Research Policy 54 (2025) 105142

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Short communication

The Government Patent Register: A new resource for measuring U.S. government-funded patenting☆

Daniel P. Gross [a,c,*], Bhaven N. Sampat [b,c]

[a] Duke University, Fuqua School of Business, Durham, NC, USA
[b] Arizona State University, Tempe, AZ, USA
[c] NBER, USA

ARTICLE INFO

Dataset link: https://doi.org/10.7910/DVN/BDCCN9

JEL classification:
N42
N72
O31
O34
O38

Keywords:
Public R&D
Patents
Patent policy
Patent data

ABSTRACT

We introduce new historical administrative data identifying U.S. government-funded patents since the early twentieth century. In addition to the funding agency, the data report whether the government has title to the patent ("title" patents) or funded a patent assigned to a private organization ("license" patents). The data include a large number of "license" patents that cannot be linked to government funding from patent text or other sources. Combining the historical data with modern administrative sources, we present a public, consolidated data series measuring U.S. government-funded patents — including funding agencies — through 2020, and we provide code to extend this series in the future. We use the data to document long-run patterns in U.S. government-funded patents and federal patent policy, propose ways in which these data can be used in future research, and discuss limitations of the data.

1. Introduction

Since World War II, the U.S. federal government has been the world's largest funder of research and development (R&D). In 2022, the U.S. government spent nearly $200 billion on R&D, and in each year over the postwar era federal R&D expenditures have accounted for between 0.5% and 2% of U.S. GDP (Anderson and Moris, 2023). Social scientists and policymakers have long sought ways to assess the impacts of federal R&D investments on innovation, jobs, health, security, regional development, and other outcomes. There is significant variation over time in the level and composition of federal R&D funding, including large shocks like the Cold War and Space Race, that could be a source of evidence to inform R&D policy. However, a shortage of granular data linking federal R&D investments to specific outcomes over long horizons has posed challenges for harnessing these opportunities, and data availability has often constrained research to the recent past, limiting what has been learned.[1]

In this paper, we introduce a new long-run, administrative record of U.S. publicly-funded invention. At the heart of the paper is a newly-discovered historical data source: the U.S. Patent and Trademark Office (USPTO) Register of Government Interest in Patents (historically referred to as the Government Patent Register, which we will henceforth in this paper call the "historical GPR"), which we digitized from USPTO's archival records. The historical GPR was for most of the twentieth century the U.S. government's official record of patents which

☆ We are grateful for helpful comments from the editor and two reviewers, as well as from participants at the NBER Innovation and Information Initiative Fall 2023 workshop and the 2023 Mid-Atlantic Patent Works-in-Progress Colloquium at the Georgetown University Law Center. This material is based upon work supported by the National Science Foundation, United States under Grant No. 1951470, which funded the digitizing of the historical data and creation of a public dataset. The data repository for this paper can be found on the Harvard Dataverse titled "The Government Patent Register", URL: https://doi.org/10.7910/DVN/BDCCN9. All errors are our own.

* Corresponding author at: Duke University, Fuqua School of Business, Durham, NC, USA.
E-mail addresses: daniel.gross@duke.edu (D.P. Gross), bhaven.sampat@asu.edu (B.N. Sampat).

[1] While research in bibliometrics, economics, the science of science and other fields has linked publications to funding through in-text funding acknowledgments (e.g., Narin, 1976; Sampat and Lichtenberg, 2011; Azoulay et al., 2019; Fleming et al., 2019; Lin et al., 2023, among others), most of these analyses focus on the post-1980 period when data become more reliable. More progress has been made with historical patents (Fleming et al., 2019), but as we will show, existing text-based approaches have left substantial gaps.

https://doi.org/10.1016/j.respol.2024.105142
Received 18 January 2024; Received in revised form 23 October 2024; Accepted 24 October 2024
Available online 14 November 2024

were developed with government funding, until USPTO transitioned to electronic records. We combine this resource with analogous modern sources to create an extensible dataset of publicly-funded U.S. patents through 2020, including the funding agency and the government's legal interest in the invention. This dataset can support novel research on the evolution and impacts of U.S. research policy.

The data we introduce complement and extend prior data sources by measuring significantly more patents, and more of their features, than can elsewhere be observed. This extension is particularly important for the mid-twentieth century, when federal R&D was growing quickly and according to qualitative accounts was particularly impactful (e.g., Mowery and Rosenberg, 1982; Flamm, 1987, 1988). Previous efforts to measure government-funded patents rely on information in patent text, such as government assignees and interest statements (Fleming et al., 2019; De Rassenfosse et al., 2019). The main reason why this approach undercounts, we discovered, is that many government-funded patents before the 1980s were produced by government contractors and grantees who received title to inventions and wrote no government interest statement in the patent itself, as these were not widely required or consistently used in this era.[2] Numerous publicly-funded patents leave no paper trail outside of USPTO's own internal records.

Administrative records of government-funded patents originate in World War II-era efforts to track patented inventions in which the government held legal interest. The historical GPR was formally created in 1944 by President Franklin Roosevelt's Executive Order 9424 ("Establishing in the United States Patent Office a Register of Government Interests in Patents and Applications for Patents"), which instructed USPTO to produce these records and which remains in effect today. From the 1940s to 1990s, the USPTO's Assignment Branch kept a card index where government interest patents were recorded, with information on the patent, assignor, funding agency, and legal interest, where this legal interest is either title (i.e., government ownership) or license (contractor/grantee ownership, subject to a royalty-free government license). Around 1990, the Assignment Branch transitioned to electronic records, after which government interests can be identified in USPTO assignment data, measuring assignments of patent interests to U.S. government agencies.

Comparisons against existing sources (e.g., Fleming et al., 2019 or PatentsView) suggest that although the historical GPR identifies many patents that other sources miss, and measure these more precisely, it misses some patents too. For example, government interest statements identify some patents not indicated as government-funded in the historical GPR or its modern equivalent. We thus incorporate data from Fleming et al. (2019) and PatentsView to measure government-interest patents independently of administrative USPTO records, and introduce a scalable, large language model (LLM) approach to extracting funding agencies from interest statements. Combining these sources, we provide what we believe is a highly precise dataset on U.S. government-funded patents since 1900 which significantly expands on what was previously available.

Although patents are imperfect proxies for public R&D investment — measuring outputs rather than inputs, and only those patentable and worth patenting — they include information on the timing, geography, and topics of invention, providing a versatile lens on public R&D investment across space and time. These data can be used to study a range of questions, including research connecting U.S. government R&D investments to innovation and other outcomes, within and across agencies, technologies, regions, firms, diseases, and more. Beyond questions around the impacts of federal R&D investments — which the richness of patent data make them well-suited for — a detailed record of federally-funded patents resulting from both "title" and "license" agencies (which kept or deferred title to patents) may also facilitate evaluation of these patent policies themselves. Since World War II, the question of whether the public (i.e., the government) or contractors and grantees should hold title to inventions produced by publicly-funded research has been contended. Although the 1981 Bayh–Dole Act created a uniform federal patent policy by shifting all "title" agencies to a "license" policy, this continues to be controversial. The data we introduce can be used to study the impact of these policies and help inform these ongoing debates.

We proceed as follows. Section 2 provides background on the USPTO's historical Government Patent Register (the centerpiece of our data) and the evolution of government patent policy since World War II. In Section 3, we discuss the contents of the historical GPR, and in Section 4, we describe how we extend it with analogous data from the more recent past. Section 5 documents patterns in the data. In Section 6 we describe the contents of the dataset being released with this paper. We conclude in Section 7 by discussing use cases, as well as limitations and potential gaps in the data. The online data repository accompanying this paper provides (i) the historical GPR (ii) a consolidated dataset of government interest patents through 2020, which we call the Government Patent Register (without the historical qualifier), and (iii) documentation, code, and instructions for extending these data as more patents are issued.

2. Institutional background

2.1. Origins of the register

Five days after Pearl Harbor, President Franklin Roosevelt established a "National Patent Planning Commission" (NPPC) to study various aspects of the patent system—at the time, before the wartime expansion of government R&D funding, "the only provision of the government for the promotion of invention and discovery" (Commission, 1941, p. 7). Among the subjects it considered was patents the government had rights in, which until the war were mainly inventions produced by government employees. The NPPC considered various questions that would become prominent during and after the war, including whether government-owned patents were desirable at all (for the public interest), and the costs and benefits of exclusive licensing of these patents. But it also noted a paucity of information on exactly how many patents the government had rights in to begin with. To that end, it recommended the creation of a "central source" of information on patents where the government held a legal interest.

Following the NPPC's recommendation, amid the war, President Roosevelt's Executive Order 9424 (February 18, 1944) created the Government Patent Register:

> WHEREAS there exists among the several executive departments and agencies a need for a more adequate source of information with respect to patent rights and interests owned or controlled by the United States Government; and
>
> WHEREAS the establishment in the United States Patent Office, Department of Commerce, of a separate register for the recording of such patent rights and interests would meet this need and would be in the public interest
>
> ...
>
> The Secretary of Commerce shall cause to be established in the United States Patent Office a separate register for the recording of all rights and interests of the Government in or under patents and applications for patents.

Roosevelt's Executive Order also instructed government departments and agencies to forward to the Commissioner of Patents information on any patents (or applications) where the government had rights, including not just those where the government agency was an assignee, but also those that were government-funded but held by grantees or

[2] Government interest statements are haphazardly reported even today (Rai and Sampat, 2012).

contractors, and to which the government had license. The register was to be maintained by the USPTO's Assignment Branch. Though the Executive Order was issued in 1944, the Assignment Branch made efforts to backfill information (Watson and Holman, 1964), such that the register (and the data we collect from it) included patents issued in the 1920s and 1930s, and as far back as 1890.[3]

Though we were unable to find official documentation of USPTO's procedure for reporting and recording government interests, Watson and Holman provide contemporary insight, explaining that inventions were reported to federal agencies' patent departments "by Government laboratories and other research facilities and by contractors" (Watson and Holman, 1966). These agencies were required to then "forward to the [USPTO] Assignment Branch complete information" about these patents (Watson and Holman, 1964). They also note three potential sources of "small errors" in these records: "First, government agencies might not send information about every patent to the Assignment Branch; secondly, clerks in the Assignment Branch can make many different kinds of mistakes; finally, people using the indexes can misplace the cards because they are loose" (Watson and Holman, 1964). In Section 3 we address the second source of error (clerical mistakes) by cross-validating data against Google Patents.

2.2. The evolution of government patent policy

With the explosion of federal R&D during World War II, most through contracts to universities and private firms, questions over who should own patents resulting from publicly-funded research grew in both importance and prominence. The wartime research effort, coordinated and led by the newly-established Office of Scientific Research and Development (OSRD), funded extramural research at levels unimaginable prior to the war, which often resulted in patentable inventions—and necessitated a policy on patent rights.

The patent terms which OSRD initially wrote into R&D contracts held that the government would retain title to any patents that resulted from its funding, reflecting a presumption that the fruits of publicly-funded research should belong to the public. But after some firms showed reluctance to engage in OSRD-funded work due to patent rights (see Gross and Sampat, 2023b), OSRD adopted for some contracts what became known as the "long form" patent clause, which allowed contractors to retain title to patents, provided the government received a royalty-free license for wartime use. In other cases, the original "short form" clause continued to be used (giving the government presumptive title)—particularly when public interest required government ownership (such as atomic fission, or medical research with public health benefits), when OSRD (rather than its contractors) supplied necessary research equipment and personnel for the contractual R&D, and for research performed under contracts with universities.

Given the success of the wartime R&D effort, it was widely recognized that the federal government would continue to be a significant funder of extramural research after the war ended. How, and in what form, was a point of legislative contention. A specific point of contention (among others) was patent policy: OSRD's choices had been controversial, with critics objecting that the long-form clause had given away the fruits of publicly-funded research (Sampat, 2020).

Most major legislation for postwar research funding contemplated one major research funder (called the National Research Foundation in some bills, the National Science Foundation (NSF) in others). While these proposals were mired in Congressional debates over details — including who should get title to patents resulting from publicly-funded research — other government agencies absorbed the wartime portfolio. The Public Health Service (PHS) picked up wartime medical research (through the National Institutes of Health, or NIH); the Department of Defense (DoD), military research; and the Atomic Energy Commission (AEC), research on nuclear fission. The postwar R&D system was thus fragmented, with a large number of research funding agencies, by the time the NSF was created in 1950 (Kevles, 1977).

One consequence of this splintering is that each agency evolved its own patent policy "without any central guidance or overall coordination" (Federal Council for Science and Technology, 1976, p. 1). As Eisenberg (1996) explains, DoD and NSF initially had "license" policies similar to OSRD's long-form clause, where contractors and grantees could retain title and the government received a royalty-free license. Other agencies, including the AEC (later the Department of Energy, or DOE), Department of Agriculture (USDA), Department of the Interior (DOI), and Department of Health, Education, and Welfare (HEW, which included the NIH) had "title" policies under which the government retained title, like the OSRD short-form policy.

There was also variation within agencies over time. In response to President John F. Kennedy's 1963 Statement on Government Patent Policy, some agencies — most notably, DoD — made changes to their policies. Prior to these revisions, special permission was required by DoD in keeping title to patents; after the policy change, title was asserted as the norm. The National Aeronautics and Space Administration (NASA, founded in 1958), on the other hand, initially had a title policy, but in 1963 shifted to a license policy for most of its patents. And every agency had exceptions to its standard policy, such as license policy patents which reverted to government ownership if the contractor or grantee chose not to file, or petitions for title from contractors of license policy agencies. Some agencies had no formal policy, and instead "simply ignored the issue ... which in effect permitted contractors to retain all rights to inventions" (Federal Council for Science and Technology, 1976, p. 1). One implication is that the Government Register became an essential resource for keeping track of which patents the federal government controlled or had a legal right to use.

Uniformity was finally achieved by the 1981 Bayh–Dole Act, which created a uniform "license" policy under which contractors and grantees would own inventions created in the course of publicly-funded research, with the government retaining a license for its own use. Originally limited to universities and small businesses — reflecting concerns about giving away government-funded inventions to large firms — the Act was extended to all recipients of federal R&D funding in 1983 by an Executive Order from President Ronald Reagan. Bayh–Dole also required contractors and grantees to include "government interest statements" in the text of patents, which was not common practice beforehand, and even since then has suffered from significant non-compliance (Rai and Sampat, 2012).

The history of government patent policy thus suggests that simply looking at patents assigned to a government agency would miss a large number of government-funded patents, as would searching for government interest statements in the patent text—especially pre-1980. In Section 5 we verify this empirically, using information from the Government Register.

3. The historical government register

Though a few contemporary studies of the Government Register were produced in the 1960s — including several Congressional reports on patenting practices at specific agencies, which consulted the register — it has largely been passed over since.[4] As far as we know, the only previous empirical work using the register was a set of papers by economists Mary Holman and Donald Watson (Watson and Holman, 1964, 1966, 1967), who introduced the register as "a valuable data source hitherto unexploited" (Watson and Holman, 1964).

[3] Given this, we consider the data most reliable post-1944—which also coincides with the emergence of the modern federal R&D funding system.

[4] See, for example, U.S. Congress Joint Committee on Atomic Energy (1959), U.S. Senate Judiciary Committee, Subcommittee on Patents, Trademarks, and Copyrights (1959, 1961), U.S. House of Representatives, Committee on Science and Astronautics (1966) .

Fig. 1. Example index cards in the historical Government Register.
Notes: Figure shows example index cards from the historical Government Patent Register. Examples illustrate variation in funders, research performers, and patent policies.

These studies, in tandem with Executive Order 9424, revealed to us the historical GPR's existence, and suggested where we might find it. As Watson and Holman (1964) explain, the register was maintained by the USPTO Assignment Branch, in three index card sets recording government interest patents, each with different sequencing (to facilitate manual searches). Though these records were available in the 1960s for public inspection, this is no longer the case today—which led us into a hunt for the index cards Holman and Watson describe. We eventually found these records at the U.S. National Archives (NARA), where they were accessioned to its USPTO collection (Record Group 241, "Records of the Patent and Trademark Office"), in a 174-box set (mis)titled as "Index to Patent Assignments by Government Licensees, 1/1/1890-12/31/1955".[5] Appendix Figure A.1 provides the National Archives' box list. Inspection of the records confirmed that they were what we sought, included both government-assigned and -licensed patents, and extended into the 1990s (despite their title claiming 1955). As Appendix Figure A.1 shows, it indeed consisted of three sets of index cards, each with different index sequencing: one in alphabetical order by assignor, one by funding agency, and one by patent number. We picked one of these sets (the first set, by assignor), and digitized its complete contents, with a total of 127,852 index cards.

Fig. 1 provides examples from these records, with patents: (i) from the 1940s, 1950s, 1960s, and 1980s; (ii) funded by different agencies (OSRD, the Army, the Navy, the Air Force, and NASA); and (iii) where the government interest took the form of a title or license on extramural invention, and title to intramural (employee) invention (the latter denoted as "Act of 1883" or "U.S.C. 266" — i.e., 35 U.S. Code Section 266 — which were legal statutes which determined government rights in employee inventions).[6] Each index card provides several pieces of information, including the patent number, filing and issue date, and title; the inventor and assignor (e.g., the contractor or grantee filing the patent); and the government interest (title or license).

We scanned and transcribed, cleaned, and regularized these data, including by hand-checking values of numeric fields with non-numeric characters, confirming that all identifying information is internally

[5] The records can be found in the NARA online catalog at https://catalog.archives.gov/id/159071266.

[6] Government rights in employee inventions were first established by legislation in 1883 (P.L. 47-103, labeled in the historical GPR as "Act of 1883") and amended by legislation in 1928 (P.L. 70-325), which was written into the U.S. Code at 35 U.S.C. §266 ("USC 266"). The text of USC 266 stipulated that the U.S. government would obtain title to all inventions made by government employees where the invention was made using government resources or resulted from the inventor's official duties. However, the statute also provided exceptions when "the contribution of the Government" was insufficient to justify assignment, or when the government was deemed to have insufficient interest in the invention. In these cases, title was given to the employee, subject to the provision of a "non-exclusive, irrevocable, royalty-free license" to the government. Under this framework, the U.S. government could have title or license in employee inventions—explaining why we see both in the register.

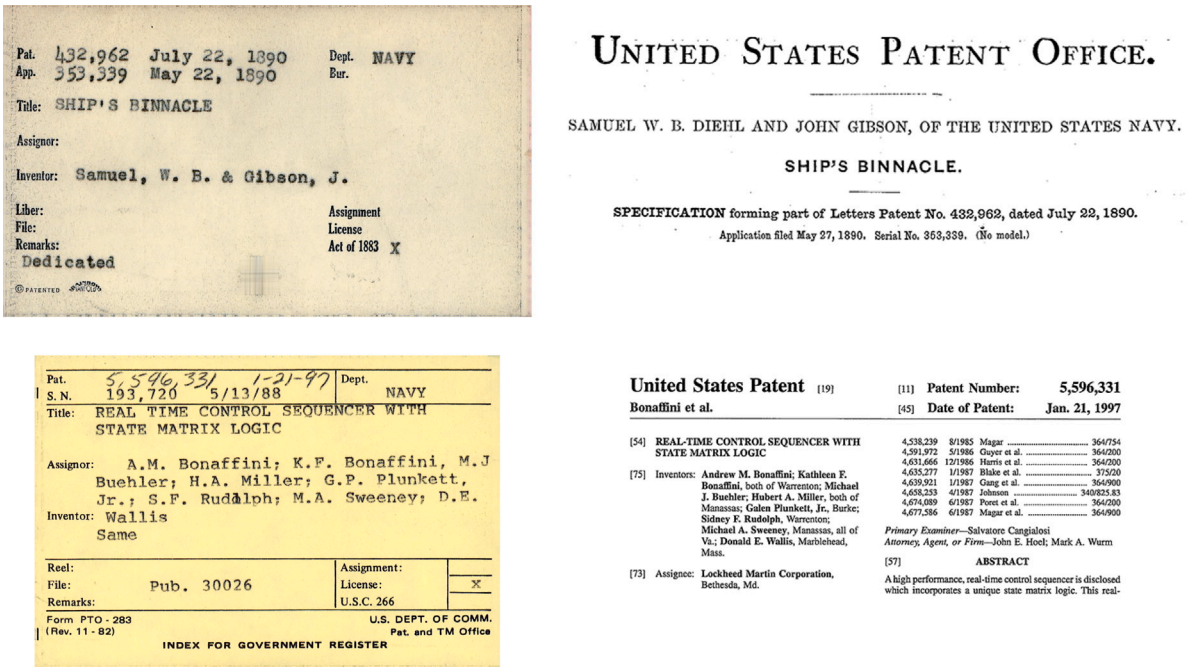

Fig. 2. First and last patents in the historical Government Register.
Notes: Figure shows the first and last index cards in the historical GPR (issued in 1890 and 1997, respectively), juxtaposed against the associated patent publications.

consistent (manually resolving inconsistencies), and harmonizing government agency names and spellings, aggregating them up to modern cabinet-level departments where possible (e.g., Army, Navy, Air Force, War Department all become DoD; AEC becomes DOE; HEW, PHS, NIH all become the modern Department of Health and Human Services (HHS); etc.). We drop index cards recording patent applications which were later abandoned, government interest patents at foreign patent offices, design patents, reissues, and a handful of index cards which identified firms that gave the U.S. government license to all of their patents for the duration of World War II only, which we interpret more as public service than a contractual legal interest. These special cases comprise only a small fraction (3%) of all index cards.[7]

Even with accurate transcription, a residual challenge is the possibility that the physical index cards themselves may sometimes have typographical errors in their printed contents. To a first order, we expect these will be rare and non-systematic—but if our goal is a complete accounting, we would like to identify and fix source errors. To do so, we cross-validate the information on every card—patent numbers, serials, filing and issue dates, inventors, and titles—to ensure they match, by comparing them to Google Patents. When the patent number does not agree with other fields, but other fields agree with each other, we replace the provided patent number with that implied by other fields—a correction which implicates 694 patents (around 0.5% of the sample). We also drop a handful of patents (<100) for which we were unable to reconcile the provided information.

The first patent in the final working data is U.S. patent number 432,962 (“Ship’s Binnacle”, issued July 22, 1890 to two inventors in the U.S. Navy), and the last is 5,596,331 (“Real-time Control Sequencer with State Matrix Logic”, issued January 21, 1997 to Lockheed Martin), which has no indication of government interest in the printed patent but which the register indicates is a license patent (to DoD). The associated index cards and patents are seen in Fig. 2.

The final data we construct from the historical GPR include 110,158 unique patents, and identify patents funded by the following agencies (ordered alphabetically by acronym, as in the accompanying dataset): the Departments of Commerce (DOC), Defense (DoD), Energy (DOE), Interior (DOI), Justice (DOJ), and Transportation (DOT); Environmental Protection Agency (EPA); Department of Health and Human Services (HHS); National Aeronautics and Space Administration (NASA); National Science Foundation (NSF); Department of the Treasury (TREAS); Department of Agriculture (USDA); and Veterans Administration (VA). Some patents are associated with multiple agencies, either because multiple agencies were printed on the card, or (more often) because they had multiple associated index cards with different agencies printed. Of 110,158 patents in the final data, 109,336 (99.3%) have one associated funding agency, 637 (0.6%) have 2+ agencies, and 185 (0.2%) have no agency listed. Fig. 3(A) provides the share of patents in the data associated with each of these agencies through 1975.

This information is provided in the data released with this paper. The dataset also indicates whether each patent was found in the register with a title, license, or unknown (unmarked) interest. It also identifies patents marked as employee inventions (U.S.C. 266). Some patents had multiple cards in the card index with different recorded interests, or multiple interests on the same card. Patents where both title and license are indicated are ambiguous cases, and in our analysis below we treat these as having an unknown government interest.

4. Modern data on government-interest patents

The number of patents in the historical GPR begins to dwindle in the 1980s (declining from 1650 patents in 1980, to 870 in 1990, to

[7] The presence of these special cases raises the question of whether the register may cover non-government financed inventions to which the federal government holds title or license by other means, including donations, purchases, or other agreements. Historical evidence, including the implementation guidance for E.O. 9424, suggests against this possibility; if anything, the exceptions appear to mainly be these few World War II cases. Consistent with this understanding, Watson and Holman’s early study of government-financed R&D used the register, explaining that “Nearly all of the patents owned by and licensed to the federal government are one of the outcomes of the research and development performed by the government and by its contractors and grantees” (Watson and Holman, 1964, p. 208). The 1968 Harbridge House study of government patent policy—which was influential in the debates leading to the Bayh–Dole Act—also sampled from the register, using it to identify patents “which resulted from work performed by contractors under government financing” (Harbridge House 1968).

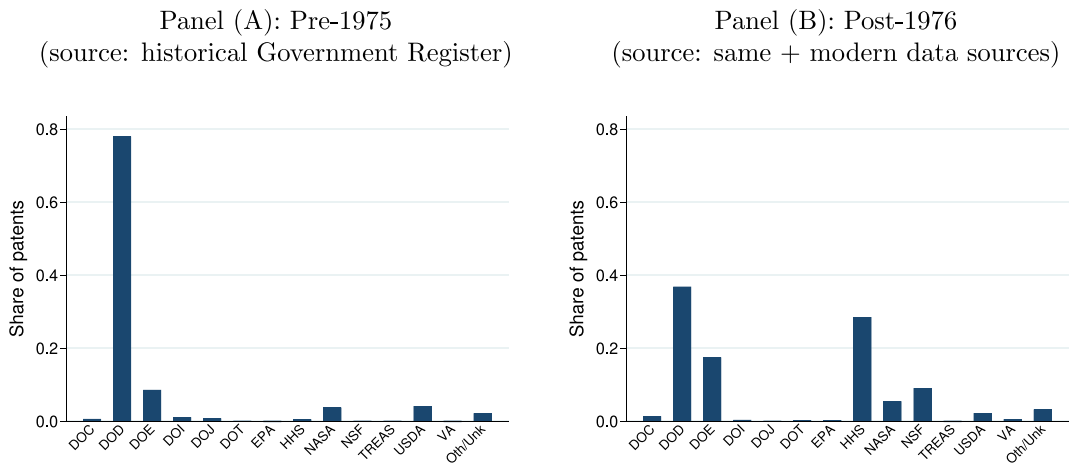

Fig. 3. Agency shares of government-interest patents.
Notes: Table shows each of 13 federal agencies' share of government interest patents, pre-1975 (i.e., the era before modern electronic records, using the historical GPR; Panel A) and post-1976 (using the historical GPR and modern data sources such as the UPAD and PatentsView; Panel B).

264 in 1995) and formally ends in 1997. But E.O. 9424 remained in effect, and still does today. A second goal of this paper is thus to combine the historical GPR with more recent administrative data sources for more complete and extensible data on government interest patents that fills in gaps left by text-based approaches. We harness information in the USPTO's Patent Assignment Dataset (Graham et al., 2018; also see Marco et al., 2015, Marco and Vishnubhukat, 2015), which reports some later E.O. 9424 patents. We supplement this with data from PatentsView on patents with government interest statements, introducing a new method of extracting funding agencies from the interest statement text. From each of these sources we retrieve data for patents issued through 2020. We then merge in measures of government-assigned patents based on Fleming et al. (2019) and PatentsView assignee data.

4.1. The USPTO patent assignment dataset

The USPTO Patent Assignment Dataset (UPAD) is a dataset which reports transactions conveying U.S. patent interests (title or license) between assignors and assignees (Graham et al., 2018). The UPAD is derived from records of the USPTO Assignment Branch (the same office which maintained the historical GPR), includes patent transactions since 1980, and is currently updated annually. As Graham et al. (2018) explain, the UPAD should subsume government interest reporting under E.O. 9424: assignments "required to be filed by Executive Order 9424 are recorded in the Office's assignment records and, with some exceptions, will appear in the UPAD" (p. 348).

Potential government interests are observable in the UPAD in a few ways. One is through an author-coded "conveyance type" variable, which sometimes takes the value "govern" (for government interests). Another is through the conveyance text directly: the UPAD contains many instances of conveyances reported as "EXECUTIVE ORDER 9424, CONFIRMATORY LICENSE" and similar variants. Closer inspection suggests that when recorded in UPAD, government-license patents will generally be identified by this or similar conveyance text, whereas government-assigned patents will instead have traditional assignment text seen for other transactions (including between private parties), such as "ASSIGNMENT OF ASSIGNORS INTEREST". In all of these cases, we can identify the specific government agency which funded the research as the recipient of the conveyed interest (see Appendix Figure A.2–A.3 for an example conveyance).

Because we would like to build a modern Government Patent Register from the UPAD, we aim to measure both. This presents some challenges, however, including measurement error: not all government interest conveyances in the UPAD take these forms, and not all conveyances of these forms have government assignees. We thus cast a wider but still-precise net. We began by identifying all assignees (i.e., interest recipients) of conveyances with the text "Executive Order 9424" or "Confirmatory License", or which are coded with a "govern" conveyance type. We then evaluated this list to specifically identify government assignees in the UPAD by name, and to associate them to cabinet-level agencies (as we did for the historical GPR in Section 3). We then reviewed all remaining UPAD assignees for federal agencies via careful manual string matching. We subsequently retrieved the conveyance text of all conveyances to these entities and manually classified them as title or license. Putting the pieces together, we code transactions as (i) conveying interest to a government entity, and (ii) whether that interest is title or license.

This approach returns a broad, precisely-measured sample of government interest patents, with the funding agency and the legal interest. Even so, there are reasons why it may be incomplete. One is the possibility of underreporting (despite the requirements of E.O. 9424). Another is errors in the source. Two features of the data bolster our confidence: (i) 90% of government interest patents which we identify through the UPAD are also measured in other sources, and (ii) similar to the historical GPR, effectively 100% of these patents have internally-consistent patent numbers, serials, dates, inventors, and titles when compared to other sources.

We were nevertheless cautious about the roughly ten thousand government interest patents identified by UPAD that none of our other sources measure—which comprise 5% of our post-1976 sample (see Section 6). To better understand whether they are accurately measured, we probed these cases further by reading the interest conveyance letters associated with these patents, using the USPTO's Patent Assignment Search website. In a random sample of 100 such patents, we found that we accurately measure the underlying interest in 88% of cases, and that mistakes were due to human error in the conveyance transaction — generally, mistakes during the data entry process — which resulted in the conveyance transaction being linked to the wrong patent. We consider this error rate low—especially given that it applies only to UPAD-only cases, which are a small share of our sample, and that these errors appear random. Nevertheless, in certain contexts (e.g., litigation relating to specific patents), users of our data may want to confirm the accuracy of individual UPAD-only cases by reviewing the associated legal interest conveyance letter.[8]

[8] These can be found at https://assignment.uspto.gov/patent/index.html.

4.2. PatentsView data on government interest statements

A second resource for measuring modern government interest patents is PatentsView, which covers the post-1976 period and provides a patent-level data file with government interest statements (`g_gov_interest.tsv`), and an accompanying file listing government agencies extracted from them (`g_gov_interest_org.tsv`).[9] Patents in these lists only partially coincide with the historical GPR and the UPAD, in part because interest statements are not always included in the text of government interest patents. PatentsView interest statement data are nevertheless useful as a complement to other sources, since each includes patents which others omit.

The agencies identified in the `g_gov_interest_org.tsv` file can be crosswalked to the same agencies we identify in other sources (the historical GPR and the UPAD). After probing the data and comparing them to the underlying interest statements, we observed they mostly agree, but also noticed occasional imprecisions and discrepancies.[10] This led us to explore a new approach to extracting funding agencies from interest statements using large language models (LLMs): feeding the interest statement text provided in the `g_gov_interest.tsv` file to OpenAI's GPT-4 (in February 2024), and prompting it to identify the U.S. federal agencies acknowledged, or the U.S. government (generically) if the statement claims a U.S. government interest but no specific agency is named or thanked for funding.

After running this prompt on all ~171,000 patents in the PatentsView file, we received back a list of government funders. We manually processed this list to map named entities to our focal agencies and remove foreign and state governments and private organizations that co-appear with public ones. Using these data, we identify government interest patents and funding agencies. Manual inspection of the results reinforced our confidence in an LLM-based approach, under which we identify 169,360 patents with a U.S. government interest.[11] Additional advantages to this approach are that it is cheap (at a cost of about $1.50 per 1000 interest statements processed at the time of execution, which has since declined further) and can be easily extended when the input file is updated.

4.3. Government-assigned patents

A third resource is patent assignment data, which we use to directly measure government-assigned patents. We use PatentsView assignee data to identify government assignees from 1976 onwards, which we manually review and associate to government agencies. To extend our measurement backwards, we use assignee data from Fleming et al. (2019)—a prior effort to measure government-funded patents through patent text. Following our approach to the PatentsView data, we review (Fleming et al., 2019)'s extracted assignee text to identify U.S. government entities and to crosswalk these government assignees to specific federal agencies.[12]

4.4. Additional cases from Fleming et al. (2019)

The final resource we use to identify government interest patents are Fleming et al. (2019)'s measures of "government acknowledging" patents, which indicate patents which appear to have government interest statements in their text, based on their contents (identified by algorithmically detecting specific keywords — like "government" — near the front of the patent).

In the 1960s and early 1970s, there are a few thousand such patents in the Fleming et al. data which are not reported by the historical GPR. A manual review of 50 of these cases, however, revealed that many were false positives: patents measured as government-acknowledging which mention government uses, standards, or regulations that motivated the invention, but which did not actually have interest statements or indications of government funding.

Some subset, however, were accurate cases which were missing from the historical GPR. To isolate these cases, we retrieved the full text of roughly 4000 government-acknowledging patents in the Fleming et al. data but not the historical GPR filed between 1945 to 1975, and implemented a semi-automated processing pipeline, using GPT-4 to identify patents in this sample with a potential interest statement and manually reviewing these cases to identify the true positives. We then used a similar semi-automated procedure to extract the funding agency, where provided. This procedure contributes an additional 818 government interest patents to our data—a relatively small number in the wider dataset, but which fills modest gaps in the historical GPR, especially in the 1960s.

4.5. A "Government Patent Register" for research

We combine these sources to create a modern Government Patent Register that runs from the early twentieth century (the beginning of the historical GPR) through 2020, and which we are releasing with this paper, along with code for future extensions. To our knowledge these data are the most complete long-run accounting of U.S. government-funded patents available—particularly because, as we show in Section 6, each of these sources fills gaps in others.

The posted dataset begins with a base layer of USPTO patents issued between 1836 and 2020, which we obtained from Google Patents. We then merge in measures of government interest patents, the nature of the government interest (where known), and the funding agency from (i) the historical GPR, (ii) the UPAD, (iii) government interest statements (from PatentsView), and (iv) government assignees (from Fleming et al. and PatentsView). We then measure government interest patents as the union of these samples.[13] We measure a patent as government title or license, and associate it to our focal agencies, if it is ever measured as such (in any source).

The final data file includes roughly 275,000 government interest patents issued through 2020. Fig. 3 shows the agency distribution, pre-1976 (left panel, based on the historical GPR only) and post-1976 (right panel). We are able to associate nearly all government interest patents with their funding agency. Historically, nearly 80% of these were DoD-funded. More recently, U.S. defense, energy, and biomedical research funding agencies' shares have been closer to parity. We will discuss these differences more in the next section, as we explore other patterns in government-funded patenting and patent policy.

[9] Available at https://patentsview.org/download/data-download-tables.

[10] For example, patents where the U.S. Geological Survey or Air Force Office of Scientific Research is acknowledged but only the "U.S. Government" is measured (rather than the Department of Interior or Defense).

[11] This LLM-based list largely matches that provided by PatentsView (`g_gov_interest_org.tsv`), but also includes subtle differences and improvements. The LLM-based list includes 605 patents that are not in the PatentsView list, and omits 2321 that are. Manual inspection of these cases identifies false positive and negatives. Among the 168,755 patents in both lists, the associated government agencies agree in 95% of cases. Where there are discrepancies, it appears it is usually because the LLM-based approach corrects errors or fills gaps. For example, it accurately tags the patents described in footnote 10 (to DOI and DoD, respectively).

[12] Though the authors provide an indicator for government-assigned patents, we manually reviewed assignees to correct occasional errors—which typically either had foreign government assignees or were U.S. government-assigned but not measured as such because the OCR text was garbled.

[13] We omit patents identified by Fleming et al. (2019) as government-acknowledging for precision, as these are more prone to misclassification (e.g., if a patent has the word "government" in its preamble, we found it sometimes gets tagged as government-acknowledging—even when not), and because the true positives in government-acknowledging patents appear to be accounted for by our other sources (e.g., PatentsView).

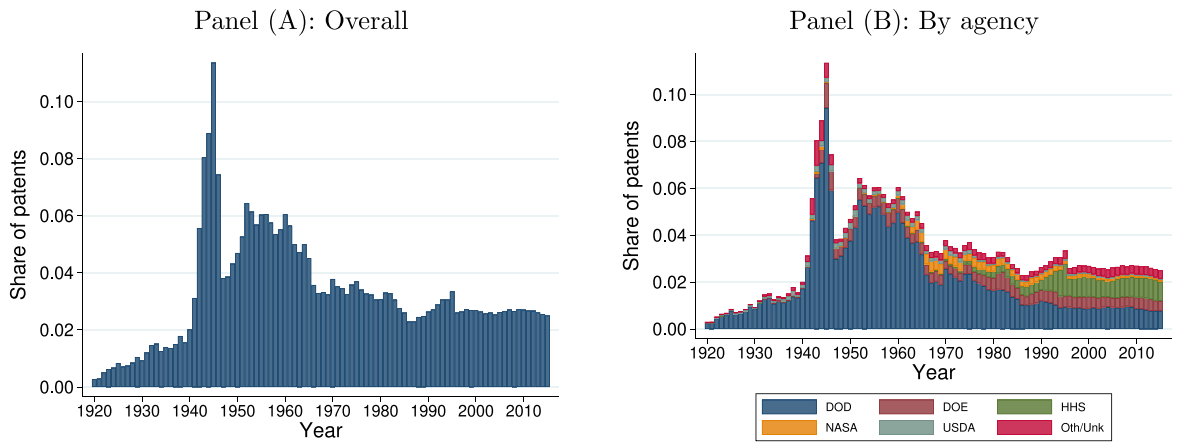

Fig. 4. Share of U.S. patents produced with government funding, overall and by agency, shown by filing year (1920–2015).
Notes: Figure shows the government-funded share of annual U.S. patenting (across filing years), overall (Panel A) and by funding agency for major R&D funders (Panel B).

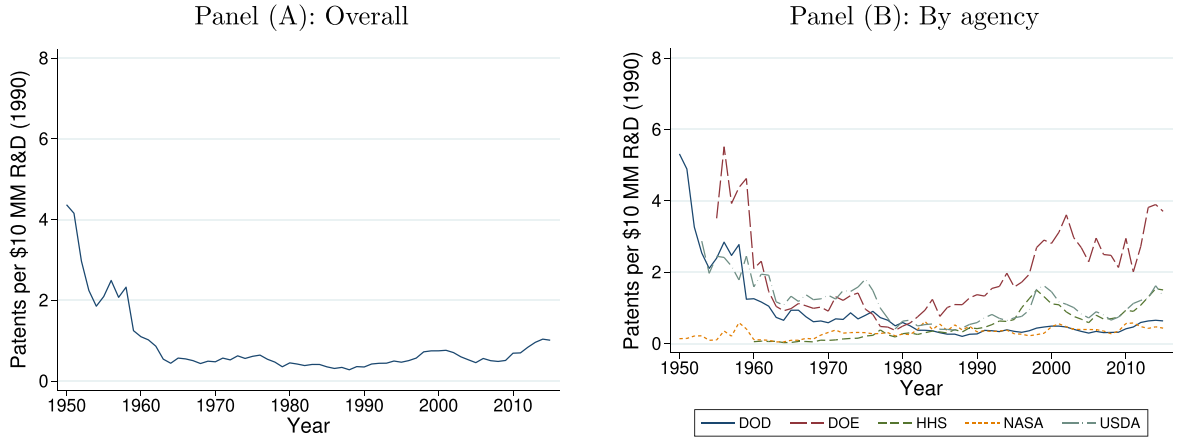

Fig. 5. Government-funded patents per R&D dollar, overall and by agency, shown by filing year (1950–2015).
Notes: Figure plots annual government-funded patents per $10 million dollars of government R&D expenditure (in 1990 dollars), overall (Panel A) and by funding agency for major R&D funders (Panel B).

5. Patterns in the Government Patent Register

Our data reveal several hitherto unseen features of government-funded patents. In this section, we share several patterns we have found in the data, emphasizing those which we think raise or may be useful in studying new research questions. Because this analysis provides only a partial lens into the data this paper provides, which are rich in variation, in Section 7 we suggest other cuts which might yield intriguing findings or raise questions for further study.

5.1. Government-funded patenting

Our starting point is to use the Government Patent Register to evaluate the frequency of government-funded patenting. Fig. 4(A) shows the share of annual U.S. patent filings from 1920 to 2015 which our data identify as government-supported. Contrary to the perception that government-funded technological innovation peaked in the Space Race (1960s), the federal government's share of invention was in fact much higher in World War II, at roughly 11% of USPTO patent filings (see Gross and Sampat, 2023a). It remained elevated through the early years of the Cold War (1950s) at 5%–6% of filings, but has since steadily fallen and by 1990 was around 2% filings—down nearly 70% since the Cold War and 80% since its World War II peak.

Fig. 4(B) breaks this patenting out by funding agency. DoD-funded patents comprised 75%–90% of government-funded patents in every year from 1920 to 1965, but subsequently began to decline. By the early 2000s, the DoD share had fallen to under 30%, roughly matching the growing shares of DOE and HHS. NASA had its peak share of government-funded patents in 1969 (15%), but was even then only a quarter as large as DoD's share at the time.

One advantage of agency-level patent counts (a measure of innovation outputs) is that they can be compared to agency R&D spending (the inputs), which are available from 1949 onwards.[14] Combining the two, we can calculate and examine how efficiently (or at what intensity) each agency converts R&D into patented inventions. Though this is inevitably an imperfect metric — not all public R&D yields inventions, patentable inventions, or patentable inventions that patents are taken out on — it can reveal differences and trends. Fig. 5(A) shows that in the 1950s, government-funded research produced 4 patents for every $10 million in R&D (in 1990 USD). By the mid-1960s, this patent efficiency had declined to roughly 0.5 patents per $10 million, and by 1990 under 0.25 patents per $10 million—a pattern which is directionally similar to that for the overall U.S. economy, albeit lower in magnitude and plateauing earlier (see Appendix Figure B.2 for economy-wide patterns; also see Griliches, 1990). Fig. 5(B) shows

[14] See, for example, historical tables accompanying the President's FY2024 budget. Available (at the time of writing) at https://www.whitehouse.gov/omb/budget/historical-tables/.

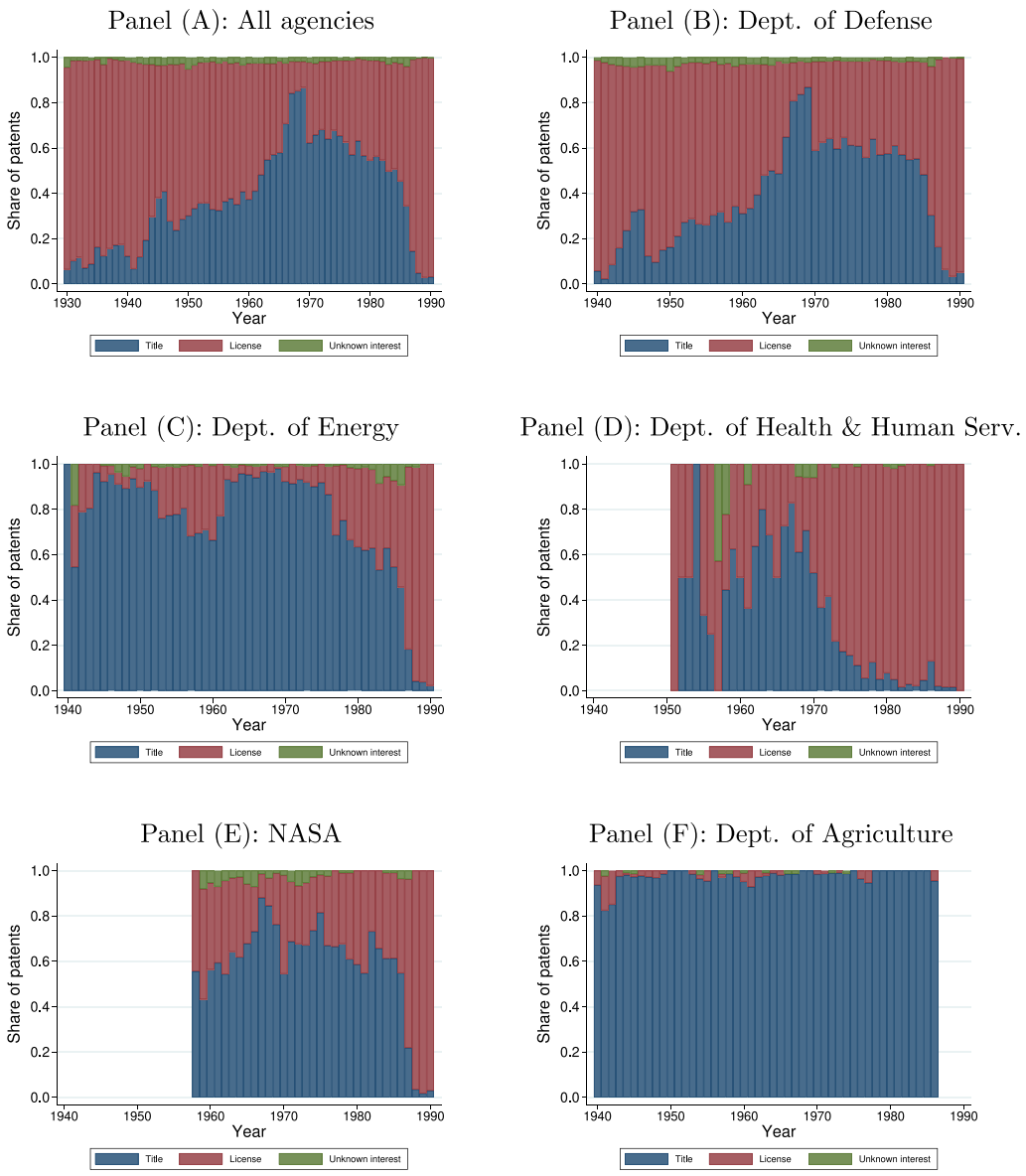

Fig. 6. Title vs. license shares in the historical Government Register, by filing year (1930–1990).
Notes: Figure shows the share of patents in the historical GPR with title, license, and unknown government interest, overall (Panel A) and for select R&D funding agencies (Panels B to F). Sample is restricted to patents with either title or license indicated in the historical GPR (95% of all), and "unknown interest" measures patents with both title and license indicated. Title patents are those to which the funding agency retained ownership; license patents, those to which the R&D contractor or grantee kept title, providing the U.S. government an irrevocable royalty-free license for government use. See text for discussion. The data shown for HHS begin in 1950 (when NIH extramural grants began to grow) and for NASA in 1958 (when the agency was created). The data shown for USDA end in 1986 (after which USDA has very few patents in the historical GPR).

that whereas DoD, DOE, and HHS were relatively active at turning R&D into patents in the 1950s and 1960s, they subsequently converged to the rates of other agencies like NASA, NSF, and USDA. One reason may be that these agencies' research programs grew more basic in nature—though we think this unlikely, as defense R&D (as one example) is generally fairly applied, and responsive to mission needs. To us, Fig. 5 raises questions around what was different in the immediate postwar era that led to more patents per R&D dollar and what can be learned from it today, as well as what is distinctive about DOE, which has since rebounded to its 1960 efficiency.

5.2. Government patent policy

We can also use the Government Patent Register to measure the relative share of title and license policies, overall or at individual agencies. We do so in Fig. 6, focusing on the 1930 to 1990 period, when policies varied across agencies (prior to the Bayh–Dole Act and its expansion). Fig. 6, Panel (A) shows that between 1930 and 1960, a plurality (if not outright majority) of government-funded patents were licensed to (rather than owned by) the U.S. government. The title share rises throughout this period, however, passing 50% in the 1960s, before declining again in the 1980s. After the Bayh–Dole Act in 1981 and its expansion to all federally-funded invention in 1983, patents on which the U.S. government held title were limited to intramural employee government inventions and patents which contractors and grantees chose not to file, but the government did (after first refusal).

Panel (B) shows that these patterns are driven by changes in defense patent policy (due to its large share of government-funded patents in this era). Panels (C) to (F) illustrate the patent policies of other agencies, highlighting cross-agency variation. We can visibly see DOE and HHS transitioning from title to license policies, and USDA holding title to most patents it funded throughout this era (because it performed much of its R&D intramurally).

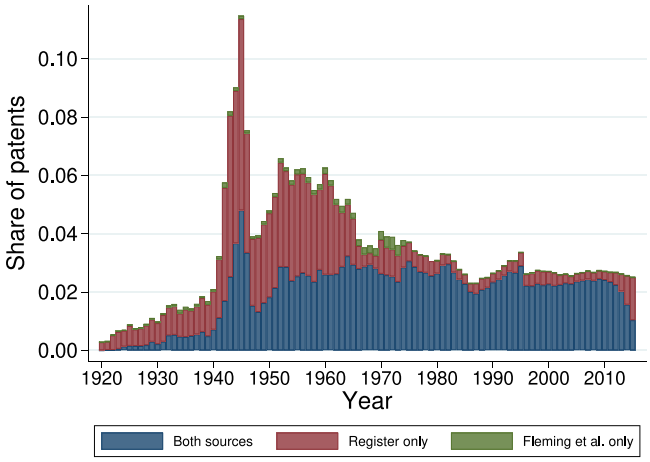

Fig. 7. Comparison of historical Government Register and Fleming et al. data, by filing year (1920–2015).
Notes: Figure compares our data to previous measures of government interest patents from Fleming et al. (2019). We reduce differences against the Fleming et al. data in the pre-1976 historical GPR by measuring government interest patents via patent assignees and in-text interest statements. Remaining differences are due to false positives in the Fleming et al. measures. See text for discussion.

5.3. Comparisons to previous measures

As noted in Sections 1 and 4, past efforts to measure government-funded patenting have used the contents of published patents to do so, via assignees and interest statements. To evaluate the Government Patent Register we have compiled and what new information it may (or may not) offer, we found it useful to compare it against existing measures. In principle, administrative data like the historical GPR and the UPAD may provide more complete, and more precise, measurement than observational approaches which rely on voluntary disclosure or compliance with interest statement reporting and are more subject to error (both over- and under-classification) in algorithmic measurement methods, even in complete samples.

We focus our comparisons to data from Fleming et al. (2019), who have produced the longest time-series to date (extending back to 1926, using freshly-OCRed patent text). We count government interest patents in (i) the historical GPR, (ii) the Fleming et al. data, and (iii) both sources from 1930 to 1990. We do so for all patents and separately for title and license patents, interpreting the Fleming et al. measures as indicating title if the patent is government-owned (i.e., has a government assignee), and as license if acknowledging government support (i.e., has an interest statement). For this exercise, we treat patents in each source as title patents if both title and license are indicated, which occurs rarely in the historical GPR (95% of title patents only have title indicated), and frequently in the Fleming et al. data (50% of government-assigned patents also have an interest statement explicitly claiming government rights in the invention).

Fig. 7 shows how these sources compare, plotting the government interest share of patents in one, the other, or both. The historical GPR fills significant gaps in what is available under the Fleming et al. (2019) approach—especially in the mid-century (through roughly 1970). Further investigation reveals why: though these sources largely overlap in measuring government-assigned patents, the historical GPR data measure many more patents which were contractor-owned and government-licensed. The magnitude of the difference is often substantial: in some years, the historical GPR measures up to 15 times as many license patents as can be measured through patent text alone—most likely because interest statements were not widely used at the time. Casual inspection of a few patents where these sources disagree reinforces this interpretation.

There are several implications of this evidence. Prior to the collection of these data, the government footprint in technology, for example, was significantly underestimated. This can be seen in Appendix Figure B.1, which reproduces the main chart of Fleming et al. (2019) with our updated data: we find that a substantially larger fraction of postwar U.S. patents were related to government-funded science and invention than was previously known. These data present more opportunities to study what technology was being developed with public support at this time, and what impacts that has had since—including today. Patent productivity of public R&D (e.g., Fig. 5) would be significantly underestimated for this period using measures derived from assignments and interest statements alone, and indeed, in separate analysis we have found that it would look quite similar to patent productivity of public R&D today — but with our newly-collected data, we can see differences — raising questions over what has changed.

6. Distribution dataset

The key dataset accompanying this paper is the "Government Patent Register" described in Section 4, which lists U.S. patents issued between 1836 and 2020 and merges in the historical GPR, UPAD, PatentsView, and assignee-based measures. We include source-specific indicators of title and license patents and funding agencies, as well as overall indicators (aggregating across sources). The data are posted to the Harvard Dataverse (Gross and Sampat, 2024).[15]

Table 1 counts how many of the 275,000 government-funded patents in this dataset appear in every combination of sources, showing what each source contributes to measurement individually and jointly. We split the sample into the pre- and post-1976 eras to compare periods before and after electronic data become available. Of the nearly 90,000 government interest patents pre-1976, 65% are only found in the historical GPR. Another 32% are in both the historical GPR and our assignee-based measures, and 3% in assignee-based measures only. Post-1976, our sources diversify and include more redundancy. Fig. 8 complements this table, showing the incremental contribution that each data source makes to our measurement of government interest patents, showing that the historical GPR, UPAD, and PatentsView data can account for nearly all known cases.

7. Use cases, caveats, and concluding remarks

The Government Patent Register has the potential to open up new opportunities for research on the development and impacts of R&D policy on the U.S. innovation system with comprehensive, administrative, long-run data. A historical lens is not only helpful for understanding the evolution of U.S. innovation: it can also be a source of natural experiments that can inform current practice. There are relatively few sources of longitudinal data on U.S. public R&D investments beyond broad aggregate data such as that provided in the NSF's annual "Federal Funds for Research and Development" volume. Moreover, despite recurring debates around government patent policy, and tensions between incentives for firms and scientists to engage in publicly-funded R&D vs. policy goals of securing the benefits of publicly-funded research for the public, the impacts of government patent policy have not been systematically evaluated with government-wide data or harnessing the rich policy variation in the postwar era. While the downsides of patent data are well-known (among them are that not all patents are inventions, not all inventions are patents, and the propensity to patent can vary across fields, agencies, and time), patent documents provide rich information on inventive activity. Beyond standard "front page" information

[15] With this paper we are also releasing several other data resources, including a digitized version of the historical GPR. We also provide datasets derived from other sources. We complement these with code and instructions for future extensions of measures we produce from the UPAD and from PatentsView data files, which are regularly updated by their creators.

Table 1
Government interest patent counts from 1920 to 2020, by data source.

Data source combinations					Associated patents					
	Historical				All years		Pre-1976		Post-1976	
	Register	UPAD	PatentsView	Assignee data	Patents	Share	Patents	Share	Patents	Share
1.	Y	–	–	–	60,849	22.2%	57,057	64.6%	3792	2.0%
2.	Y	–	–	Y	30,898	11.3%	27,608	31.3%	3290	1.8%
3.	Y	–	Y	–	8305	3.0%	0	0.0%	8305	4.5%
4.	Y	–	Y	Y	3795	1.4%	0	0.0%	3795	2.0%
5.	Y	Y	–	–	361	0.1%	10	0.0%	351	0.2%
6.	Y	Y	–	Y	1314	0.5%	0	0.0%	1314	0.7%
7.	Y	Y	Y	–	901	0.3%	0	0.0%	901	0.5%
8.	Y	Y	Y	Y	3406	1.2%	0	0.0%	3406	1.8%
9.	–	–	–	Y	6278	2.3%	2774	3.1%	3504	1.9%
10.	–	–	Y	–	63,978	23.3%	0	0.0%	63,978	34.5%
11.	–	–	Y	Y	3731	1.4%	0	0.0%	3731	2.0%
12.	–	Y	–	–	9876	3.6%	34	0.0%	9842	5.3%
13.	–	Y	–	Y	11,632	4.2%	0	0.0%	11,632	6.3%
14.	–	Y	Y	–	48,943	17.9%	0	0.0%	48,943	26.4%
15.	–	Y	Y	Y	18,927	6.9%	0	0.0%	18,927	10.2%
16.	–	–	–	–	818	0.3%	818	0.9%	0	0.0%
Total					*274,012*	*100.0%*	*88,301*	*100.0%*	*185,711*	*100.0%*

Notes: Table lists counts of government interest patents in our data identified by different combinations of data sources, illustrating where these data sources are overlapping versus additive. The 818 patents included in our data but not measured in these sources were identified by manually reviewing pre-1976 patents which were measured by Fleming et al. (2019) as having an interest statement in the patent text. We separately list totals before and after 1976, when USPTO's electronic records begin. Roughly 20% of all government interest patents, and over 50% of pre-1976 government interest patents, are identifiable only by the historical GPR (as seen in the first row of the table).

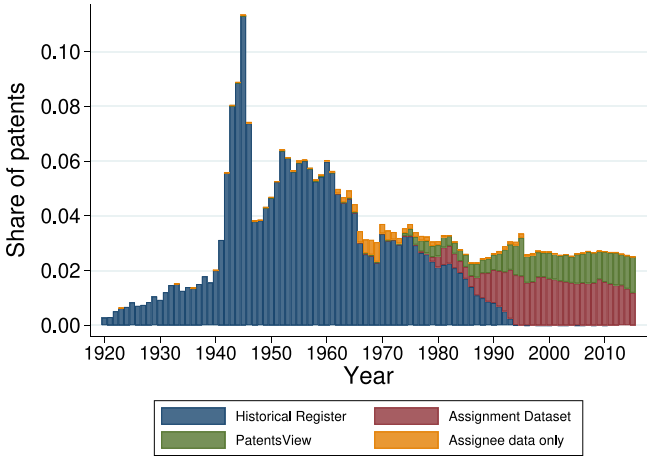

Fig. 8. Government interest patents contributed by each data source, by filing year (1920–2015).
Notes: Figure tallies the government interest patents successively added by each of the following sources: the historical GPR, the UPAD, PatentsView government interest statement data, and assignee-based measures (based on Fleming et al. (2019) assignees pre-1976, and PatentsView assignees post-1976).

(e.g., assignee, class, citations), the rich text in patent documents (in-text citations, topics), now extractable via computational methods, may make our data even more valuable.

These data could be used to study a wide range of questions—such as to examine the determinants of government R&D investment, or to study the composition and evolution of individual agencies' R&D portfolios. Changes in patent policy can also be related to participation in the federal R&D enterprise, the technological fruits of federally-funded research (in the spirit of De Rassenfosse et al., 2019), or the commercialization of federally-funded invention. The data can also be used as a control variable: in our own research examining the long-run effects of World War II R&D on the U.S. innovation system (Gross and Sampat, 2023a), we used them to control for postwar publicly-funded R&D at geographically disaggregated levels. Similarly, these data may be useful in other historical exercises assessing specific R&D shocks, especially when DoD research is a focus (or a potential confounder), since standard patent assignment data severely undercounts DoD-financed patenting in the 1950s and 1960s.

These data may also be useful in descriptive and/or causal analyses examining how government funded inventions percolate through the innovation system, and the division of labor between the public and private sectors. What topics does each sector specialize in? How often are government-funded inventions novel or disruptive, relative to private inventions? How do shocks to agency funding affect government-funded patenting, and what are the effects of these patents on private patenting in similar topics or adjacent regions? These questions have held long-running interest in academic research and applied research policy, and have recently begun to attract new, detailed analysis (e.g., Fieldhouse and Mertens, 2023a,b; Dyevre, 2023) which these data may complement or support.

Other important questions relate to the determinants of knowledge diffusion spanning the public and private sectors. These data are uniquely useful for assessing how patent policies may influence diffusion, since it covers the era (before Bayh–Dole) when there was cross-agency variation in patent policy. In addition, it tracks both "title" and "license" patents within agencies, which is useful since there were sometimes exceptions to, and procedures to get waivers around, some agencies' nominal patent policies (Eisenberg, 1996). Of course, measuring diffusion, commercialization, and impact is difficult, but new advances along other dimensions — such as access to full text patent documents, full-text scientific articles, technology transfer office data, and natural language processing tools — may help facilitate such analyses (e.g., Masclans et al., 2023).

In addition to new patent data, the tools and methods we develop (e.g., for identifying government interests in the USPTO's Patent Assignment Dataset (UPAD) or extracting funding agencies from government interest statements using large language models) will make it straightforward to extend these data in the future, as new questions present themselves. The data and code accompanying this paper include pre-set programs and instructions for doing so.

Like all patent data, the Government Patent Register has its limits. Most importantly, it captures patents, not government R&D spending

(the input) or inventions (an output of R&D, and input to patents). In addition to the well-known fact that not all inventions are patentable (Griliches, 1990), there is likely significant variation in the propensity to patent across agencies, fields, and time. For example, there were historically strong norms against patenting publicly-funded medical research, so it may provide a distorted lens on historical NIH investment patterns. In addition, just as there appears to be significant underreporting of government interest statements in patent text today (Rai and Sampat, 2012) and even in the UPAD (Graham et al., 2018), there was likely some non-compliance with E.O. 9424 as well. Compliance rates are unknown historically, but for modern data, one can triangulate information across multiple sources, including government interest statements, assignment data, and disclosure to funding agencies—which, as Table 1 shows, can help fill in gaps left by any one source alone. More fully understanding compliance with reporting requirements is useful for research using data on government-funded patents. Though it is possible our data are missing some examples — there may be cases not reported anywhere — we believe the data significantly expand on prior sources.

CRediT authorship contribution statement

Daniel P. Gross: Writing – review & editing, Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Bhaven N. Sampat:** Writing – review & editing, Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.respol.2024.105142.

Data availability

Data presented in this article are available from the Harvard Dataverse at https://doi.org/10.7910/DVN/BDCCN9.

References

Anderson, Gary, Moris, Francisco, 2023. Federally funded R&D declines as a share of GDP and total R&D. National Center for Science and Engineering Statistics (NCSES), Report no. NSF 23-339, available at https://ncses.nsf.gov/pubs/nsf23339/.

Azoulay, Pierre, Graff Zivin, Joshua S., Li, Danielle, Sampat, Bhaven N., 2019. Public R&D investments and private-sector patenting: Evidence from NIH funding rules. Rev. Econ. Stud. 86 (1), 117–152.

Commission, National Patent Planning, 1941. Second Report of the National Patent Planning Commission. Government Printing Office, Washington.

De Rassenfosse, Gaétan, Jaffe, Adam, Raiteri, Emilio, 2019. The procurement of innovation by the US government. PLoS One 14 (8), e0218927.

Dyevre, Arnaud, 2023. Public R&D spillovers and productivity growth. Working paper.

Eisenberg, Rebecca S., 1996. Public research and private development: Patents and technology transfer in government–sponsored research. Va. L. Rev. 82, 1663.

Federal Council for Science and Technology, 1976. Report on Government Patent Policy: Combined December 31, 1973 and December 31, 1974 and December 31, 1975 and September 30, 1976. Government Printing Office, Washington.

Fieldhouse, Andrew J., Mertens, Karel, 2023a. A narrative analysis of federal appropriations for research and development. Working paper.

Fieldhouse, Andrew J., Mertens, Karel, 2023b. The returns to government R&D: Evidence from U.S. appropriations shocks. Working paper.

Flamm, Kenneth, 1987. Targeting the Computer: Government Support and International Competition. Brookings Institution Press, Washington.

Flamm, Kenneth, 1988. Creating the computer: Government, industry, and high technology. Brookings Institution Press, Washington.

Fleming, Lee, Greene, Hillary, Li, G, Marx, Matt, Yao, Dennis, 2019. Government-funded research increasingly fuels innovation. Science 364 (6446), 1139–1141.

Graham, Stuart J.H., Marco, Alan C., Myers, Amanda F., 2018. Patent transactions in the marketplace: Lessons from the USPTO patent assignment dataset. J. Econ. Manag. Strategy 27 (3), 343–371.

Griliches, Zvi, 1990. Patent statistics as economic indicators: A survey. J. Econ. Lit. 28 (4), 1661–1707.

Gross, Daniel P., Sampat, Bhaven N., 2023a. America, jump-started: World war II R&D and the takeoff of the US innovation system. Amer. Econ. Rev. 113 (12), 3323–3356.

Gross, Daniel P., Sampat, Bhaven N., 2023b. The world war II crisis innovation model: What was it, and where does it apply? Res. Policy 52 (9), 104845.

Gross, Daniel P., Sampat, Bhaven N., 2024. The Government Patent Register: Data and Documentation. Harvard Dataverse, http://dx.doi.org/10.7910/DVN/BDCCN9.

Kevles, Daniel J., 1977. The National Science Foundation and the debate over postwar research policy, 1942–1945: A political interpretation of science–the endless frontier. Isis 68 (1), 5–26.

Lin, Zihang, Yin, Yian, Liu, Lu, Wang, Dashun, 2023. SciSciNet: A large-scale open data lake for the science of science research. Sci. Data 10 (1), 315.

Marco, Alan C, Myers, Amanda, Graham, Stuart, D'Agostino, Paul, Apple, Kirsten, 2015. The USPTO Patent Assignment Dataset: Descriptions and analysis. SSRN working paper, available at https://ssrn.com/abstract=2849634.

Marco, Alan C., Vishnubhukat, Saurabh, 2015. USPTO government interest patents. Mimeo.

Masclans, Roger, Hasan, Sharique, Cohen, Wesley M., 2023. The commercial potential of science and its realization: Evidence from a measure using a large language model. Available at SSRN: https://ssrn.com/abstract=4481928.

Mowery, David C., Rosenberg, Nathan, 1982. Government policy and innovation in the commercial aircraft industry, 1925–1975. In: Nelson, Richard R. (Ed.), Government and Technical Progress: A Cross-Industry Analysis. Pergamon Press, New York.

Narin, Francis, 1976. Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity. Computer Horizons, Inc..

Rai, Arti K., Sampat, Bhaven N., 2012. Accountability in patenting of federally funded research. Nature Biotechnol. 30 (10), 953–956.

Sampat, Bhaven N., 2020. Whose drugs are these? Issues Sci. Technol. 36 (4), 42–48.

Sampat, Bhaven N., Lichtenberg, Frank R., 2011. What are the respective roles of the public and private sectors in pharmaceutical innovation? Health Aff. 30 (2), 332–339.

U.S. Congress Joint Committee on Atomic Energy, 1959. Selected Materials on Atomic Energy Patents. Government Printing Office, Washington.

U.S. House of Representatives, Committee on Science and Astronautics, 1966. An Evaluation of the Patent Policies of the National Aeronautics and Space Administration. Government Printing Office, Washington.

U.S. Senate Judiciary Committee, Subcommittee on Patents, Trademarks, and Copyrights, 1959. Patent Practices of the Government Patents Board. Government Printing Office, Washington.

U.S. Senate Judiciary Committee, Subcommittee on Patents, Trademarks, and Copyrights, 1961. Patent Practices of the Department of Defense. Government Printing Office, Washington.

Watson, Donald Stevenson, Holman, Mary A., 1964. Patents from government-financed research and developments. Pat. Trademark Copy. J. Res. Ed. 8, 199.

Watson, Donald Stevenson, Holman, Mary A., 1966. The federal government's propensity to patent. Pat. Trademark Copy. J. Res. Ed. 10, 61.

Watson, Donald Stevenson, Holman, Mary A., 1967. Concentration of patents from government financed research in industry. Rev. Econ. Stat. 375–381.