foundation.

Page, H. E. (1968): "The Science Development Program," in *Science Policy and the University*, ed. by H. Orlans and C. K. Arnold, The Brookings Institution, 101–122.

Pezzoni, M., F. Lissoni, and G. Tarasconi (2014): "How to Kill Inventors: Testing the Massacrator© Algorithm for Inventor Disambiguation," *Scientometrics*, 101, 477–504.

Poege, F., D. Harhoff, F. Gaessler, and S. Baruffaldi (2019): "Science quality and the value of inventions," *Science Advances*, 5, eaay7323.

President's Science Advisory Committee (1960): "Scientific progress, the universities, and the federal government," *The White House*.

Priem, J., H. Piwowar, and R. Orr (2022): "OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts," *arXiv preprint arXiv:2205.01833*.

Rosenberg, N. (1982): *Inside the black box: technology and economics*, Cambridge University Press.

Rosenberg, N. and R. R. Nelson (1994): "American universities and technical advance in industry," *Research Policy*, 23, 323–348.

Russell, L. C. and M. J. Andrews (2024): "Historical Place-Based Investments and Contemporary Economic Mobility: Impacts of University Establishment," *Mimeo*.

Russell, L. C., L. Yu, and M. J. Andrews (2024): "Higher education and local educational attainment: evidence from the establishment of US colleges," *Review of Economics and Statistics*, 106, 1146–1156.

Sampat, B. N. (2006): "Patenting and US academic research in the 20th century: The world before and after Bayh-Dole," *Research Policy*, 35, 772–789.

Schweiger, H., A. Stepanov, and P. Zacchia (2022): "The long-run effects of R&D place-based policies: evidence from Russian science cities," *American Economic Journal: Economic Policy*, 14, 322–351.

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

Tabakovic, H. and T. G. Wollmann (2019): "The impact of money on science: Evidence from unexpected NCAA football outcomes," *Journal of Public Economics*, 178, 104066.

Thursby, J. G. and M. C. Thursby (2002): "Who is selling the ivory tower? Sources of growth in university licensing," *Management Science*, 48, 90–104.

Tolbert, C. M. and M. Sizer (1996): "US commuting zones and labor market areas: A 1990 update," *Economic Research Service Staff Paper 9614*.

Trajtenberg, M., G. Shiff, and R. Melamed (2006): "The "names game": Harnessing inventors' patent data for economic research," *NBER WP 12479, National Bureau of Economic Research*.

USPTO (2024): "The USPTO PatentsView database [dataset]," *https://patentsview.org/download/data-download-tables*.

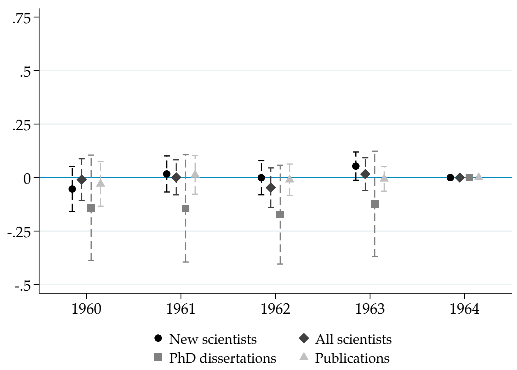
Verluise, C., G. Cristelli, K. Higham, and G. de Rassenfosse (2020): "The Missing 15 Percent of Patent Citations," *Mimeo*.

Zucker, L. G., M. R. Darby, and J. S. Armstrong (2002): "Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology," *Management Science*, 48, 138–153.

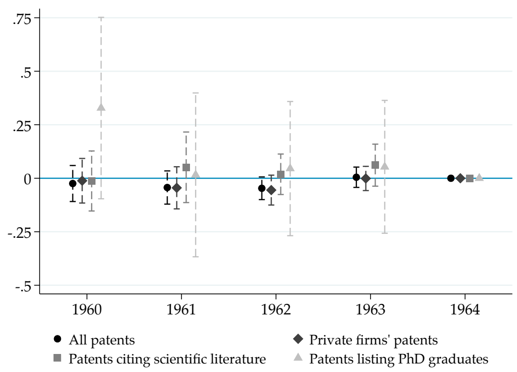
Zucker, L. G., M. R. Darby, and M. B. Brewer (1998): "Intellectual human capital and the birth of US biotechnology enterprises," *American Economic Review*, 88, 290–306.

Figures

Figure 1: University and Commuting Zone Trends Prior to the Science Development Program



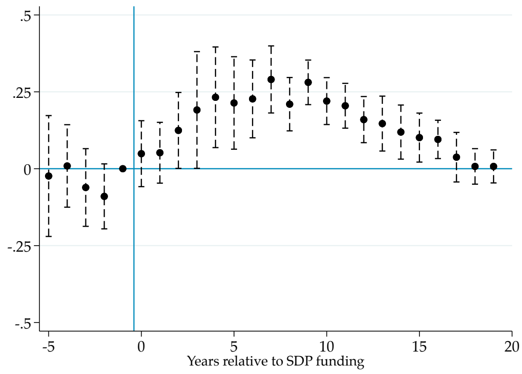
(a) Universities: natural sciences and engineering outcomes



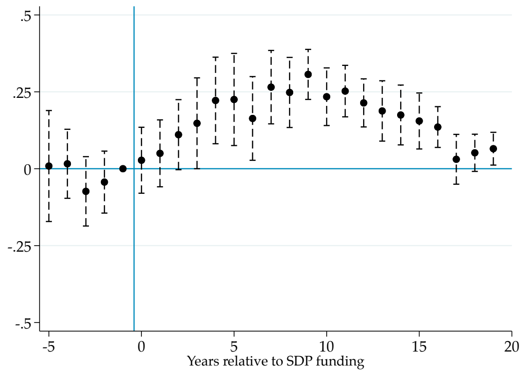
(b) Commuting zones: patenting outcomes

Notes: Panel (a) reports the coefficients from regressions where a university outcome in the natural sciences and engineering is regressed on year dummies interacted with an indicator equal to 1 for universities funded by the SDP and equal to 0 for top-ranked universities. All regressions include year and university fixed effects and standard errors are clustered at the university level. Panel (b) reports the coefficients from regressions where a commuting zone patenting outcome is regressed on year dummies interacted with an indicator equal to 1 for commuting zones hosting an SDP-funded university and equal to 0 for those hosting a top-ranked university. All regressions include year and commuting zones fixed effects and standard errors are clustered at the commuting zone level. For all regressions, the baseline year is 1964. Vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

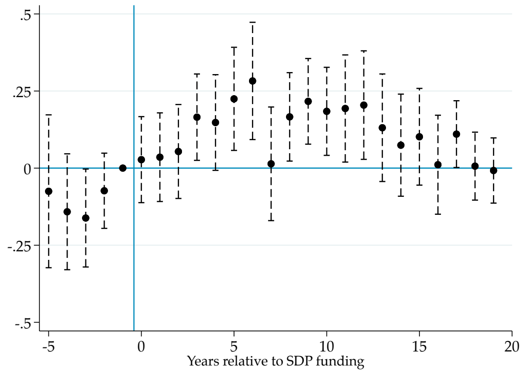
Figure 2: The NSF Science Development Program and University Outcomes



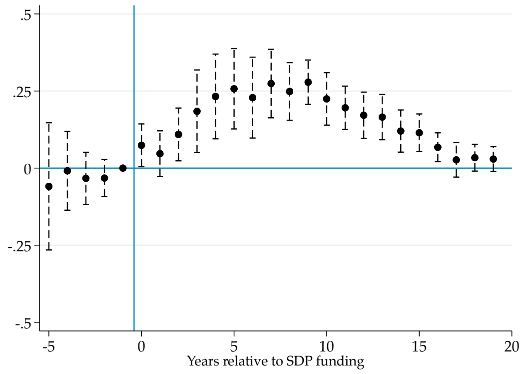
(a) New Scientists



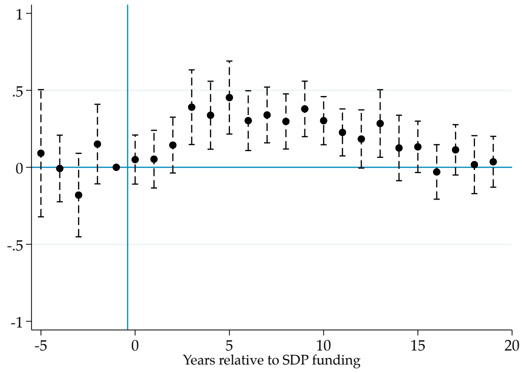
(b) All Scientists



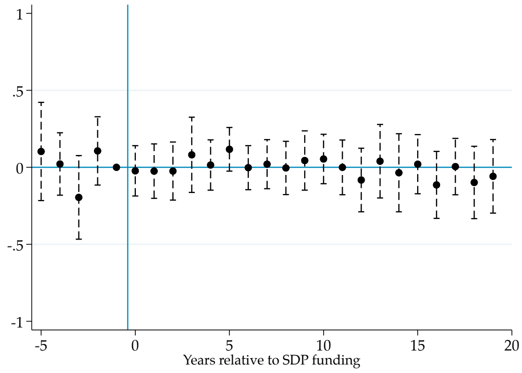
(c) PhD Dissertations



(d) Publications



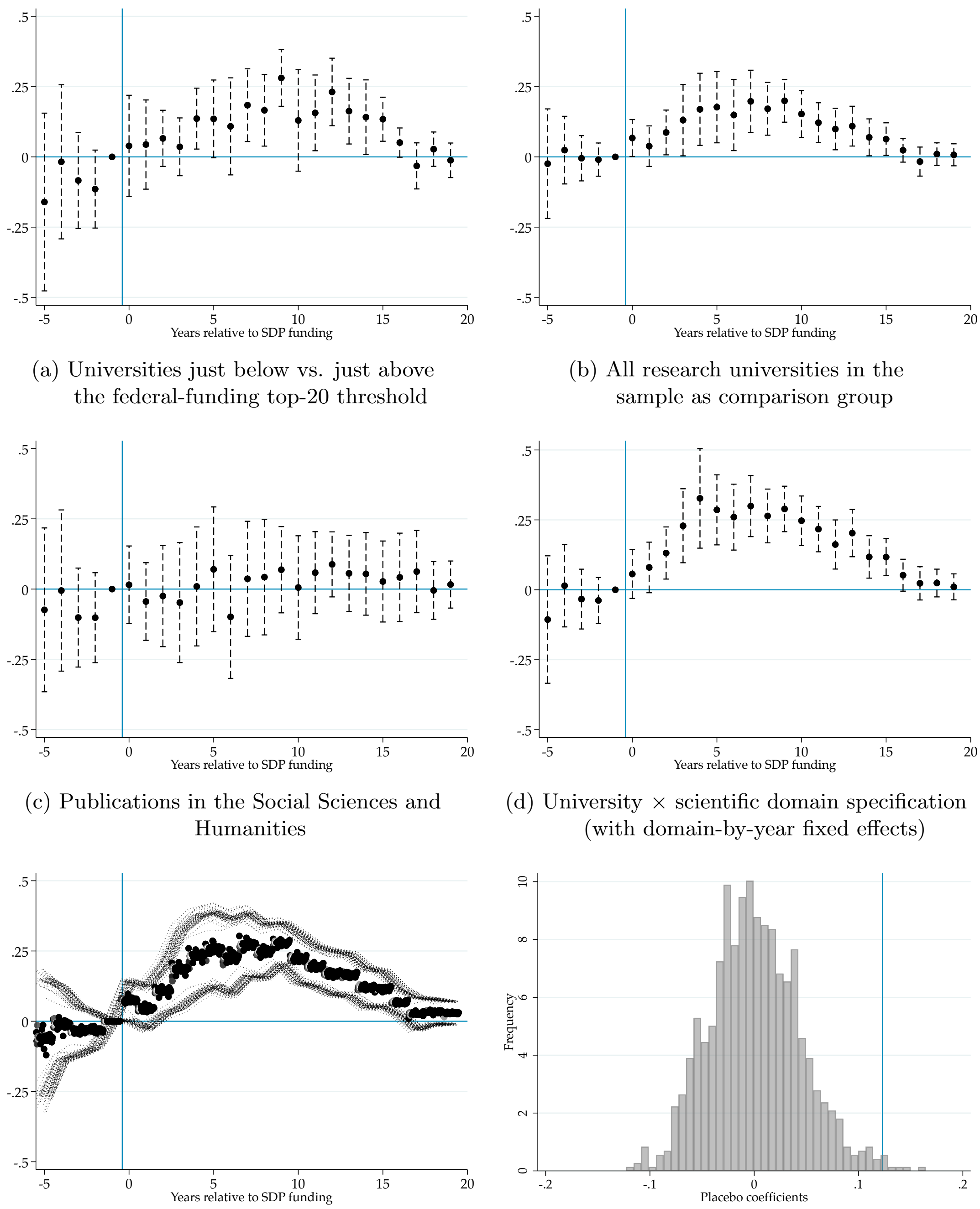
(e) Citation-weighted publications



(f) Average citations

Notes: In panel (a) the dependent variable is the number of new scientists publishing in the natural sciences and engineering affiliated university u in year t, defined as those who never published for university u prior to year t. In panel (b) the dependent variable is the number of all scientists publishing in the natural sciences and engineering affiliated to university u in year t. In panel (c) the dependent variable is the number of new PhD dissertations in the natural sciences and engineering published by university u in year t. In panel (d) the dependent variable is the number of publications in the natural sciences and engineering from university u in year t. In panels (e) the dependent variable is the number of citation-weighted STEM publications from university u in year t. In panel (f) the dependent variable is the average number of citations to STEM publications from university u in year t. The baseline period is $\tau = -1$. Standard errors are clustered at the university level. Vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

Figure 3: The NSF Science Development Program and University Outcomes: Robustness Checks



(a) Universities just below vs. just above the federal-funding top-20 threshold

(b) All research universities in the sample as comparison group

(c) Publications in the Social Sciences and Humanities

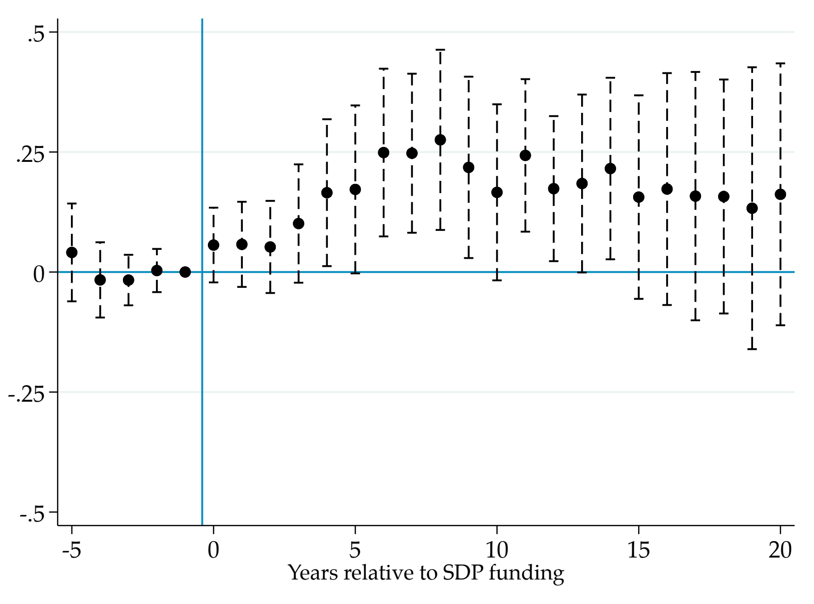
(d) University × scientific domain specification (with domain-by-year fixed effects)

(e) Iteratively exclude universities

(f) Random SDP funding assignment

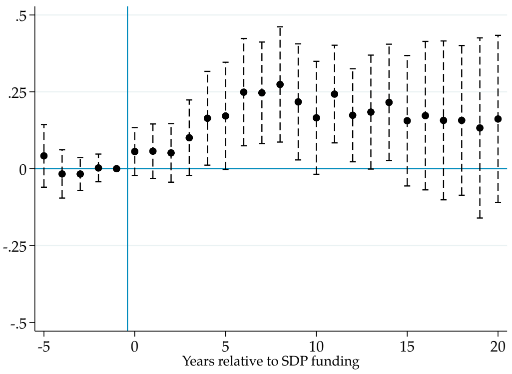
Notes: In all panels, except (c) and (d) the dependent variable is the number of publications in the natural sciences and engineering from university u in year t. In panel (c) the dependent variable is the number of publications in the social sciences and humanities (i.e., Economics, History, Philosophy, Political Science, and Sociology) from university u in year t. In panel (d) the dependent variable is the number of publications for university u in a natural science or engineering field d in year t. The regressions includes year and university fixed effects, along with controls for other institutional funding programs that vary by commuting zone and year. Panel (d) includes also scientific domain-by-year fixed effects. Standard errors are clustered at the university level. In panels (a) to (e) the baseline period is $\tau = -1$ and the vertical bars represent 95% confidence intervals (except in panel (e), where they are denoted by dashed black lines). Panel (f) reports the distribution of coefficients from 1,000 permutation tests, where I estimate a two-period difference-in-differences randomly assigning a placebo SDP-funded status and timing (between 1965 and 1969) to all research universities in my sample not funded by the SDP program. The coefficient from my true estimate is denoted by a vertical blue line. All estimations by Poisson pseudo-maximum likelihood.

Figure 4: The NSF Science Development Program and Local Patenting

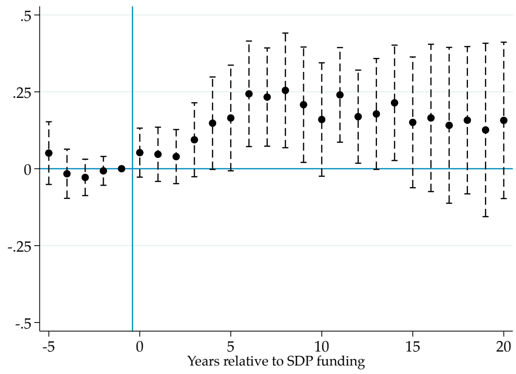


Notes: The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

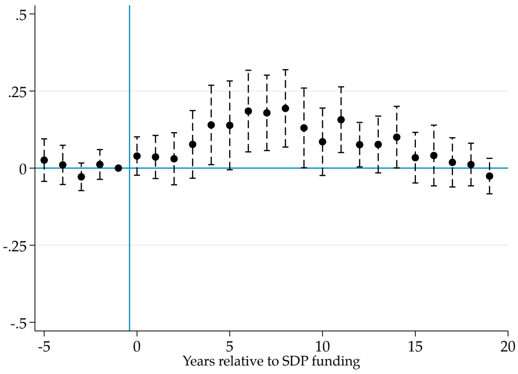
Figure 5: The NSF Science Development Program and Local Patenting: Robustness Checks



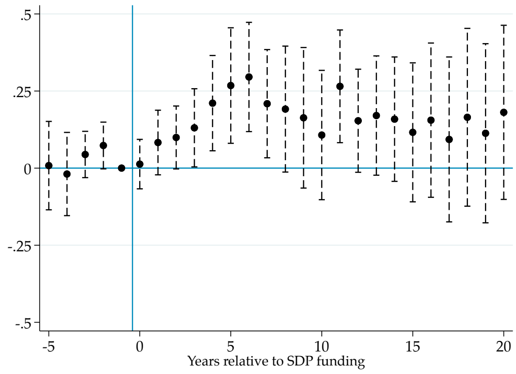
(a) Control for federally-funded patenting share



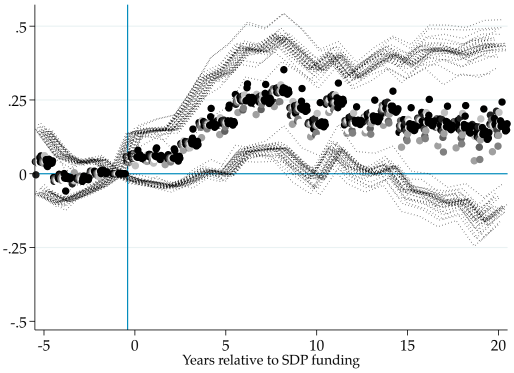
(b) Exclude federally-funded patents



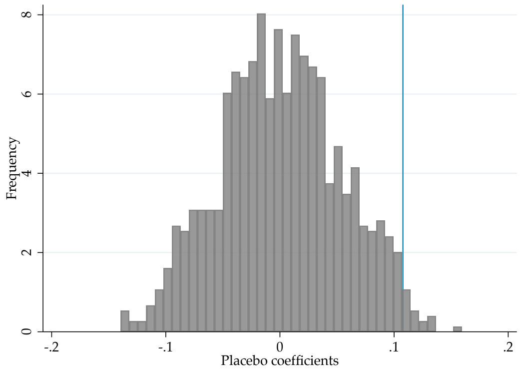
(c) Add commuting zone-specific time trends



(d) Mahalanobis-matched control group



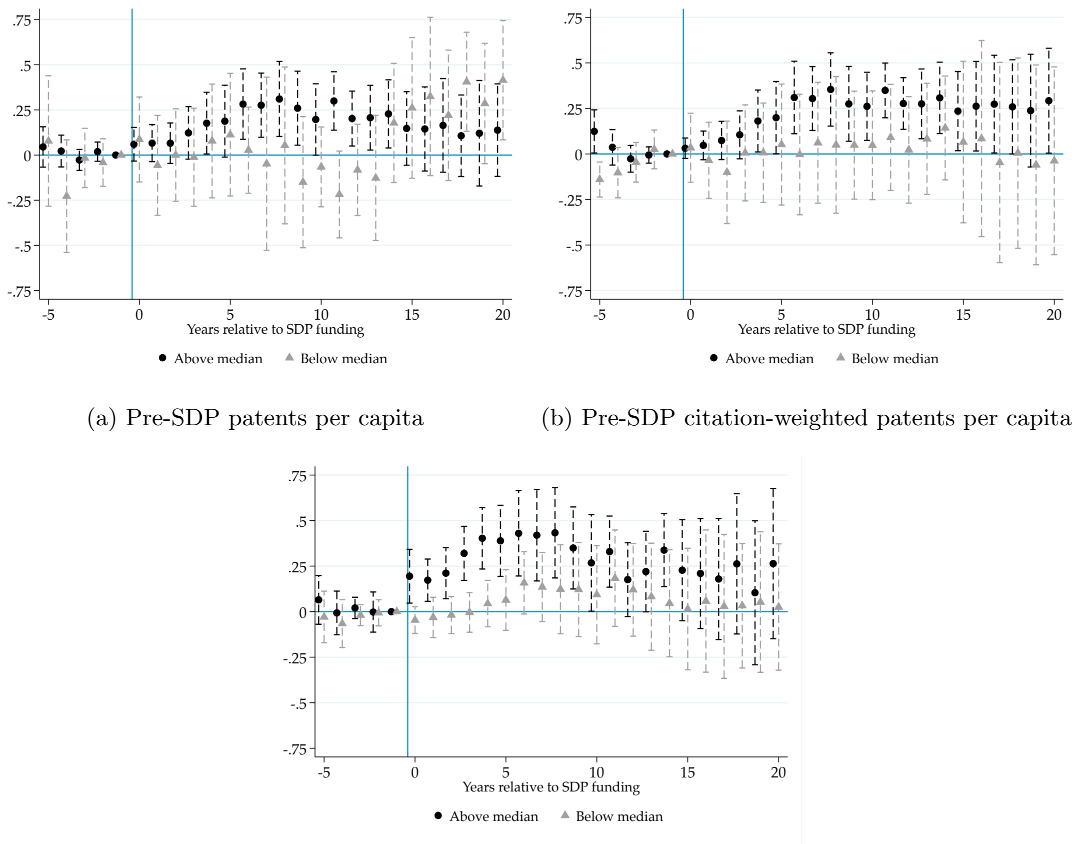
(e) Iteratively exclude commuting zones



(f) Random SDP funding assignment

Notes: In all panels, the dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regressions includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. Standard errors are clustered at the commuting zone level. In panels (a) to (e) the baseline period is $\tau = -1$ and the vertical bars represent 95% confidence intervals (except in panel (e), where they are denoted by dashed black lines). Panel (f) reports the distribution of coefficients from 1,000 permutation tests, where I estimate a two-period difference-in-differences randomly assigning a placebo SDP-funded status and timing (between 1965 and 1969) to all commuting zones not hosting any true funded university. The coefficient from my true estimate is denoted by a vertical blue line. All estimations by Poisson pseudo-maximum likelihood.

Figure 6: The NSF Science Development Program and Local Patenting: Effects by Commuting Zone Characteristics



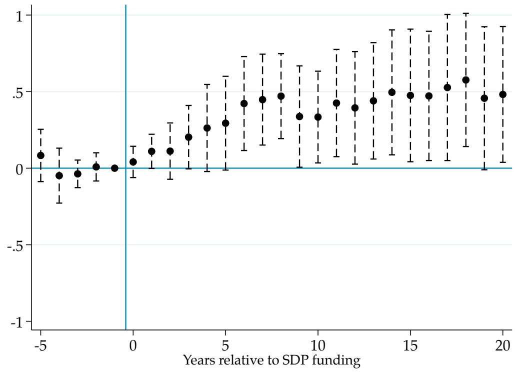
(a) Pre-SDP patents per capita

(b) Pre-SDP citation-weighted patents per capita

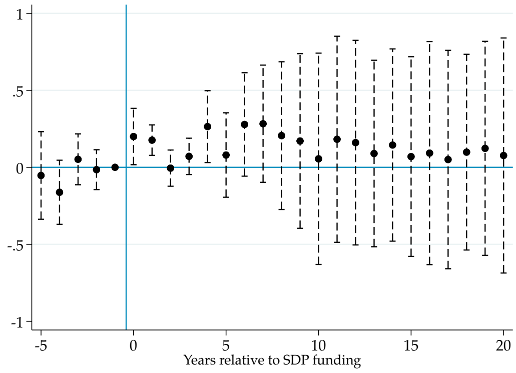
(c) Pre-SDP share of patents citing the scientific literature

Notes: The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. All regressions includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. Each panel presents separate estimates for commuting zones with either above- or below-median values of a given characteristic measured prior to the SDP introduction (i.e., between 1960 and 1964). Panel (a) focuses on commuting zones' total patents per capita; panel (b) on total citation-weighted patents per capita; and panel (c) on the share of total patents citing the scientific literature. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

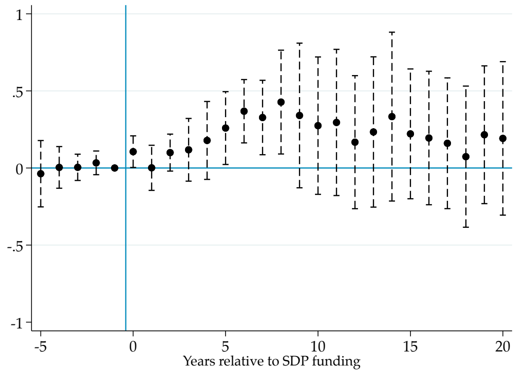
Figure 7: The NSF Science Development Program and Local Patenting: Effects by Distance to the Local University



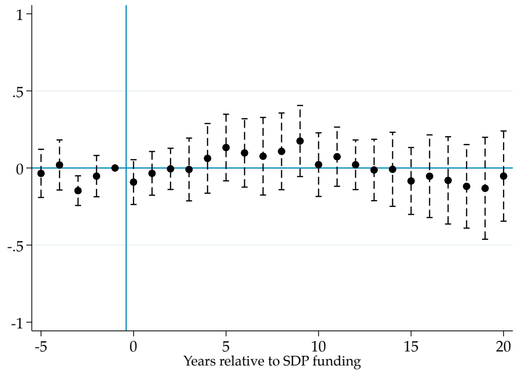
(a) Up to 5 miles



(b) 5-10 miles



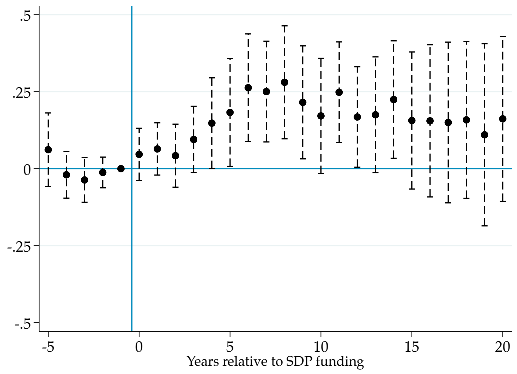
(c) 10-20 miles



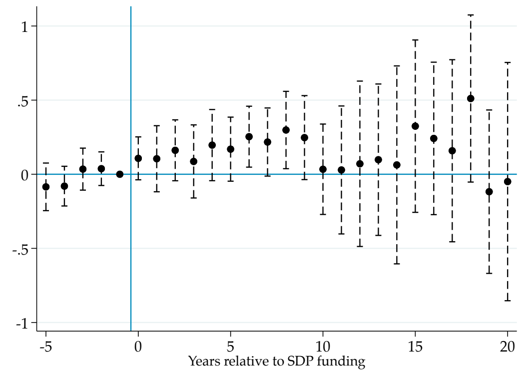
(d) Beyond 20 miles

Notes: The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. Each panel reports separate estimates for patent counts at different distance bands measured from the inventors' city and the nearest research university in the commuting zone. All regressions includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, plus controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

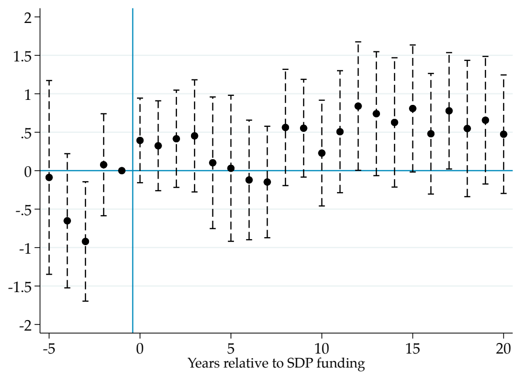
Figure 8: The NSF Science Development Program and Local Patenting: Effects by Assignee Type



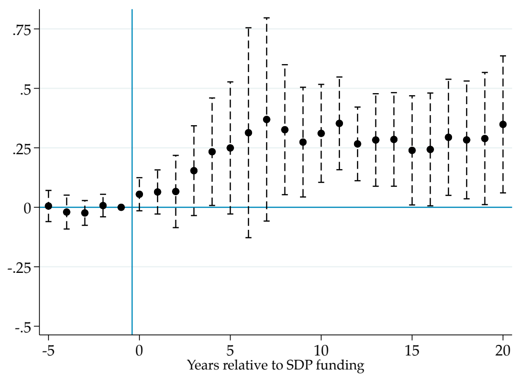
(a) Firms



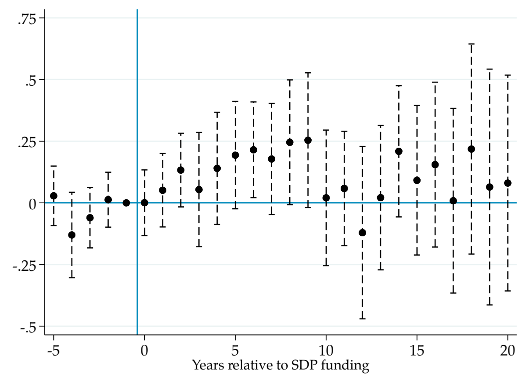
(b) Independent inventors



(c) Universities



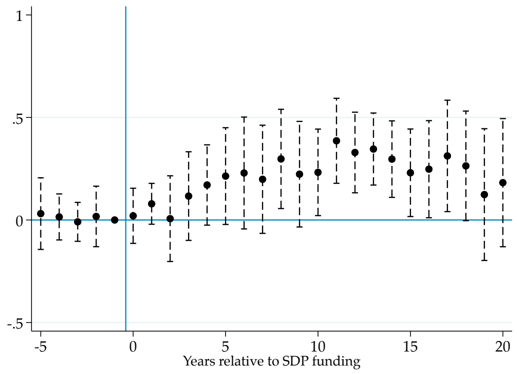
(d) Incumbents



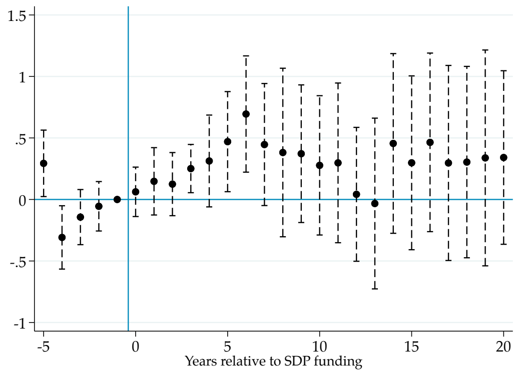
(e) New assignees

Notes: The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

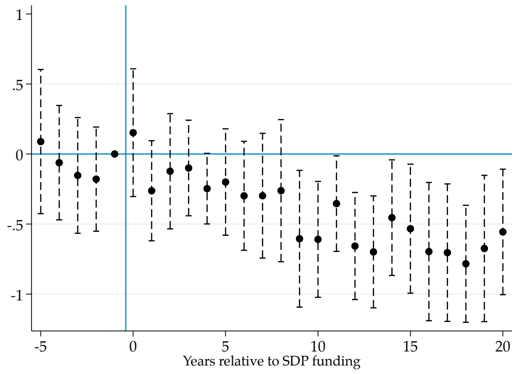
Figure 9: The NSF Science Development Program and Local Patenting: Effects by Technology Category



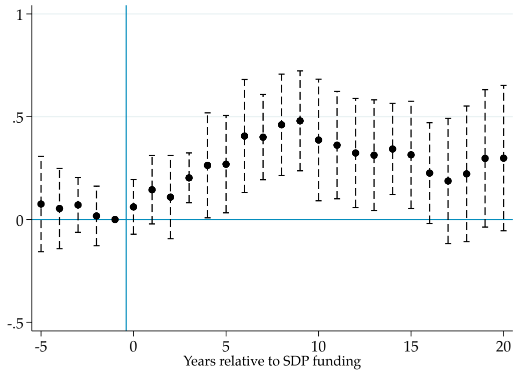
(a) Chemical



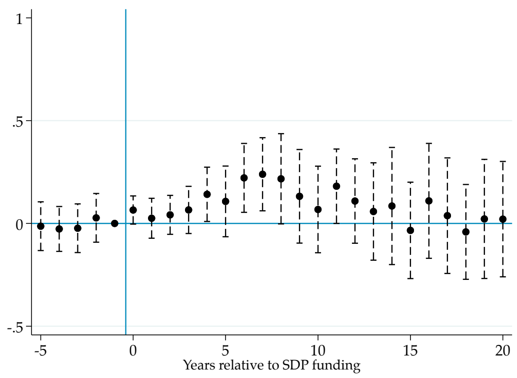
(b) Computers & Communications



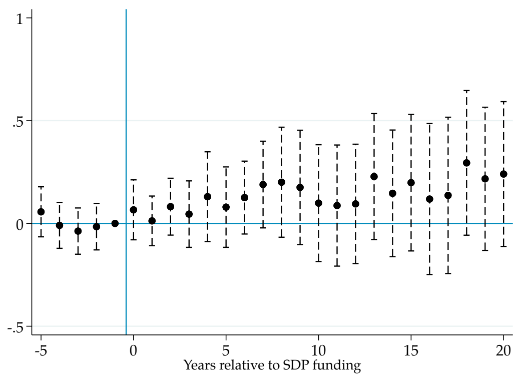
(c) Drugs & Medical



(d) Electrical & Electronics



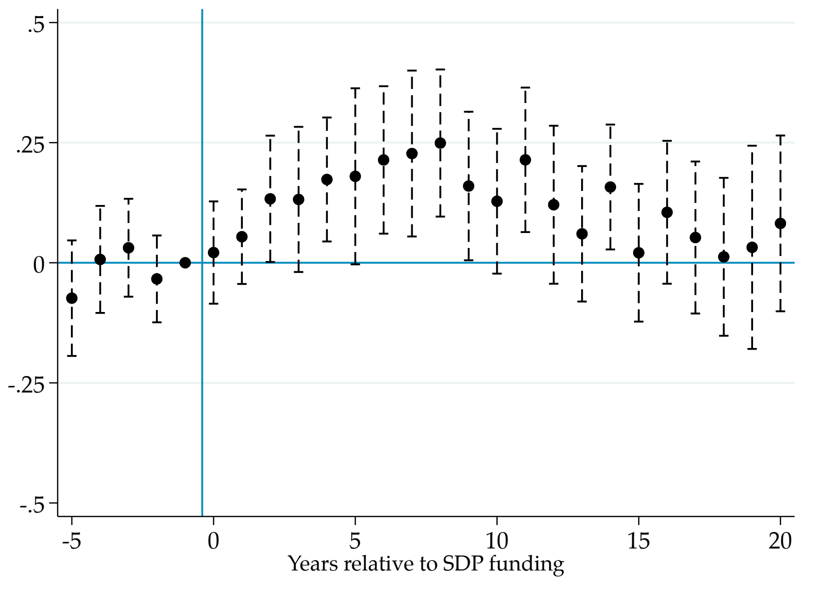
(e) Mechanical



(f) Others

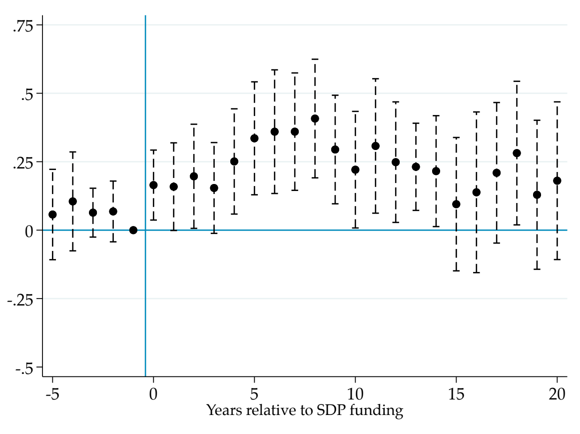
Notes: The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

Figure 10: The NSF Science Development Program and Local Patenting: Technology Field Exposure

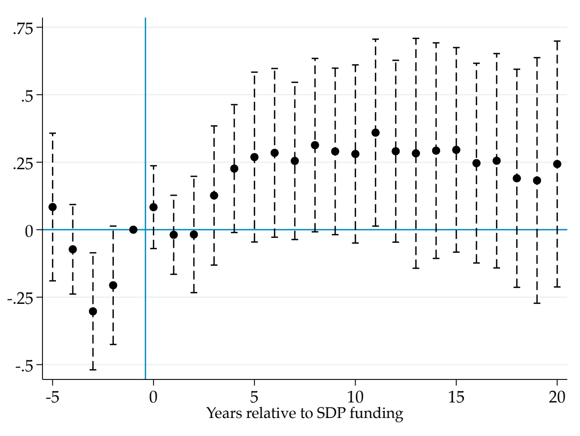


Notes: The figure plots β_τ estimates from a specification equivalent to equation (1), where $I_\tau \times SDP_u$ is further interacted with $Exposure_{ic}$, which measures the extent to which each technology field $\times$ commuting zone pair was exposed to the research activities of co-located universities prior to the Science Development Program introduction. The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

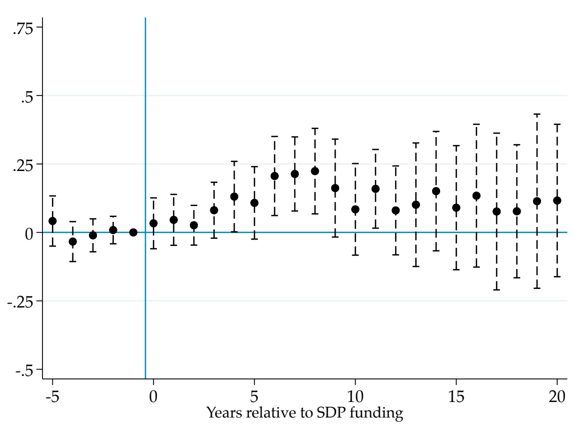
Figure 11: The NSF Science Development Program and Local Patenting: Reliance on the Scientific Literature



(a) Patents directly citing the scientific literature ($D = 1$)



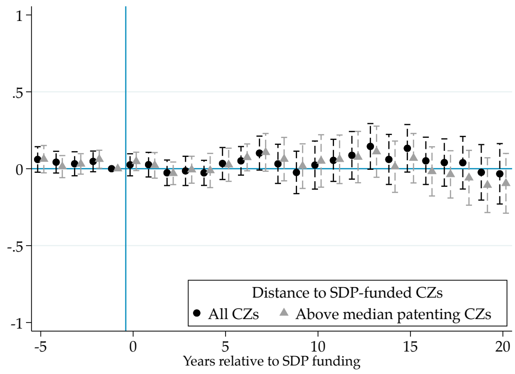
(b) Patents indirectly linked to the scientific literature ($D \in \{2, 3, 4\}$)



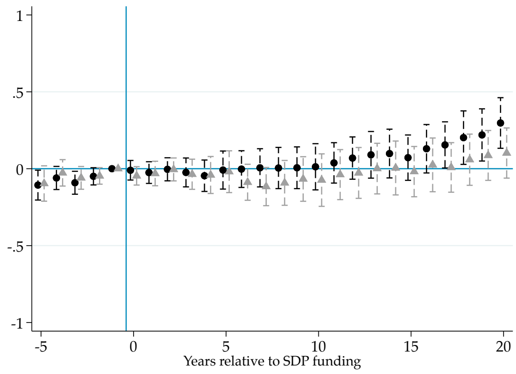
(c) Patents remotely linked to the scientific literature ($D \geq 5$)

Notes: The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

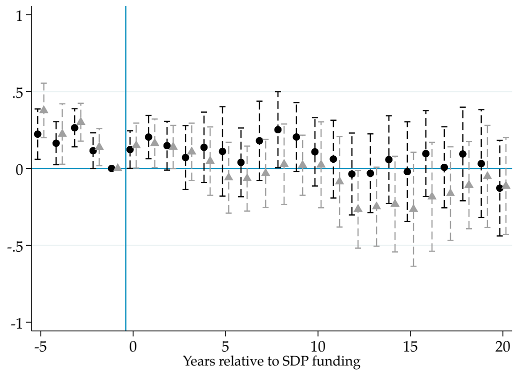
Figure 12: Patenting in Nearby Commuting Zones by Distance to Closest SDP-funded University



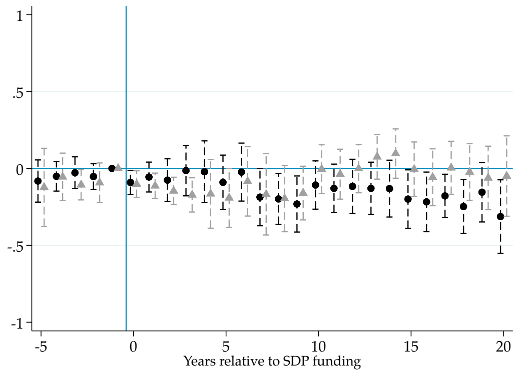
(a) Up to 100 miles



(b) 100-200 miles



(c) 200-300 miles



(d) Beyond 300 miles

Notes: Each panel plots difference-in-differences estimates comparing commuting zones lying within a specified radius of the nearest SDP-funded university with commuting zones at the same distance from the closest top-ranked university. Black markers use distances to the full set of SDP-funded universities, whereas gray markers consider only SDP-funded universities situated in commuting zones with above-median pre-SDP patenting. The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

Tables

Table 1: Universities Funded by the NSF Science Development Program

	Start Date	Award (Thous. $)	Award % of Total Federal Funding in 1963 and 1964
University of Oregon	5/1965	6,748	35
University of Colorado Boulder	6/1965	5,431	19
Rice University	6/1965	2,390	30
University of Rochester	6/1965	5,705	16
Washington University in St. Louis	6/1965	7,009	25
Case Western Reserve University	6/1965	9,160	29
University of Arizona	7/1965	7,227	49
University of Florida	7/1965	5,928	26
University of Virginia	7/1965	5,684	41
Louisiana State University	11/1965	6,216	40
Polytechnic Institute of Brooklyn	11/1965	4,542	-
University of Southern California	11/1965	7,473	31
North Carolina State University	5/1966	4,555	29
Purdue University	5/1966	3,900	14
Rutgers University	5/1966	4,708	28
Tulane University	5/1966	4,720	21
Duke University	12/1966	3,177	10
University of Texas at Austin	12/1966	6,600	15
Carnegie Mellon University	5/1967	4,789	35
University of Maryland, College Park	5/1967	4,703	15
University of North Carolina at Chapel Hill	5/1967	6,537	28
University of Notre Dame	5/1967	5,666	72
Vanderbilt University	5/1967	5,403	35
Indiana University Bloomington	7/1967	9,356	41
University of Georgia	8/1967	5,995	43
University of Iowa	8/1967	6,727	33
Florida State University	7/1968	6,020	48
Michigan State University	9/1968	5,487	23
University of Washington	9/1968	6,500	13
New York University	6/1969	6,160	12
University of Pittsburgh	9/1969	4,450	12

Notes: Data on SDP awards are drawn from the NSF Science Development Documentary Reports I–II (NSF 1977a,b). The right-most column shows each institution's SDP award as a percentage of total federal funds received in fiscal years 1963 and 1964, based on data from NSF (1967).

Table 2: Top-ranked Universities in the U.S. in 1964

	Cartter (1966) ranking			Total federal research funds in 1964 ranking	
	Biological Sciences	Physical Sciences	Engineering	National Science Foundation	General Accounting Office
A. Comparison group					
Brown University	28	29	11	60	-
California Institute of Technology	6	3	5	26	-
Columbia University	19	7	18	3	3
Cornell University	19	10	16	13	15
Harvard University	1	1	12	5	5
Johns Hopkins University	12	24	20	16	12
Massachusetts Institute of Technology	4	4	1	1	1
Ohio State University	39	20	29	19	19
Princeton University	33	5	14	28	20
Stanford University	7	6	4	4	4
University of California, Berkeley	2	2	2	8	9
University of California, Los Angeles	23	14	27	7	6
University of Chicago	37	9	-	9	7
University of Illinois Urbana-Champaign	10	11	7	6	8
University of Michigan	9	17	8	2	2
University of Minnesota	17	17	12	11	16
University of Pennsylvania	22	24	19	17	11
University of Wisconsin-Madison	7	13	9	10	10
B. SDP-funded					
New York University	27	8	24	14	13
Purdue University	31	24	10	29	-
University of Texas at Austin	18	22	21	15	17
C. SSDP-funded					
Northwestern University	40	23	15	27	-

Notes: The first three columns present each university's rank in the biological sciences, physical sciences, and engineering, based on evaluation scores from Cartter (1966). The universities listed are those ranked in the top twenty in at least one scientific domain and were categorized as "distinguished" or "strong" across all three domains—the highest-rated categories by Cartter (1966) and the only ones assigned scores. The University of Chicago is included because it could not be ranked in engineering disciplines (lacking engineering departments in the 1960s) and because it was explicitly mentioned by Page (1968) as an elite institution excluded from the SDP. The fourth and fifth columns provide each university's rank by total federal research funding received in 1964, based on data from the NSF (1967) and Comptroller General of the U.S. (1976), respectively. Since the Comptroller General of the U.S. (1976) ranked only the top twenty universities, few institutions miss this information.

Table 3: Commuting Zones Hosting SDP-funded and Top-ranked Universities

Commuting Zone Location	Universities
A. Hosting SDP-funded universities	
Athens, GA	University of Georgia
Austin, TX	University of Texas at Austin
Baton Rouge, LA	Louisiana State University
Bloomington, IN	Indiana University Bloomington
Boulder and Denver, CO	University of Colorado
Chapel Hill, Durham, and Raleigh, NC	Duke University; North Carolina State University; UNC Chapel Hill
Charlottesville, VA	University of Virginia
Cleveland, OH	Case Western Reserve University
Eugene, OR	University of Oregon
Gainesville, FL	University of Florida
Houston, TX	Rice University
Iowa City, IA	University of Iowa
Lafayette, IN	Purdue University
Lansing, MI	Michigan State University
Nashville, TN	Vanderbilt University
New Orleans, LA	Tulane University
Pittsburgh, PA	Carnegie Mellon University; University of Pittsburgh
Rochester, NY	University of Rochester
Seattle, WA	University of Washington
St. Louis, MO	Washington University in St. Louis
Tallahassee, FL	Florida State University
Tucson, AZ	University of Arizona
Washington, DC	University of Maryland, College Park
B. Hosting top-ranked universities (comparison group)	
Baltimore, MD	Johns Hopkins University
Berkeley and San Francisco, CA	University of California, Berkeley
Boston, MA	Harvard University; Massachusetts Institute of Technology
Champaign, IL	University of Illinois Urbana–Champaign
Chicago, IL	University of Chicago
Columbus, OH	Ohio State University
Detroit, MI	University of Michigan
Ithaca and Elmira, NY	Cornell University
Madison, WI	University of Wisconsin–Madison
Minneapolis, MN	University of Minnesota
Philadelphia, PA	University of Pennsylvania
Providence, RI	Brown University
Santa Clara and San Jose, CA	Stanford University
C. Hosting both SDP-funded and top-ranked universities	
Los Angeles, CA	Caltech; University of California, Los Angeles; University of Southern California
Newark, NJ	Princeton University; Rutgers University
New York City, NY	Columbia University; New York University; Polytechnic Institute of Brooklyn

Notes: Commuting zones are defined using the 1980 boundaries of Tolbert and Sizer (1996) and Autor and Dorn (2013).

Table 4: The NSF Science Development Program and Local Patenting: Two-Period Difference-in-Differences Results and Reliance on the Scientific Literature

				Patents linked to the scientific lit.			Share of patents linked to the scientific lit.		
	All patents	Firms' patents	Incumbents' patents	Direct $D=1$	Indirect $D \in \{2,3,4\}$	Remote $D \geq 5$	Direct $D=1$	Indirect $D \in \{2,3,4\}$	Remote $D \geq 5$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. All periods									
$Post_\tau \times SDP_c$	0.123*	0.157**	0.169**	0.148*	0.301**	0.093	0.013*	0.004	-0.017*
	(0.072)	(0.072)	(0.080)	(0.080)	(0.124)	(0.063)	(0.007)	(0.007)	(0.008)
Observations	32,215	31,463	31,311	29,458	27,197	31,633	23,316	23,316	23,316
Pseudo R^2 and R^2	0.842	0.832	0.825	0.703	0.692	0.805	0.371	0.297	0.411
Mean dep.var.	8.21	6.55	5.56	1.61	1.26	5.33	0.20	0.14	0.66
B. Until period 14									
$Post_\tau \times SDP_c$	0.108**	0.136**	0.145**	0.141**	0.278***	0.080*	0.016**	0.002	-0.018**
	(0.055)	(0.053)	(0.070)	(0.069)	(0.102)	(0.046)	(0.006)	(0.006)	(0.007)
Observations	24,421	23,671	23,809	21,819	19,645	24,059	17,887	17,887	17,887
Pseudo R^2 and R^2	0.852	0.841	0.837	0.689	0.685	0.819	0.334	0.316	0.383
Mean dep.var.	8.52	6.65	5.99	1.55	1.11	5.87	0.19	0.12	0.70
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
CZ × technology field FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Technology field × year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Institutional grants controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Standard errors are clustered at the commuting zone level and shown in parentheses. Estimations by Poisson pseudo-maximum-likelihood for patent counts and OLS for shares. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: The NSF Science Development Program and Local Patenting: Human Capital Creation and University-Industry Formal Collaborations

	Patents co-filed by local PhD grad. (1)	Patents not co-filed by local PhD grad. (2)	Patents co-filed by local academic (3)	Publications co-authored by local inventor and academic (4)
A. All periods				
$Post_\tau \times SDP_c$	0.273**	0.122*	-0.608	0.713
	(0.134)	(0.072)	(0.406)	(0.923)
Observations	14,299	32,173	6,894	1,618
Pseudo R^2	0.387	0.841	0.355	0.174
Mean dep. var	0.12	8.08	0.05	0.01
B. Until period 14				
$Post_\tau \times SDP_c$	0.259**	0.106*	-0.765	0.512
	(0.126)	(0.055)	(0.478)	(0.969)
Observations	9,450	24,381	3,128	565
Pseudo R^2	0.354	0.851	0.268	0.099
Mean dep. var	0.10	8.42	0.03	0.01
Year FEs	✓	✓	✓	✓
CZ × technology field FEs	✓	✓	✓	✓
Technology field × year FEs	✓	✓	✓	✓
Institutional grants controls	✓	✓	✓	✓

Notes: Standard errors are clustered at the commuting zone-level and shown in parentheses. Estimations by Poisson pseudo-maximum-likelihood. *** p<0.01, ** p<0.05, * p<0.1.

Appendix for

"Funding the Ivory Tower: The Effects of NSF Institutional Grants on Universities and Local Innovation"

by Gabriele Cristelli

June 2025

Contents

A Historical Context **2**

B Inventor Disambiguation **14**

C Additional Descriptive Evidence **21**

D Additional Estimates **25**

A. Historical Context

Table A1: Research Universities in the U.S. in 1964

University	Location
American University	Washington, D.C.
Boston University	Boston, MA
Brandeis University	Waltham, MA
Brown University	Providence, RI
Bryn Mawr College	Bryn Mawr, PA
California Institute of Technology	Pasadena, CA
Carnegie Institute of Technology	Pittsburgh, PA
Case Institute of Technology	Cleveland, OH
Catholic University of America	Washington, D.C.
Claremont Graduate University	Claremont, CA
Columbia University	New York, NY
Cornell University	Ithaca, NY
Duke University	Durham, NC
Emory University	Atlanta, GA
Florida State University	Tallahassee, FL
Fordham University	New York, NY
George Washington University	Washington, D.C.
Georgetown University	Washington, D.C.
Georgia Institute of Technology	Atlanta, GA
Harvard University	Cambridge, MA
Illinois Institute of Technology	Chicago, IL
Indiana University Bloomington	Bloomington, IN
Iowa State University	Ames, IA
Johns Hopkins University	Baltimore, MD
Kansas State University	Manhattan, KS
Lehigh University	Bethlehem, PA
Louisiana State University	Baton Rouge, LA
Loyola University Chicago	Chicago, IL
Massachusetts Institute of Technology	Cambridge, MA
Michigan State University	East Lansing, MI
New Mexico State University	Las Cruces, NM
The New School	New York, NY
New York University	New York, NY
North Carolina State University	Raleigh, NC
Northwestern University	Evanston, IL
Ohio State University	Columbus, OH
Oregon State University	Corvallis, OR
Pennsylvania State University	University Park, PA
Polytechnic Institute of Brooklyn	Brooklyn, NY
Princeton University	Princeton, NJ
Purdue University	West Lafayette, IN
Rensselaer Polytechnic Institute	Troy, NY
Rice University	Houston, TX
Rockefeller University	New York, NY
Rutgers, The State University of New Jersey	New Brunswick, NJ
Stony Brook University	Stony Brook, NY
St. John's University	Queens, NY
Stanford University	Stanford, CA
Syracuse University	Syracuse, NY
Temple University	Philadelphia, PA
Texas A&M University	College Station, TX
Tufts University	Medford, MA
Tulane University	New Orleans, LA
University at Buffalo, The State University of New York	Buffalo, NY
University of Alabama	Tuscaloosa, AL
University of Arizona	Tucson, AZ
University of Arkansas	Fayetteville, AR

University of California, Berkeley	Berkeley, CA
University of California, Davis	Davis, CA
University of California, Los Angeles	Los Angeles, CA
University of Chicago	Chicago, IL
University of Cincinnati	Cincinnati, OH
University of Colorado Boulder	Boulder, CO
University of Connecticut	Storrs, CT
University of Delaware	Newark, DE
University of Denver	Denver, CO
University of Florida	Gainesville, FL
University of Houston	Houston, TX
University of Illinois Urbana-Champaign	Champaign, IL
University of Iowa	Iowa City, IA
University of Kansas	Lawrence, KS
University of Kentucky	Lexington, KY
University of Maryland, College Park	College Park, MD
University of Massachusetts Amherst	Amherst, MA
University of Michigan	Ann Arbor, MI
University of Minnesota	Minneapolis, MN
University of Missouri	Columbia, MO
University of Nebraska-Lincoln	Lincoln, NE
University of New Mexico	Albuquerque, NM
University of North Carolina at Chapel Hill	Chapel Hill, NC
University of North Dakota	Grand Forks, ND
University of Notre Dame	Notre Dame, IN
University of Oklahoma	Norman, OK
University of Oregon	Eugene, OR
University of Pennsylvania	Philadelphia, PA
University of Pittsburgh	Pittsburgh, PA
University of Rochester	Rochester, NY
University of Southern California	Los Angeles, CA
University of Tennessee, Knoxville	Knoxville, TN
University of Texas at Austin	Austin, TX
University of Utah	Salt Lake City, UT
University of Virginia	Charlottesville, VA
University of Washington	Seattle, WA
University of Wisconsin-Madison	Madison, WI
University of Wyoming	Laramie, WY
Vanderbilt University	Nashville, TN
Virginia Tech	Blacksburg, VA
Washington State University	Pullman, WA
Washington University in St. Louis	St. Louis, MO
Wayne State University	Detroit, MI
West Virginia University	Morgantown, WV
Western Reserve University	Cleveland, OH
Yale University	New Haven, CT
Yeshiva University	New York, NY

Notes: List of the research universities surveyed and ranked by Cartter (1966).

Table A2: Total Federal Obligations to the 40 Universities Receiving the Largest Amount in 1964

Rank	University	Total federal funding (1964) (Thousand $)
1	Massachusetts Institute of Technology	70,681
2	University of Michigan	45,862
3	Columbia University	43,863
4	Stanford University	41,794
5	Harvard University	39,568
6	University of Illinois Urbana-Champaign	36,339
7	University of California, Los Angeles	35,301
8	University of California, Berkeley	34,623
9	University of Chicago	33,545
10	University of Wisconsin-Madison	31,111
11	University of Minnesota	28,544
12	University of Washington	27,839
13	Cornell University	27,023
14	New York University	26,468
15	University of Texas at Austin	26,086
16	Johns Hopkins University	25,821
17	University of Pennsylvania	25,319
18	Yale University	24,919
19	Ohio State University	20,084
20	Duke University	19,868
21	University of Maryland, College Park	18,439
22	University of Rochester	18,437
23	University of Pittsburgh	18,224
24	University of Colorado Boulder	15,218
25	Washington University in St. Louis	15,002
26	California Institute of Technology	14,940
27	Northwestern University	14,195
28	Princeton University	14,032
29	Purdue University	13,893
30	Western Reserve University	13,520
31	Howard University	13,509
32	Yeshiva University	13,497
33	University of Southern California	13,234
34	University of North Carolina at Chapel Hill	13,110
35	University of Florida	12,808
36	Pennsylvania State University	12,388
37	University of Tennessee, Knoxville	12,290
38	Indiana University Bloomington	12,015
39	Tulane University	11,944
40	Michigan State University	11,516

Notes: Data from NSF (1967).

Table A3: NASA Sustaining University Program for Training, Research, and Facilities Grants (1962-1971)

University	Grant (Thousand $)
Boston University	383
Brandeis University	375
Brown University	1,192
California Institute of Technology	2,635
Carnegie Mellon University	1,136
Case Western Reserve University	4,569
Catholic University of America	1,046
Columbia University	3,008
Cornell University	3,117
Duke University	1,359
Emory University	197
Florida State University	682
Fordham University	230
George Washington University	2,158
Georgetown University	498
Georgia Institute of Technology	4,347
Harvard University	276
Illinois Institute of Technology	977
Indiana University	933
Iowa State University	1,037
Johns Hopkins University	900
Kansas State University	1,037
Lehigh University	767
Louisiana State University	1,285
Loyola University Chicago	672
Massachusetts Institute of Technology	10,989
Michigan State University	895
New York University	1,866
North Carolina State University	1,024
Northeastern University	435
Northwestern University	1,570
Ohio State University	1,022
Oregon State University	759
Pennsylvania State University	2,702
Princeton University	1,383
Purdue University	4,148
Rensselaer Polytechnic Institute	3,145
Rice University	4,010
Stanford University	6,029
Syracuse University	1,932
Temple University	177
Texas A&M University	3,194
Tufts University	229
Tulane University	795
University of Alabama	3,769
University of Arizona	2,633
University of Arkansas at Fayetteville	725
University of California, Berkeley	6,675
University of California, Davis	706
University of California, Los Angeles	7,041
University of Chicago	3,336
University of Cincinnati	1,296
University of Colorado Boulder	1,722
University of Connecticut	622
University of Delaware	578
University of Denver	2,633
University of Florida	3,852
University of Houston	2,394
University of Illinois Urbana-Champaign	3,366
University of Iowa	1,736

University of Kansas	3,871
University of Kentucky	538
University of Maryland, College Park	5,612
University of Massachusetts Amherst	293
University of Michigan	2,825
University of Minnesota	6,134
University of Missouri	1,495
University of Nebraska–Lincoln	366
University of New Mexico	1,058
University of North Carolina at Chapel Hill	1,362
University of Notre Dame	760
University of Oklahoma	1,234
University of Pennsylvania	1,974
University of Pittsburgh	4,869
University of Rhode Island	337
University of Rochester	1,935
University of Southern California	2,996
University of Tennessee, Knoxville	1,975
University of Texas at Austin	901
University of Virginia	1,660
University of Washington	2,623
University of Wisconsin–Madison	7,515
Vanderbilt University	830
Virginia Tech	1,429
Washington State University	426
Washington University in St. Louis	3,411
Yale University	1,506
Yeshiva University	672

Notes: Total grants awarded by NASA under the Sustaining University Program for Training, Research, and Facilities (NSF 1977a).

Table A4: Department of Defense, project THEMIS Grants (1967-1971)

University	Grant (Thousand $)
Catholic University of America	2,527
Case Western Reserve University	600
Florida State University	3,293
Georgetown University	808
Georgia Institute of Technology	1,441
Illinois Institute of Technology	1,594
Indiana University	796
Iowa State University	1,738
Kansas State University	1,868
Lehigh University	2,218
Louisiana State University	1,706
Michigan State University	400
North Carolina State University	1,689
Oregon State University	1,435
Rensselaer Polytechnic Institute	2,497
Rice University	1,419
Texas A&M University	2,456
University at Buffalo, State University of New York	1,525
University of Alabama	605
University of Arizona	808
University of Connecticut	814
University of Florida	1,333
University of Houston	910
University of Iowa	1,575
University of Kansas	1,393
University of Kentucky	812
University of Massachusetts Amherst	840
University of Minnesota	2,630
University of Missouri	769
University of New Mexico	852
University of North Dakota	1,051
University of Notre Dame	925
University of Oklahoma	768
University of Tennessee, Knoxville	1,540
University of Vermont	864
University of Virginia	2,369
Vanderbilt University	902
Washington University in St. Louis	1,614
West Virginia University	809
Yeshiva University	780

Notes: Total grants awarded by the Department of Defense under Project THEMIS (NSF 1977a).

Table A5: National Institutes of Health, Health Science Advancement Awards (1966-1974)

University	Grant (Thousand $)
Cornell University	1,780
Duke University	2,516
Purdue University	2,542
Rice University	2,130
University of California, Davis	2,469
University of Colorado Boulder	2,655
University of Kansas	2,638
University of Oregon	2,097
University of Virginia	2,200
Vanderbilt University	2,491
Washington University in St. Louis	2,731

Notes: Total grants awarded by the National Institutes of Health under its Health Science Advancement Awards program (NSF 1977a).

Table A6: Ford Foundation, Special Program in Education ("Challenge") Grants (1960-1966)

University	Grant (Thousand $)
Brandeis University	6,000
Brown University	7,500
Columbia University	25,000
Duke University	8,000
Emory University	6,000
Illinois Institute of Technology	2,764
Johns Hopkins University	6,000
New York University	25,000
St. Louis University	4,000
Stanford University	25,000
Tulane University	6,000
University of Chicago	25,000
University of Denver	5,000
University of Notre Dame	12,000
University of Southern California	6,500
Vanderbilt University	15,000
Washington University in St. Louis	15,000

Notes: Total grants awarded by the Ford Foundation under its Special Program in Education (NSF 1977a).

Table A7: Cartter (1966) Biological Sciences Rankings

	Bacteriology/Microbiology	Biochemistry	Botany	Entomology	Pharmacology	Physiology	Psychology	Zoology	Average Score	Rank
Harvard University	4,04	4,63	4,25	-	4,35	4,52	4,58	4,56	4,42	1
University of California, Berkeley	4,51	4,54	4,63	4,56	3,52	3,90	4,35	4,67	4,34	2
Rockefeller University	4,31	4,34	-	-	-	4,18	-	4,31	4,29	3
Massachusetts Institute of Technology	-	4,25	-	-	-	-	-	-	4,25	4
University of Utah	-	-	-	-	3,98	-	-	-	3,98	5
California Institute of Technology	4,11	4,15	3,73	-	-	3,85	-	-	3,96	6
Stanford University	3,92	4,47	3,70	-	3,49	3,21	4,56	4,09	3,92	7
University of Wisconsin-Madison	-	4,30	4,02	3,76	-	3,42	3,97	3,81	3,92	7
University of Michigan	3,32	3,47	4,17	-	4,11	3,69	4,40	3,80	3,85	9
University of Illinois Urbana-Champaign	4,21	3,75	3,39	3,93	3,44	3,49	4,08	3,60	3,74	10
University of Iowa	-	-	-	-	3,75	-	3,66	-	3,71	11
Yale University	3,52	3,37	3,30	-	4,02	3,29	4,35	3,96	3,69	12
Johns Hopkins University	3,63	3,93	-	-	3,12	3,97	3,44	4,02	3,69	12
Emory University	-	-	-	-	3,68	-	-	-	3,68	14
Indiana University Bloomington	3,39	-	3,81	-	-	3,24	3,62	3,95	3,60	15
Brandeis University	3,33	3,85	-	-	-	-	-	-	3,59	16
University of Minnesota	3,63	-	3,16	3,61	3,55	3,73	3,98	3,21	3,55	17
University of Texas at Austin	3,51	-	3,75	-	-	-	-	3,31	3,52	18
Columbia University	3,19	3,70	-	-	-	3,62	3,54	3,50	3,51	19
Western Reserve University	3,65	3,63	-	-	-	3,39	-	3,36	3,51	19
Duke University	-	3,54	3,80	-	-	3,41	3,34	3,44	3,51	19
University of Pennsylvania	3,41	3,42	3,01	-	4,07	3,70	3,63	3,29	3,50	22
University of California, Los Angeles	3,04	3,37	3,58	-	-	3,54	3,58	3,85	3,49	23
University of California, Davis	3,47	3,16	3,92	3,32	-	-	-	-	3,47	24
University of Washington	3,55	3,58	-	-	3,02	3,96	-	3,20	3,46	25
Cornell University	3,06	3,19	3,46	3,87	3,60	3,52	3,33	3,48	3,44	26
New York University	3,26	3,60	-	-	-	-	-	-	3,43	27
Brown University	-	-	-	-	-	-	3,73	3,09	3,41	28
Rutgers, The State University of New Jersey	3,41	-	-	-	-	-	-	-	3,41	28
University of Rochester	-	-	-	-	3,26	3,55	-	-	3,41	28
University of Kansas	-	-	-	3,71	3,05	-	-	-	3,38	31
Purdue University	3,57	-	3,41	-	-	-	-	3,15	3,38	31
University at Buffalo	-	-	-	-	-	3,37	-	-	3,37	33
Princeton University	-	3,21	-	-	-	3,50	3,14	3,62	3,37	33
Yeshiva University	3,04	3,42	-	-	3,72	3,25	-	-	3,36	35
Washington University in St. Louis	-	3,35	-	-	3,61	3,11	-	-	3,36	35
University of Chicago	3,32	3,40	3,02	-	3,10	3,57	3,37	3,70	3,35	37
Tufts University	-	3,33	-	-	-	-	-	-	3,33	38
Ohio State University	-	-	-	3,38	-	-	3,24	-	3,31	39
Northwestern University	-	-	-	-	-	3,05	3,43	3,18	3,22	40
Michigan State University	-	-	3,41	-	-	-	3,02	-	3,22	40
University of Pittsburgh	3,07	3,21	-	-	-	-	-	-	3,14	42
North Carolina State University	-	-	3,26	-	-	-	3,01	-	3,14	42
Claremont Graduate University	-	-	3,08	-	-	-	-	-	3,08	44
Iowa State University	-	-	-	3,06	-	-	-	-	3,06	45
Kansas State University	-	-	-	3,04	-	-	-	-	3,04	46
University of Oregon	-	-	-	-	-	3,02	-	-	3,02	47
Vanderbilt University	-	-	-	-	3,02	-	-	-	3,02	47

Notes: Scores are derived from Cartter's (1966) evaluation of universities' "quality of graduate faculty," based on a 1964 survey of U.S. scholars.

Table A8: Cartter (1966) Physical Sciences Rankings

	Astronomy	Chemistry	Geology	Mathematics	Physics	Average Score	Rank
Harvard University	4,08	4,95	4,45	4,85	4,71	4,61	1
University of California, Berkeley	4,10	4,68	4,38	4,81	4,78	4,55	2
California Institute of Technology	4,81	4,72	4,38	3,66	4,77	4,47	3
Massachusetts Institute of Technology	-	4,55	3,96	4,39	4,45	4,34	4
Princeton University	4,62	3,67	3,98	4,79	4,60	4,33	5
Stanford University	-	4,32	3,94	4,19	4,47	4,23	6
Columbia University	-	4,00	4,28	4,02	4,32	4,16	7
New York University	-	-	-	4,10	-	4,10	8
University of Chicago	4,12	3,91	3,30	4,60	4,00	3,99	9
Cornell University	-	3,77	-	3,70	4,07	3,85	10
University of Illinois Urbana-Champaign	-	4,13	3,32	3,74	4,10	3,82	11
Yale University	3,39	3,76	3,76	4,13	3,77	3,76	12
University of Wisconsin-Madison	3,25	4,00	3,45	3,88	3,69	3,65	13
University of California, Los Angeles	-	3,92	3,67	3,47	3,12	3,55	14
Pennsylvania State University	-	3,13	3,82	-	-	3,48	15
University of Rochester	-	-	-	-	3,46	3,46	16
University of Michigan	3,20	3,25	3,32	3,86	3,46	3,42	17
University of Minnesota	-	3,51	3,37	3,48	3,31	3,42	17
Iowa State University	-	3,40	-	-	-	3,40	19
Ohio State University	-	3,37	-	-	-	3,37	20
University of Maryland, College Park	-	-	-	-	3,35	3,35	21
University of Texas at Austin	-	3,14	3,50	-	-	3,32	22
Northwestern University	-	3,52	3,19	3,21	-	3,31	23
University of Pennsylvania	-	-	-	3,15	3,37	3,26	24
Johns Hopkins University	-	3,17	3,50	3,23	3,12	3,26	24
Purdue University	-	3,37	-	3,14	-	3,26	24
University of Washington	-	3,18	-	3,39	3,16	3,24	27
Brandeis University	-	-	-	3,24	-	3,24	27
Brown University	-	3,02	-	3,36	-	3,19	29
University of Virginia	-	-	-	3,13	-	3,13	30
Indiana University Bloomington	-	3,24	-	3,02	-	3,13	30
Carnegie Institute of Technology	-	-	-	-	3,09	3,09	32
Rice University	-	3,06	3,12	-	-	3,09	33
Polytechnic Institute of Brooklyn	-	3,08	-	-	-	3,08	34
Florida State University	-	3,06	-	-	-	3,06	35

Notes: Scores are derived from Cartter's (1966) evaluation of universities' "quality of graduate faculty," based on a 1964 survey of U.S. scholars.

Table A9: Cartter (1966) Engineering Rankings

	Chemical Engineering	Civil Engineering	Electrical Engineering	Mechanical Engineering	Average Score	Rank
Massachusetts Institute of Technology	4,36	4,17	4,78	4,61	4,48	1
University of California, Berkeley	4,24	4,52	4,38	3,83	4,24	2
University of Delaware	4,13	-	-	-	4,13	3
Stanford University	3,42	3,86	4,68	4,14	4,03	4
California Institute of Technology	3,53	4,09	3,98	4,20	3,95	5
Polytechnic Institute of Brooklyn	-	-	3,94	-	3,94	6
University of Illinois Urbana-Champaign	3,80	4,40	4,13	3,33	3,92	7
University of Michigan	4,13	3,62	3,68	3,50	3,73	8
University of Wisconsin-Madison	4,43	3,22	3,34	-	3,66	9
Purdue University	-	3,70	3,51	3,65	3,62	10
Brown University	-	-	-	3,58	3,58	11
University of Minnesota	4,25	3,08	3,24	3,72	3,57	12
Harvard University	-	-	3,60	3,54	3,57	12
Princeton University	4,25	-	3,24	3,19	3,56	14
Northwestern University	3,42	3,41	-	3,27	3,37	15
Cornell University	-	3,42	3,25	3,32	3,33	16
Carnegie Institute of Technology	3,33	-	3,33	-	3,33	16
Columbia University	-	3,37	3,34	3,26	3,32	18
University of Pennsylvania	-	-	3,29	-	3,29	19
Johns Hopkins University	-	-	3,28	3,23	3,26	20
University of Texas at Austin	3,35	3,14	-	-	3,25	21
Rice University	3,18	-	-	-	3,18	22
Syracuse University	-	-	3,16	-	3,16	23
New York University	-	-	3,13	-	3,13	24
Lehigh University	-	3,12	-	-	3,12	25
Case Institute of Technology	-	-	3,02	3,20	3,11	26
University of California, Los Angeles	-	-	3,08	3,12	3,10	27
University of Washington	3,05	3,12	-	-	3,09	28
Ohio State University	-	-	3,04	-	3,04	29

Notes: Scores are derived from Cartter's (1966) evaluation of universities' "quality of graduate faculty," based on a 1964 survey of U.S. scholars.

Figure A1: Cartter's (1966) List of Scientific Disciplines by Domain

FIELDS OF STUDY

HUMANITIES

Classics	German
English	Philosophy
French	Spanish

SOCIAL SCIENCES

Anthropology	History
Economics	Political Science
Geography	Sociology

BIOLOGICAL SCIENCES

Bacteriology/ Microbiology	Pharmacology
Biochemistry	Physiology
Botany	Psychology
Entomology	Zoology

PHYSICAL SCIENCES

Astronomy	Mathematics
Chemistry	Physics
Geology	

ENGINEERING

Chemical Engineering	Electrical Engineering
Civil Engineering	Mechanical Engineering

Notes: Reproduced from Cartter (1966, p. 20).

Figure A2: Examples of Cartter's (1966) Rankings

BACTERIOLOGY/MICROBIOLOGY

Leading Departments, by Rated Quality of Graduate Faculty

Institution†	All Respondents Rank	All Respondents Score	Chairmen Rank	Chairmen Score	Senior Scholars Rank	Senior Scholars Score	Junior Scholars Rank	Junior Scholars Score
"Distinguished"								
California, Berkeley	1	4.51	1	4.62	1	4.36	1	4.58
Rockefeller Inst.**	2	4.31	2	4.48	2	4.29	2*	4.20
Illinois	3	4.21	3	4.45	4*	4.10	4	4.14
Wisconsin**	4	4.14	4	4.25	3	4.25	6	3.92
Cal. Tech.**	5	4.11	5	4.08	4*	4.10	5	4.11
Harvard**	6	4.04	7	3.94	6*	3.96	2*	4.20
"Strong"								
Stanford**	7	3.92	6	3.96	6*	3.96	7	3.82
Western Reserve	8	3.65	8*	3.81	11	3.60	10	3.58
Johns Hopkins**	9*	3.63	16*	3.46	10	3.61	8	3.76
Minnesota**	9*	3.63	8*	3.81	8	3.68	14	3.40
Purdue**	11	3.57	19	3.44	14*	3.50	9	3.74
Washington (Seattle)	12	3.55	10	3.76	17	3.46	12	3.45
Yale	13	3.52	11	3.61	13	3.51	13	3.44
Texas**	14	3.51	12	3.56	9	3.67	17*	3.27
California, Davis**	15	3.47	16*	3.46	19	3.41	11	3.52
Pennsylvania	16*	3.41	14*	3.50	14*	3.50	20	3.22
Rutgers**	16*	3.41	14*	3.50	18	3.43	16	3.31
Indiana**	18	3.39	18	3.45	21	3.36	15	3.38
Brandeis**	19	3.33	27	3.09	12	3.55	19	3.26
Chicago	20*	3.32	13	3.52	22*	3.21	17*	3.27
Michigan	20*	3.32	21	3.40	20	3.39	21	3.17
N.Y.U.**	22	3.26	23	3.29	16	3.47	24	3.02
Columbia	23	3.19	22	3.32	22*	3.21	23	3.07
Pittsburgh**	24	3.07	20	3.41	26	3.04	28	2.87
Cornell	25	3.06	29	3.00	24*	3.20	25	2.93
U.C.L.A.**	26*	3.04	24	3.21	30	2.88	22	3.10
Yeshiva**a	26*	3.04	26	3.10	24*	3.20	26*	2.90

"Good" (17 departments arranged alphabetically):
Cincinnati**, Duke**, Emory, Iowa (Iowa City), Iowa State (Ames), Maryland**, Michigan State**, North Carolina**, Oregon**a, Oregon State**, Penn State, Rochester**, Southern California**, Tufts a, Tulane, Washington (St. Louis), Washington State**

"Adequate plus" (18 departments arranged alphabetically):
Buffalo**, Colorado**, Florida**, George Washington, Kansas**, Kansas State, Kentucky, Massachusetts, Nebraska**, Northwestern, Ohio State, Oklahoma**, St. Louis, Syracuse**, Temple**, Tennessee**, Utah, Vanderbilt

† These tabulations and those for the effectiveness of the graduate program are based on ratings of 75 institutions. See page 18 for further explanation.
* Rank and score shared with another department.
** Study offered in a department other than department of bacteriology or microbiology or offered in more than one department.
a Doctorates reported in "biology" or "biological sciences, other." No more than half the respondents rated the department.

Profile of Respondents

Respondents	Number	Average Age	Average Number of Publications Since Highest Degree: Books	Articles	Professional Meetings Attended in Last Four Years: Regional	National
All respondents	156	44.7	.6	35	6.4	6.4
Chairmen	37	50.6	.8	49	6.6	7.4
Senior scholars	62	48.0	.9	46	7.5	6.3
Junior scholars	57	36.9	.2	13	5.1	6.0

Respondents' Division of Time for Professional Activities
(in percents)

Respondents	Instruction: Under-graduate	Graduate	Total	Research and Writing	Adminis-tration	Other Profes-sional	Other
All respondents	22	26	48	27	17	6	2
Chairmen	16	20	36	18	35	8	3
Senior scholars	25	27	52	27	13	6	1
Junior scholars	23	28	51	33	9	4	3

44

PHYSICS

Leading Departments, by Rated Quality of Graduate Faculty

Institution†	All Respondents Rank	All Respondents Score	Chairmen Rank	Chairmen Score	Senior Scholars Rank	Senior Scholars Score	Junior Scholars Rank	Junior Scholars Score
"Distinguished"								
California, Berkeley	1	4.78	1	4.86	1*	4.78	1*	4.73
Cal. Tech.	2	4.77	2	4.79	1*	4.78	1*	4.73
Harvard	3	4.71	3	4.74	3	4.75	4	4.61
Princeton	4	4.60	4	4.62	4	4.56	3	4.63
Stanford	5	4.47	5	4.51	6	4.50	5	4.38
M.I.T.	6	4.45	6	4.48	5	4.55	7	4.28
Columbia	7	4.32	7	4.45	7	4.26	6	4.29
Illinois	8	4.10	8	4.18	9	4.14	8	4.00
Cornell	9	4.07	9	4.14	10	4.10	9	3.96
"Strong"								
Chicago	10	4.00	10	3.95	8	4.16	10	3.81
Yale	11	3.77	11	3.87	11	3.73	11	3.74
Wisconsin	12	3.69	12	3.77	12	3.70	12	3.59
Michigan	13*	3.46	14	3.48	13	3.52	14	3.32
Rochester	13*	3.46	16	3.42	14	3.48	13	3.45
Pennsylvania	15	3.37	15	3.45	15*	3.39	16	3.26
Maryland	16	3.35	17	3.35	15*	3.39	15	3.29
Minnesota	17	3.31	13	3.50	17	3.30	18	3.13
Washington (Seattle)	18	3.16	18	3.20	19	3.18	20	3.06
Johns Hopkins	19*	3.12	20*	3.02	18	3.19	19	3.08
U.C.L.A.	19*	3.12	20*	3.02	21	3.14	17	3.15
Carnegie Tech.	21	3.09	19	3.05	20	3.16	21*	3.02

"Good" (17 departments arranged alphabetically):
Brandeis, Brown, Case, Colorado, Duke, Indiana, Iowa State (Ames), Michigan State, N.Y.U., Northwestern, Ohio State, Pittsburgh, Purdue, Rice, Rutgers, Syracuse, Washington (St. Louis)

"Adequate plus" (11 departments arranged alphabetically):
Brooklyn Polytech., Florida State, Iowa (Iowa City), North Carolina, Notre Dame, Oregon, Penn State, Rensselaer, Southern California, Texas, Virginia

† These tabulations and those for the effectiveness of the graduate program are based on ratings of the 86 institutions in the survey universe that reported the award of one or more doctorates in physics from July 1952 through June 1962.
* Rank and score shared with another department.

Profile of Respondents

Respondents	Number	Average Age	Average Number of Publications Since Highest Degree: Books	Articles	Professional Meetings Attended in Last Four Years: Regional	National
All respondents	190	42.0	.5	28	3.3	8.3
Chairmen	43	47.5	.4	37	2.5	8.5
Senior scholars	82	46.6	.8	37	3.9	9.4
Junior scholars	65	33.1	.1	10	2.9	6.9

Respondents' Division of Time for Professional Activities
(in percents)

Respondents	Instruction: Under-graduate	Graduate	Total	Research and Writing	Adminis-tration	Other Profes-sional	Other
All respondents	19	28	47	23	22	7	1
Chairmen	13	19	32	14	47	7	1
Senior scholars	19	29	48	23	20	8	2
Junior scholars	22	35	57	29	8	5	—

68

ELECTRICAL ENGINEERING

Leading Departments, by Rated Quality of Graduate Faculty

Institution†	All Respondents Rank	All Respondents Score	Chairmen Rank	Chairmen Score	Senior Scholars Rank	Senior Scholars Score	Junior Scholars Rank	Junior Scholars Score
"Distinguished"								
M.I.T.	1	4.78	1	4.96	1	4.85	1	4.60
Stanford	2	4.68	2	4.76	2	4.75	2	4.55
California, Berkeley	3	4.38	3	4.62	3	4.34	3	4.25
Illinois	4	4.13	4	4.17	4	4.14	4	4.07
"Strong"								
Cal. Tech.	5	3.98	5	4.04	6	3.90	5	4.02
Brooklyn Polytech.	6	3.94	7	3.76	5	4.11	6	3.87
Michigan	7	3.68	8	3.70	7	3.76	9	3.57
Harvard	8	3.60	6	3.81	8	3.44	7	3.60
Purdue	9	3.51	9	3.52	9	3.43	8	3.58
Columbia	10*	3.34	18	3.20	12	3.28	10	3.56
Wisconsin	10*	3.34	14	3.30	10	3.39	15	3.28
Carnegie Tech.	12	3.33	11	3.47	14*	3.21	12	3.37
Pennsylvania	13	3.29	19	3.19	14*	3.21	11	3.48
Johns Hopkins	14	3.28	13	3.41	13	3.25	16	3.20
Cornell	15	3.25	16	3.26	11	3.34	21	3.10
Minnesota	16*	3.24	10	3.50	17	3.16	17	3.17
Princeton	16*	3.24	12	3.45	14*	3.21	20	3.13
Syracuse	18	3.16	21	3.15	19	3.02	13	3.34
N.Y.U.	19	3.13	17	3.21	18	3.07	18	3.15
U.C.L.A.	20	3.08	22	3.13	22	2.89	14	3.29
Ohio State	21	3.04	15	3.27	21	2.91	23*	3.00
Case	22	3.02	20	3.18	20	2.95	23*	3.00

"Good" (9 departments arranged alphabetically):
Brown, Iowa State (Ames), Michigan State, Northwestern, Penn State, Rensselaer, Texas, Washington (Seattle), Yale

'Adequate plus" (13 departments arranged alphabetically):
Arizona, Colorado, Florida, Georgia Tech., Ill. Inst. of Tech., Lehigh, Maryland, New Mexico, Oklahoma State, Pittsburgh, Southern California, Utah, Washington (St. Louis)

† These tabulations and those for the effectiveness of the graduate program are based on ratings of the 54 institutions in the survey universe that reported the award of one or more doctorates in electrical engineering from July 1952 through June 1962.
* Rank and score shared with another department.

Profile of Respondents

Respondents	Number	Average Age	Average Number of Publications Since Highest Degree: Books	Articles	Professional Meetings Attended in Last Four Years: Regional	National
All respondents	128	43.1	1.3	13	5.1	6.7
Chairmen	25	49.9	1.4	16	5.2	8.5
Senior scholars	53	46.9	2.2	19	6.0	7.8
Junior scholars	50	35.7	.2	6	4.1	4.7

Respondents' Division of Time for Professional Activities
(in percents)

Respondents	Instruction: Under-graduate	Graduate	Total	Research and Writing	Adminis-tration	Other Profes-sional	Other
All respondents	21	27	48	16	26	9	2
Chairmen	9	12	21	10	57	10	2
Senior scholars	20	29	49	18	22	8	3
Junior scholars	28	32	60	17	14	9	—

74

(a) Bacteriology/Microbiology (b) Physics (c) Electrical Engineering

Notes: The panels present examples of Cartter's (1966) university rankings for the "quality of graduate faculty," with one example for each broader scientific domain. Scores are assigned exclusively to institutions categorized as "distinguished" or "strong." In Table A7, Table A8, and Table A9 I use scores based on the evaluation of all survey respondents, that is, (department) "Chairmen", "Senior scholars", and "Junior scholars."

B. Inventor Disambiguation

Patent data do not natively provide unique identifiers for the same inventor across different patents. That is because intellectual property offices such as the USPTO do not require inventors to submit their information using standardized identifiers during the patent application process. Without them, researchers interested in studying individual inventors' careers face name ambiguity issues, unable to determine whether inventors sharing the same name correspond to the same person or to homonyms. Over the past two decades, scholars reliant on patent data developed disambiguation algorithms to obtain consistent inventor identifiers (e.g., Trajtenberg et al. 2006; Li et al. 2014; Pezzoni et al. 2014). These algorithms group inventors with the same name and leverage auxiliary information from patent documents, such as inventors' residential locations, co-inventors, patent assignees, and technology classes to identify inventor records referring to the same individual. In 2015, the USPTO adopted an inventor disambiguation algorithm developed by Monath et al. (2021), distributing the unique inventor identifiers generated by the algorithm in USPTO's database "PatentsView".[28]

Despite being the most comprehensive source of USTO patent data, the PatentsView database includes only "modern" patents, those granted since 1975. That does not suit the purposes of my project, which relies also on "historical" patents, those granted before 1975 (Andrews 2021a). To address this challenge, I disambiguate all inventors filing a patent at the USPTO and residing in the US between 1920 and 2015. I adopt Monath et al.'s 2021 algorithm, a machine learning procedure combining supervised classification and hierarchical agglomerative clustering (HAC), and extend it to disambiguate inventors listed on both modern and historical patents.

I start by creating a training set of inventors from modern patents disambiguated by Monath et al. (2021). Similar to Akcigit et al. (2022)—who also generate unique identifiers for historical inventors using a training set of disambiguated modern inventors—I rely on two assumptions.[29] First, I assume that Monath et al.'s 2021 disambiguation is generally correct. I find this assumption reasonable: evaluations of Monath et al.'s 2021 algorithm against manually labeled inventor data reveal minimal errors, outperforming previous disambiguation algorithms, making it the best inventor disambiguation currently available to patent data researchers. Second, I assume that the features enabling the disambiguation of modern inventors have the same predictive power for the disambiguation of inventors from historical inventors.

My training set is based on a random extraction of 2 million disambiguated inventor mentions (i.e., unique inventor-patent instances). I group inventor mentions into "canopies" (McCallum

[28] The source code for PatentsView disambiguation process based on Monath et al.'s 2021 algorithm can be found at `https://github.com/PatentsView/PatentsView-Disambiguation`.

[29] Akcigit et al. (2022) use inventor records disambiguated by Li et al. (2014).

et al. 2000), that is, clusters of inventor mentions sharing the same last name and the first two initials of their first name. Each canopy contains a set of inventor pairs which may refer to the same individual. I obtain a training set of 25,644,056 inventor pairs. I also create a development set of 64,725 inventor pairs, based on a subsequent random extraction of 100,000 disambiguated inventor mentions not found in the training set.

Next, I train a logistic classifier to predict whether each inventor pair refers to the same individual. Following Monath et al. (2021), my features include:

- Same first name (binary)
- Jaro-Winkler similarity score for the first name
- Same middle name or no middle name for both records (binary)
- Same city of residence (binary)
- Same county of residence (binary)
- Same state of residence (binary
- Same assignee string (binary)
- Assignee string Jaro-Winkler similarity > 0.9 (binary)
- At least one common co-inventor (binary)
- Same NBER technology category, USPC technology class, and USPC technology sub-class (all binary)[30]

In order to fine-tune my algorithm, I use the predicted scores generated by my logistic classifier and create a set of distance matrices, one for each canopy in the development set. Each score ranges between 0 (minimum distance) and 1 (maximum distance). I set maximum distance scores for inventor pairs with a different middle name or whose patents were filed more than 40 years apart.

I then apply hierarchical agglomerative clustering (HAC), a method that iteratively clusters inventors within the same canopy based on their respective distance matrix scores. I determine the optimal distance threshold to halt the HAC process (and finalize the set of disambiguated inventors) as the score jointly maximizing precision and recall. In my context, precision quantifies the algorithm's accuracy in grouping inventors correctly. It is calculated as the ratio of "true positives" (correctly clustered inventor pairs) to the total number of true positives and "false negatives" (incorrectly clustered inventor pairs). Recall, on the other hand, measures the algorithm's completeness in identifying all inventors that should be grouped together. It is calculated as the ratio of true positives to the sum of true positives and "false negatives"

[30] The primary difference between my features and those used by Monath et al. (2021) lies in the method for determining technological overlap between two patents. While Monath et al. (2021) use textual similarity between patent titles, I rely on technological classification codes like the NBER and USPC because titles are not readily available for historical patents.

(inventor pairs that should have been clustered together but were not).

Figure B.1 shows the frontier of the precision-recall curve, computed by comparing the disambiguated inventor identifiers generated by the HAC process at several distance threshold increments with the "true" disambiguated status of each pair in the development set. I calculate the optimal distance threshold using the F-beta score, that is, the weighted harmonic mean of precision and recall. I introduce a weight of 0.4 for precision, aiming to privilege the minimization of false positives. I determine an optimal distance threshold of 0.25, associated to a precision score of 0.986 and a recall score of 0.946.

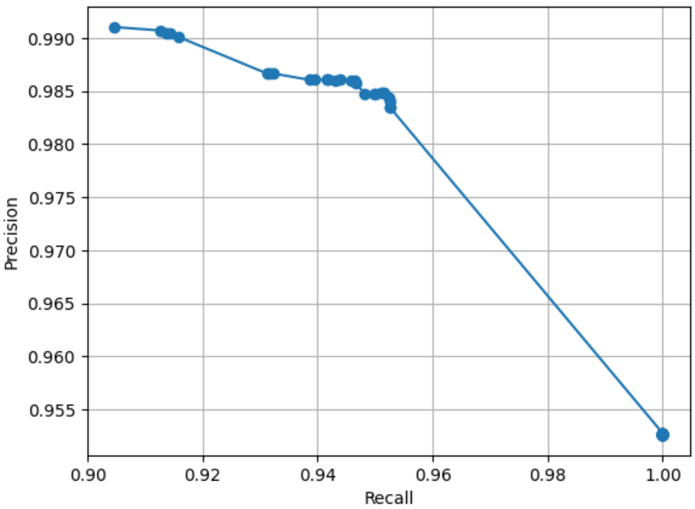
I conclude my procedure by disambiguating all inventors in my dataset, which involves creating the set of inventor canopies and similarity matrices, running the HAC algorithm until the optimal distance threshold of 0.25, and generating a unique identifier for each disambiguated inventor. I count around 2.5 million inventor mentions from historical patents and 5.4 million inventor mentions from modern patents, generating a set of 374 million inventor pairs grouped into canopies. My final dataset includes 2,352,799 disambiguated inventors associated to 4,653,426 patents.

I evaluate the performance of my algorithm through the following steps. First, I focus on inventors from modern patents and assess the overlap between my disambiguated inventor clusters and those of Monath et al. (2021) by calculating the Normalized Mutual Information (NMI) score and the Adjusted Rand Index (ARI). Intuitively, NMI quantifies the similarity between two clustering results of the same dataset, yielding a score from 0 (completely dissimilar clustering) to 1 (perfectly identical clustering). ARI measures the agreement between two clustering results while adjusting for the possibility of random agreement, with values ranging from -1 (completely dissimilar clustering) to 1 (perfectly identical clustering). I obtain an NMI score of 0.995 and an ARI score of 0.960, suggesting that my disambiguation for modern patent inventors is nearly identical to that of Monath et al. (2021).

Second, similar to Akcigit et al. (2022), I search for the top 50 most prolific inventors in my dataset in a list of the most prolific inventors known stored on Wikipedia. A significant challenge for any disambiguation algorithm is distinguishing between homonymous inventors working in the same location and similar technology fields. A low-quality disambiguation process would generate false profiles of top prolific inventors by incorrectly clustering inventors with common names under the same profile. Out of the 50 inventors, I found 45 in the Wikipedia list and confirmed the identities of the remaining five by consulting their biographical profiles on companies' or universities' websites (Table B.1 and Table B.2).

Since Wikipedia tracks inventor careers up to 2024 and includes patents filed at intellectual property offices other than the USPTO, I cannot directly compare the exact number of patents in each inventor's portfolio predicted by my algorithm. Nevertheless, when I focus only on inventors whose entire careers are covered by my dataset, I find minimal differences (Table B.3), both for inventors listed only on modern patents and those with portfolios consisting solely of historical or a mix of historical and modern patents.

Figure B1: Precision-Recall Curve



Notes: Precision and recall scores computed by comparing the "true" disambiguated status of inventor pairs in the development set with disambiguated inventors generated by the HAC process at distance thresholds between 0.01 and 1 with 0.01 increments.

Table B1: Top 50 Most Prolific Inventors: Manual Search Results

Name	**Total patents**	**Found on Wikipedia**	**Found in other sources**
Gurtej S. Sandhu	1157	✓	
Leonard Forbes	1074	✓	
Lowell L. Wood, Jr.	1021	✓	
Donald E. Weder	992	✓	
George A. Lyon	908	✓	
Melvin De Groote	882	✓	
Jay S. Walker	853	✓	
Warren M. Farnworth	763	✓	
Edward K. Jung	722	✓	
Roderick A. Hyde	722	✓	
George Spector	710	✓	
Salman Akram	699	✓	
William H. Eby	697	✓	
Austin Gurney	668	✓	
James A. Jorasch	668	✓	
William I. Wood	663	✓	
Michael J. Sullivan	643	✓	
Ahmadreza Rofougaran	630	✓	
Rick A. Hamilton II	626	✓	
Audrey Goddard	622	✓	
Clarence T. Tegreene	615	✓	
Kie Y. Ahn	604	✓	
Paul Godowski	593	✓	
Jeyhan Karaoğuz	567		✓
Mark I. Gardner	515	✓	
Lee D. Whetsel	514	✓	
Edward J. Nowak	511	✓	
Kangguo Cheng	500	✓	
John F. O'Connor	499	✓	
Ravi K. Arimilli	498	✓	
Geoffrey B. Rhoads	492		✓
Nathan P. Myhrvold	490	✓	
Clyde C. Farmer	490	✓	
Anthony J. Baerlocher	489	✓	
Edwin H. Land	485	✓	
Jack A. Mandelman	481	✓	
Frankie F. Roohparvar	478		✓
Mark A. Malamud	473	✓	
Louis H. Morin	469	✓	
Louis L. Hsu	467	✓	
Royce A. Levien	466	✓	
Muriel Y. Ishikawa	445	✓	
Robert W. Lord	442	✓	
David R. Hall	435	✓	
Niall R. Lynman	427		✓
Jeffrey P. Gambino	421	✓	
Shmuel Shaffer	420		✓
James M. Hart	419	✓	
Scott H. Wittkopp	418	✓	
John D. Rinaldo, Jr.	410	✓	

Notes: Results of my search for the top 50 most prolific inventors ranked by my disambiguation algorithm among Wikipedia's top prolific inventors (`https://en.wikipedia.org/wiki/List_of_prolific_inventors` and `https://en.wikipedia.org/wiki/Talk:List_of_prolific_inventors`; last access: December 2024).

Table B2: Biographical Details of Inventors Not Found Among Wikipedia's Prolific Inventors

Name	Biographical extract	Source
Jeyhan Karaoğuz	"Dr. Karaoğuz conducted pioneering research and development in broadband network access, wireless connectivity, and mobile devices in their early days. He is a prolific inventor with 764 US patents awarded to his name."	North Carolina State University, Department of Computer Science (2023)
Geoffrey B. Rhoads	"Mr. Rhoads, the inventor of ZuluTime's technology, started his career at Tektronix, becoming the principal designer of the first commercially available 1 GHz digitizing oscilloscope. He was the chief scientist in working with key customers/applications such as Los Alamos/Livermore in the then nascent commercial efforts toward fiber optics and laser communications systems. In 1994, Mr. Rhoads founded Digimarc Corporation as its Chief Technology Officer and inventor of digital watermarking technology. At Digimarc he led the intellectual property and technology development, directly resulting in over 600 patents awarded to the company to date."	Crunchbase (2024)
Frankie F. Roohparvar	"In his new role, Roohparvar will serve as chief strategy officer. Prior to joining Xitore, he served as VP and senior strategic advisor at HGST (a WD company). He became part of the HGST team as a result of HGST's acquisition of Skyera, an all flash array storage start-up, where he served as the CEO. Prior to joining Skyera, he was the GM and VP of the NAND business unit at Micron Technology. He is a prolific inventor and has over 500 U.S. patents."	StorageNewsletter (2016)
Niall R. Lynman	"Niall R Lynam, a prolific inventor based in Holland, MI, United States, [...] has made significant contributions to the field of vehicular control systems. With an impressive portfolio of 604 patents, Lynam's expertise lies in developing advanced technologies that enhance driver assistance and improve overall vehicle safety."	IDiyas (2024)
Shmuel Shaffer	"Before joining Cisco in 1999, Dr. Shaffer held managerial positions at Siemens ROLM where he headed the Voice over IP (VoIP) development, the U.S. Hardware organization, and the Embedded SW development. Dr. Shaffer received his Ph.D. from Stanford University in the area of Adaptive Self Optimizing Systems. He also holds two Masters Degrees from Stanford in Electrical Engineering and in Statistics. Dr. Shaffer has authored over 250 US patents."	Stanford University, Department of Electrical Engineering (2011)

Notes: The table reports the biographical details and source I used to confirm the identity of the five inventors ranked among the top 50 most prolific inventors by my disambiguation algorithm, but not found among Wikipedia's prolific inventors.

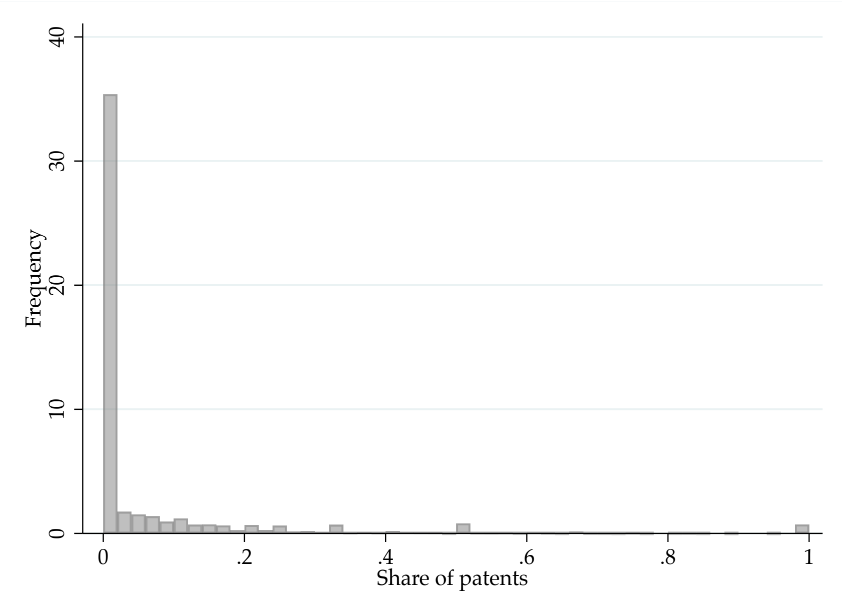
Table B3: Comparison of Total Inventor Patent Stocks: Disambiguation Algorithm vs. Wikipedia's List

Name	**Years of activity**	**Total patents**	**Total patents (Wikipedia)**
Donald E. Weder	1978-2015	992	1000
Melvin De Groote	1924-1966	882	925
George Spector	1974-1998	710	723
Kie Y. Ahn	1970-2013	604	622
Paul Godowski	1994-2010	593	579
Edwin H. Land	1933-1990	485	535
Jack A. Mandelman	1987-2014	481	481
Louis H. Morin	1924-1969	469	503

Notes: The table shows the total number of patents associated to top prolific inventors by my disambiguation algorithm and compares it to the total number of patents found on their profile on Wikipedia's list of top prolific inventors. Since Wikipedia's covers patents filed until 2024, I restrict the sample to inventors whose entire careers are covered by my dataset (1920-2015).

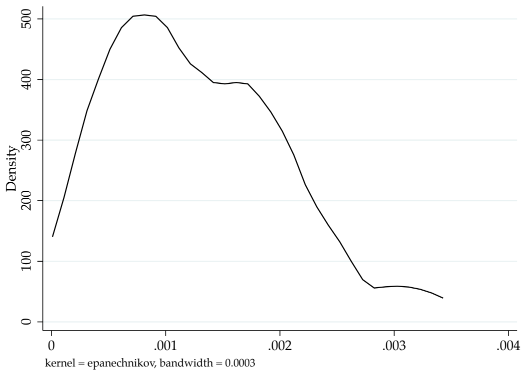
C. Additional Descriptive Evidence

Figure C1: Distribution of Patenting Share Funded by Federal Agencies

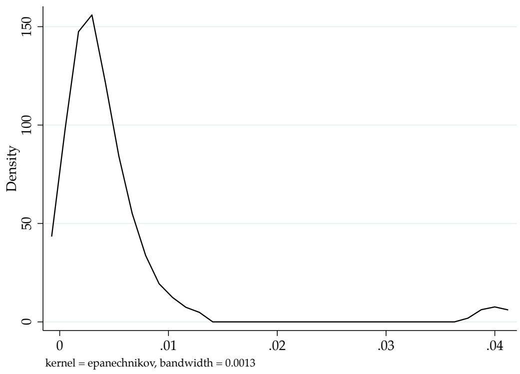


Notes: The figure displays the distribution of the share of patents funded by either the Department of Defense, Department of Energy, Department of Health and Human Services, and NASA in each technology field, commuting zone, and year cell. The distribution is conditional on there being non-zero patenting activity. The data on federally funded patents is sourced from Gross and Sampat (2024).

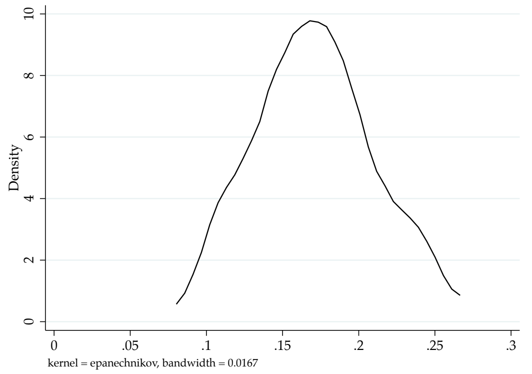
Figure C2: Commuting Zone Pre-SDP Characteristics: Kernel Density Distribution



(a) Patents per capita



(b) Citation-weighted patents per capita



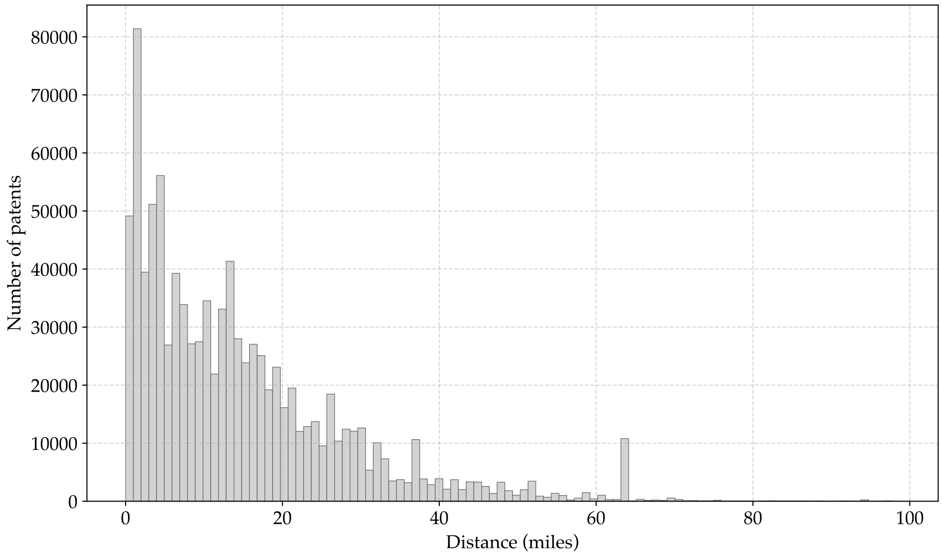
(c) Share of patents citing the scientific literature

Notes: Each panel shows the kernel density distribution of a given commuting zone characteristic measured prior to the SDP introduction (i.e., between 1960 and 1964). Panel (a) focuses on commuting zones' total patents per capita; panel (b) on total citation-weighted patents per capita; and panel (c) on the share of total patents citing the scientific literature.

Table C1: Commuting Zone Pre-SDP Characteristics: Correlation Matrix

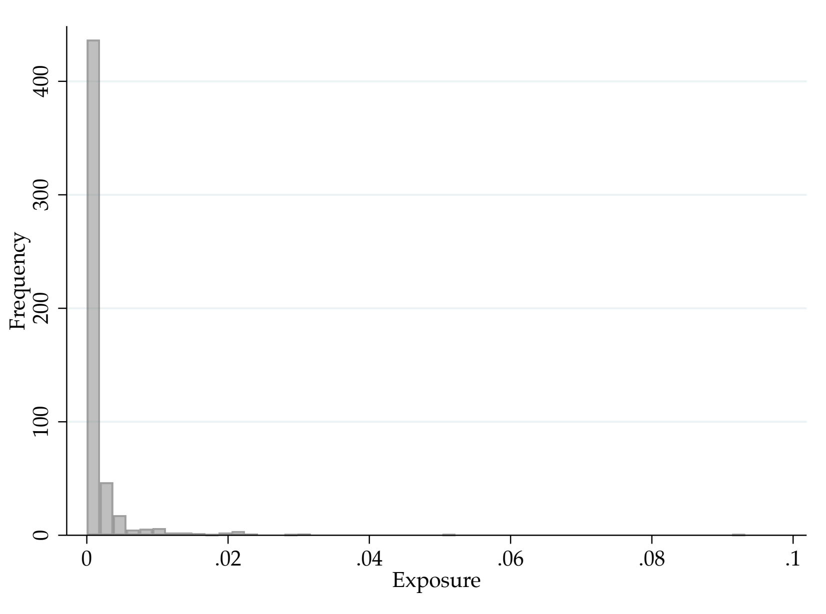
	Patents per capita	Cit.-weighted patents per capita	Share of patents citing the scientific lit.
Patents per capita	1.0000		
Cit.-weighted patents per capita	0.6776	1.0000	
Share of patents citing the scientific lit.	0.1210	0.1607	1.0000

Figure C3: Distribution of Patent Distance to Nearest Local University



Notes: The figure shows the distribution of distances, in miles, from each patent's inventor address to the nearest research university within the commuting zone. For every patent, I geocode each inventor's city, compute great-circle distances to all universities located in the commuting zone, and retain the shortest inventor–university pair. The sample includes only commuting zones hosting a research university ranked by Cartter (1966). Summary statistics: mean = 14.7 mi; quartiles = 4.6 mi (25th), 11.6 mi (50th), 20.6 mi (75th).

Figure C4: Distribution of $Exposure_{ic}$



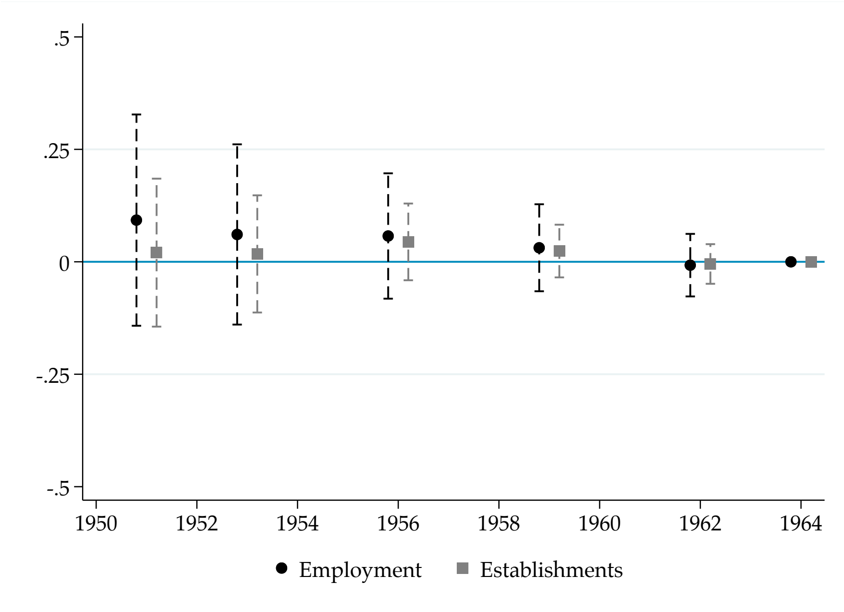
Notes: The figure displays the distribution of $Exposure_{ic} = \sum_j s_{icj} \cdot s_{cj}$, measuring the intellectual overlap between the sources of knowledge for firms active in technology field-commuting zone pairs ic and the knowledge produced by universities in the same commuting zone c, across all journals j.

Table C2: $Exposure_{ic}$ Descriptive Statistics by Technology Category

Technology category	Mean	SD	Min	Max
Chemical	0.0018	0.0037	0	0.0227
Drugs & Medical	0.0016	0.0046	0	0.0211
Computers & Communications	0.0009	0.0025	0	0.0152
Electrical & Electronic	0.0030	0.0090	0	0.0932
Mechanical	0.0010	0.0031	0	0.0311
Other	0.0011	0.0043	0	0.0514
All fields	0.0016	0.0052	0	0.0932

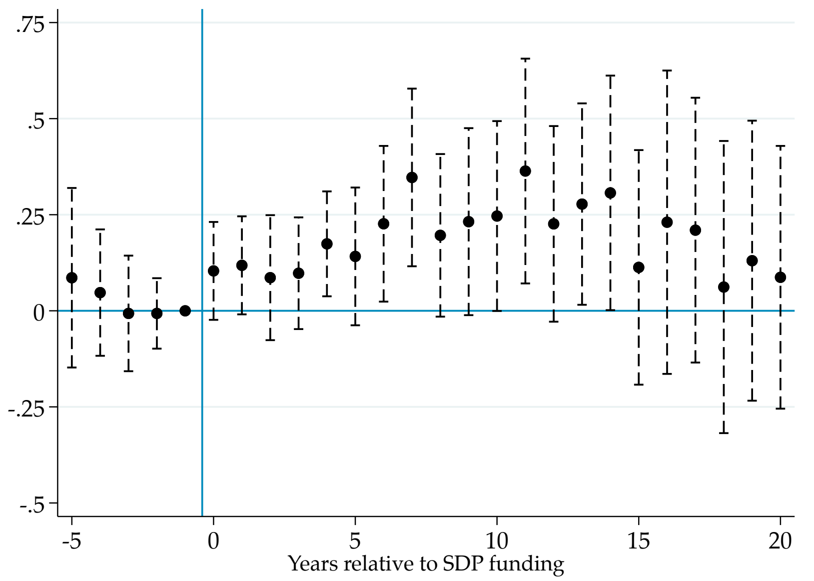
D. Additional Estimates

Figure D1: Commuting Zones Trends Prior to the Science Development Program



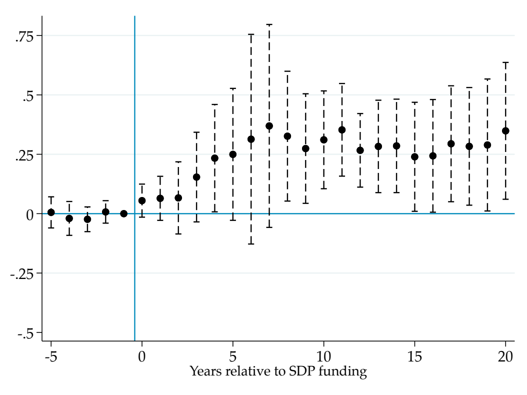
Notes: Coefficients from regressions where the total employment or the total number of establishments in a commuting zone is regressed on year dummies interacted with an indicator equal to 1 for commuting zones hosting an SDP-funded university and equal to 0 for those hosting a top-ranked university. All regressions include year and commuting zones fixed effects and standard errors are clustered at the commuting zone level. The baseline year is 1964. Vertical bars represent 95% confidence intervals. Standard errors are clustered at the commuting zone level. Estimations by Poisson pseudo-maximum likelihood. Employment and establishment data is from the County Business Pattern data digitized by Eckert et al. (2022) and available for the years: 1951, 1953, 1956, 1959, 1962, and 1964.

Figure D2: The NSF Science Development Program and Local Patenting: Citation-weighted Patents

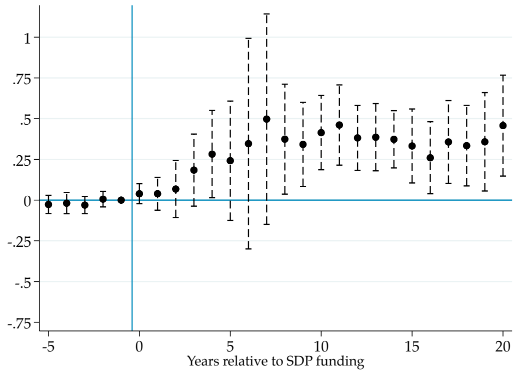


Notes: The dependent variable is the number of citation-weighted patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood. Citation counts from Kogan et al. (2017).

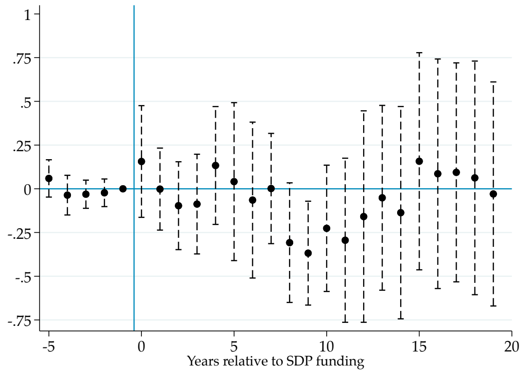
Figure D3: The NSF Science Development Program and Local Patenting: Local vs. Not-Local Assignees



(a) Incumbent assignees (all)



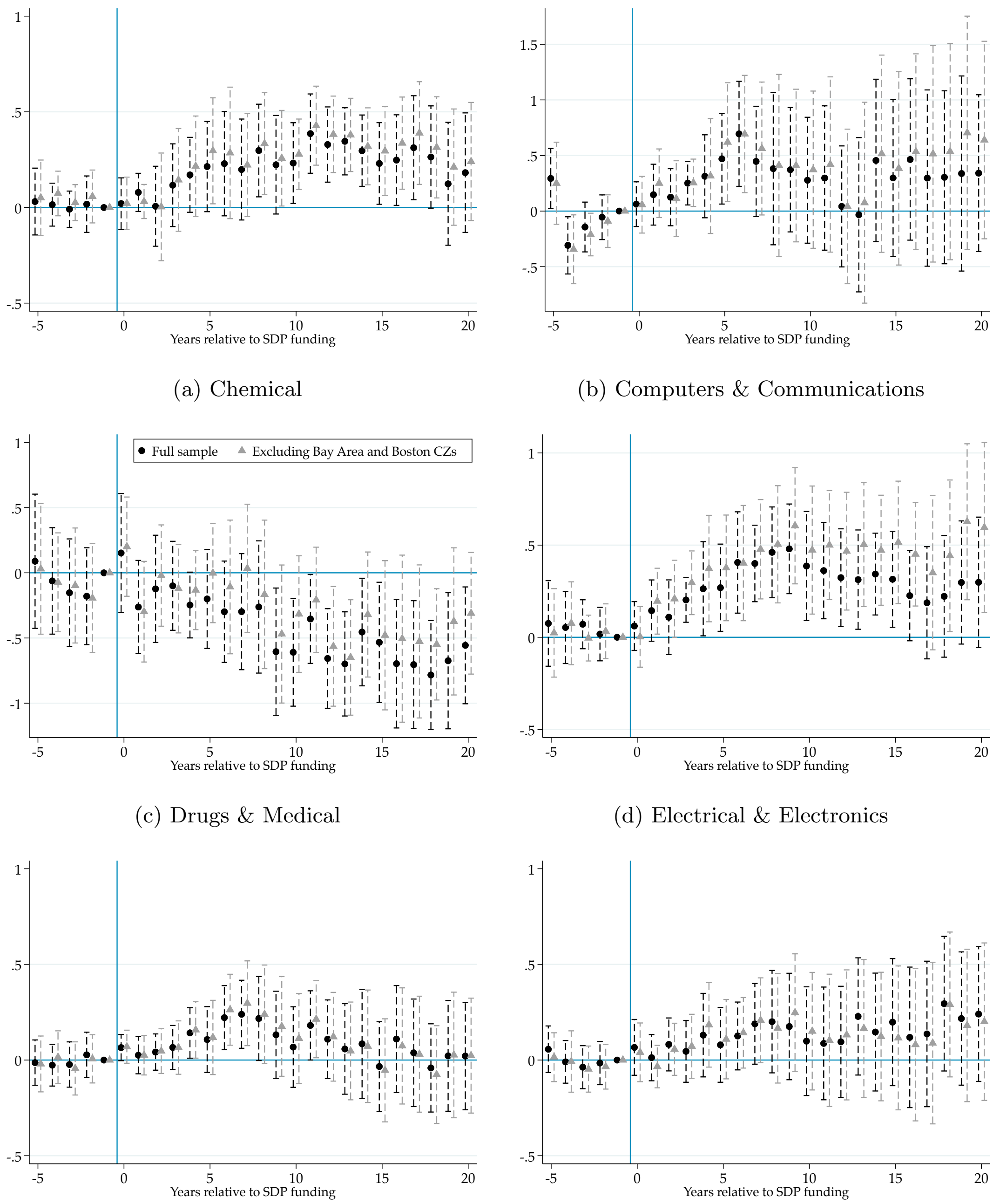
(b) Incumbent assignees (local)



(c) Incumbent assignees (not local)

Notes: The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

Figure D4: The NSF Science Development Program and Local Patenting: Effects by Technology Category (excluding Bay Area and Boston commuting zones



(a) Chemical

(b) Computers & Communications

(c) Drugs & Medical

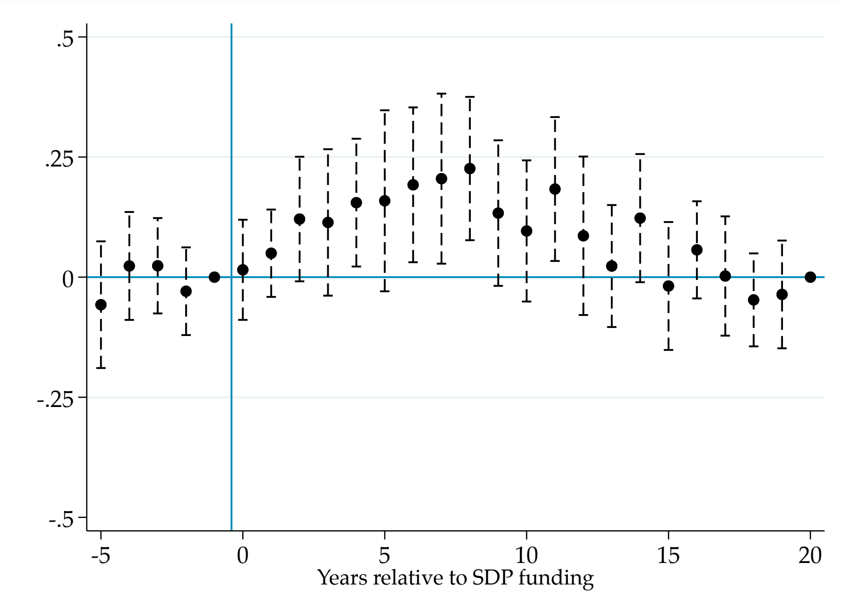
(d) Electrical & Electronics

(e) Mechanical

(f) Others

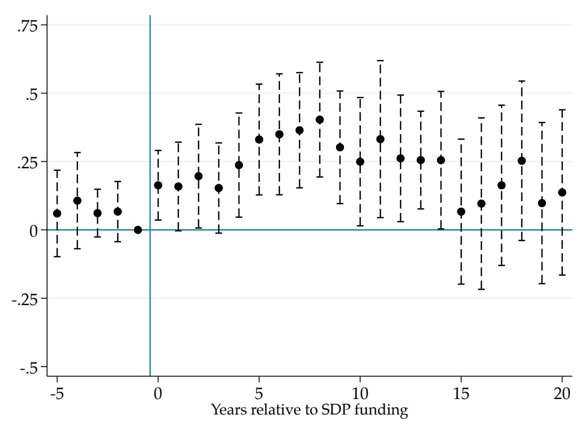
Notes: Black markers report estimates considering the full sample, while gray markers show estimates excluding commuting zones in the Bay Area and Boston from the comparison group. The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

Figure D5: The NSF Science Development Program and Local Patenting: Technology Field Exposure (including technology field-commuting zone-specific time trends)

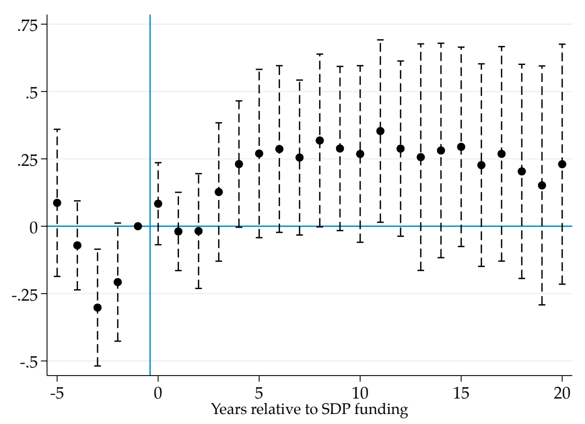


Notes: The figure plots β_τ estimates from a specification equivalent to equation (1), where $I_\tau \times SDP_u$ is further interacted with $Exposure_{ic}$, which measures the extent to which each technology field $\times$ commuting zone pair was exposed to the research activities of co-located universities prior to the Science Development Program introduction. The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. It also includes technology field-commuting zone-specific time trends. The vertical bars represent 95% confidence intervals. The coefficient for the baseline period $\tau = -1$ is set to zero and shown without confidence interval. Standard errors are clustered at the commuting zone level. Estimations by Poisson pseudo-maximum likelihood.

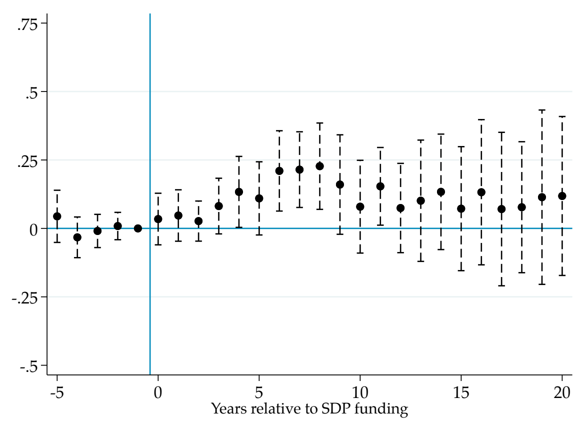
Figure D6: The NSF Science Development Program and Local Patenting: Reliance on the Scientific Literature (Only In-text Citations)



(a) Patents directly citing the scientific literature ($D = 1$)



(b) Patents indirectly linked to the scientific literature ($D \in \{2, 3, 4\}$)



(c) Patents remotely linked to the scientific literature ($D \geq 5$)

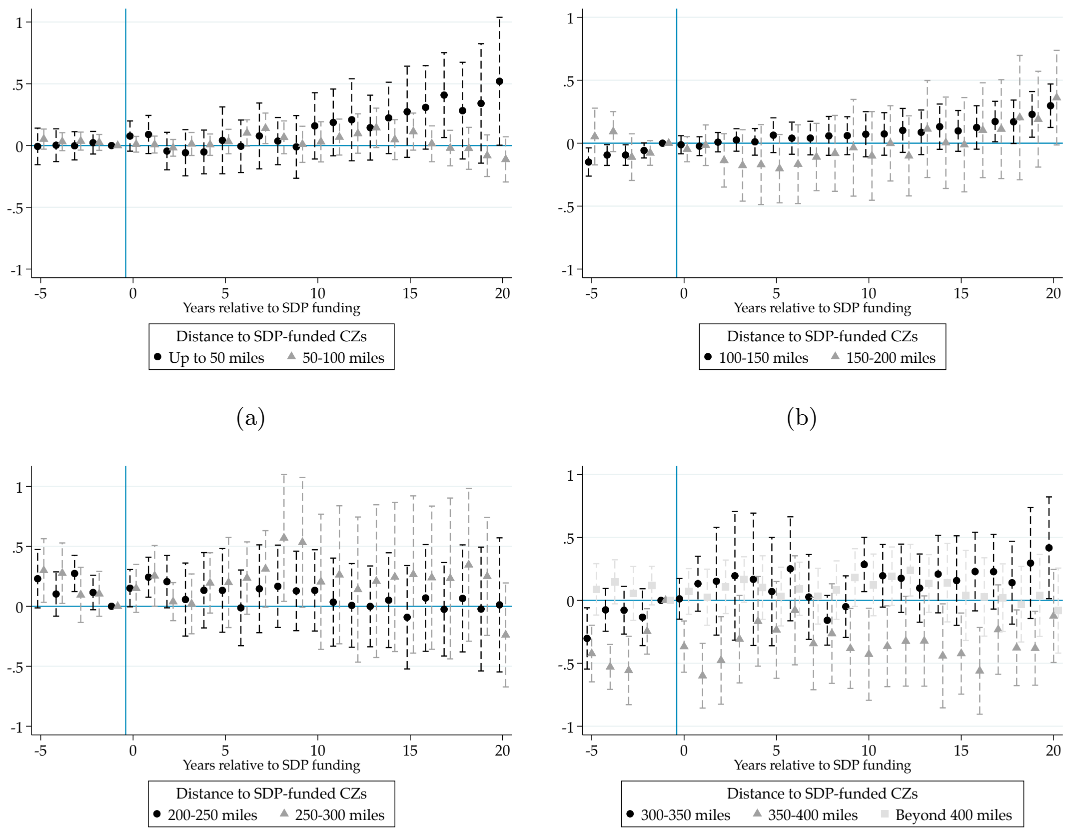
Notes: The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year, along with controls for other institutional funding programs that vary by commuting zone and year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

Table D1: The NSF Science Development Program and Local Patenting: Two-Period Difference-in-Differences Results and Reliance on the Scientific Literature (Only In-text Citations)

	All patents	Private firms' patents	Incumbents' patents	Patents linked to the scientific lit. Direct $D=1$	Patents linked to the scientific lit. Indirect $D \in \{2,3,4\}$	Patents linked to the scientific lit. Remote $D \geq 5$	Share of patents linked to the scientific lit. Direct $D=1$	Share of patents linked to the scientific lit. Indirect $D \in \{2,3,4\}$	Share of patents linked to the scientific lit. Remote $D \geq 5$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. All periods									
$Post_\tau \times SDP_c$	0.123*	0.157**	0.169**	0.140*	0.297**	0.093	0.013**	0.002	-0.015*
	(0.072)	(0.072)	(0.080)	(0.079)	(0.123)	(0.065)	(0.007)	(0.007)	(0.009)
Observations	32,215	31,463	31,311	28,813	27,207	31,815	23,316	23,316	23,316
Pseudo R^2 and R^2	0.842	0.832	0.825	0.693	0.698	0.805	0.353	0.310	0.375
Mean dep.var.	8.21	6.55	5.56	1.41	1.31	5.48	0.17	0.15	0.68
B. Until period 14									
$Post_\tau \times SDP_c$	0.108**	0.136**	0.145**	0.138*	0.278***	0.080*	0.016**	0.002	-0.018**
	(0.055)	(0.053)	(0.070)	(0.070)	(0.102)	(0.047)	(0.006)	(0.006)	(0.008)
Observations	24,421	23,671	23,809	21,473	19,784	24,103	17,887	17,887	17,887
Pseudo R^2 and R^2	0.852	0.841	0.837	0.684	0.690	0.819	0.326	0.325	0.364
Mean dep.var.	8.52	6.65	5.99	1.45	1.13	5.94	0.17	0.12	0.71
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
CZ $\times$ technology field FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Technology field $\times$ year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Institutional grants controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Standard errors are clustered at the commuting zone level and shown in parentheses. Estimations by Poisson pseudo-maximum-likelihood for patent counts and OLS for shares. *** p<0.01, ** p<0.05, * p<0.1.

Figure D7: Patenting in Nearby Commuting Zones by Distance to Closest SDP-funded University (narrower distance groups)

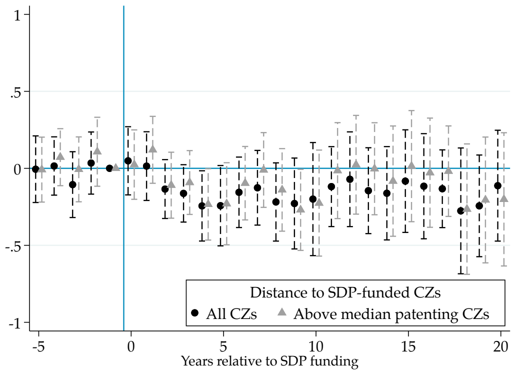


(a) (b)

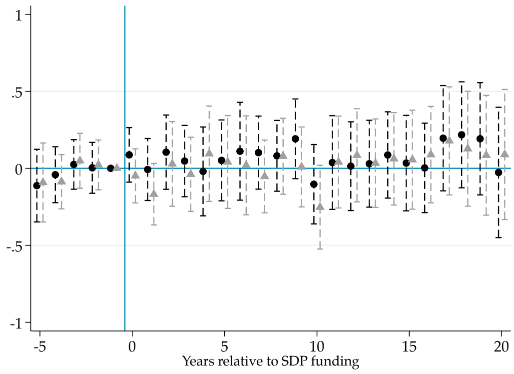
(c) (d)

Notes: Each panel plots difference-in-differences estimates comparing commuting zones lying within a specified radius of the nearest SDP-funded university with commuting zones at the same distance from the closest top-ranked university. Black markers use distances to the full set of SDP-funded universities, whereas gray markers consider only SDP-funded universities situated in commuting zones with above-median pre-SDP patenting. The dependent variable is the number of patents filed in technology field i, commuting zone c, and year t. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.

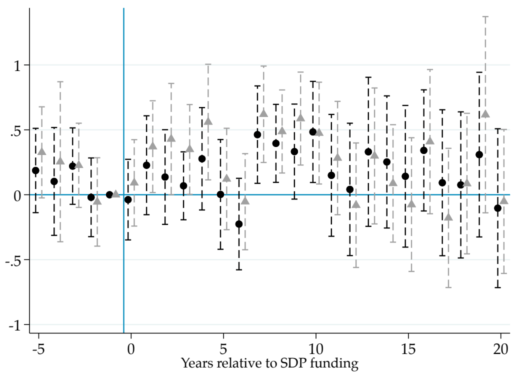
Figure D8: Inventor Outflows from Nearby Commuting Zones by Distance to Closest SDP-funded University



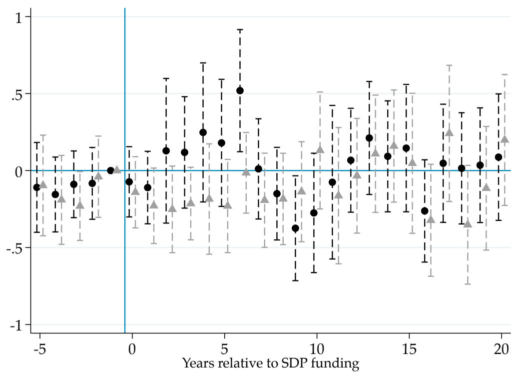
(a) Up to 100 miles



(b) 100-200 miles



(c) 200-300 miles



(d) Beyond 300 miles

Notes: Each panel plots difference-in-differences estimates comparing commuting zones lying within a specified radius of the nearest SDP-funded university with commuting zones at the same distance from the closest top-ranked university. Black markers use distances to the full set of SDP-funded universities, whereas gray markers consider only SDP-funded universities situated in commuting zones with above-median pre-SDP patenting. The dependent variable is the number of inventors whose most recent patent before year t was field in commuting zone c and field i, but who in year t were observed patenting in a different commuting zone and never patent again in c. The regression includes fixed effects for year, commuting zone-by-technology field, and technology field-by-year. The baseline period is $\tau = -1$. Standard errors are clustered at the commuting zone level. The vertical bars represent 95% confidence intervals. Estimations by Poisson pseudo-maximum likelihood.