NBER WORKING PAPER SERIES

THE GOVERNMENT PATENT REGISTER: A NEW RESOURCE FOR MEASURING U.S. GOVERNMENT-FUNDED PATENTING

Daniel P. Gross
Bhaven N. Sampat

Working Paper 32136
http://www.nber.org/papers/w32136

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2024, Revised March 2024

We are grateful for helpful comments 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 under Grant No. 1951470, which funded the digitizing of the historical data and creation of a public dataset. The data repository accompanying 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Government Patent Register: A New Resource for Measuring U.S. Government-Funded Patenting

Daniel P. Gross and Bhaven N. Sampat

NBER Working Paper No. 32136
February 2024, Revised March 2024
JEL No. N42,N72,O31,O34,O38

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.

Daniel P. Gross
Fuqua School of Business
Duke University
100 Fuqua Drive
Durham, NC 27708
and NBER
daniel.gross@duke.edu

Bhaven N. Sampat
Arizona State University
Consortium for Science, Policy, and Outcomes
1800 I Street NW, Suite 300
Washington, DC 20006
and NBER
bhaven.sampat@asu.edu

1 Introduction

Since World War II, the U.S. federal government has been the world's largest funder of research and development. In 2022, the U.S. government spent nearly $200 billion on research and development (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 (the era of electronic records), limiting what has been learned.

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 ("Government Register"), which we digitize from USPTO's archival records. The historical register was for most of the twentieth century the U.S. government's official record of patents which 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, which 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 coverage in 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.[1] Numerous publicly-funded patents

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

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 Government Register 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. 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 (both title and license) to U.S. government agencies.

Comparisons against existing sources (e.g., Fleming et al. 2019 or PatentsView) suggest that although the Government Register 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 or modern USPTO register. We thus incorporate data from Fleming et al. (2019) and PatentsView to measure government-interest patents independently of the USPTO register, 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 precise 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 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 register, 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 Government Register (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" (National Patent Planning 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 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 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.

2.2 The Evolution of Government Patent Policy

With the explosion of federal R&D during World War II, most through contracts to universities, research organizations, 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 the necessary research equipment and personnel for the contractual research, and for research performed under OSRD 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 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 (e.g., radar); 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 National Science Foundation (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) recounts, DoD and NSF 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 subsumed by 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. The National Aeronautics and Space Administration (NASA), founded in 1958, 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 title policy patents which reverted to government ownership if the contractor or grantee chose not to file, or petitions for title from contractors of license 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.

Debates over title versus license policies continued throughout this era, similar to "short-form" vs. "long- form" debates in World War II. There were various unsuccessful attempts to create a "uniform" patent policy across agencies, including through President Kennedy's Memorandum on Government Patent Policy in 1963, President Nixon's memorandum in 1971, and through several

associated pieces of legislation and executive regulations.

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 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 suffers from significant non-compliance (Rai and Sampat 2012).

The history of government patent policy 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 to examine patent policy—it has largely been passed over since.[2] 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).

These studies, in tandem with Executive Order 9424, revealed to us the Government Register's existence, and suggested where we might find it: at USPTO. 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

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

(mis)titled as "Index to Patent Assignments by Government Licensees, 1/1/1890-12/31/1955".[3] Appendix Figure A.1 provides a 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.

Figure 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).[4] 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).

[Figure 1 about here]

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 consistent (and 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,

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

[4]Government rights in employee inventions were first established by legislation in 1883 (P.L. 47-103, labeled in the Government Register 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.

which we interpret more as public service than a contractual legal interest. These special cases comprise only a small fraction (3%) of index cards in the register.

Even with accurate transcription, a residual challenge is the possibility that the physical index cards themselves may sometimes have typographical errors. 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 are seen in Figure 2.

[Figure 2 about here]

The final data we construct from the historical Government Register 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+ associated agencies, and 185 (0.2%) have no agency listed. Figure 3(A) provides the share of patents in the data associated with each of these agencies through 1975.

[Figure 3 about here]

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 recorded 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 Government Register begins to dwindle in the 1980s (declining from 1,650 patents in 1980, to 870 in 1990, to 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 register 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). We supplement this with data from PatentsView on patents with government interest statements (Jones and Madhavan 2020), 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 managed the historical register), includes patent transactions since 1980, and is currently updated annually. As Graham et al. (2018) explain, the UPAD should subsume 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".

Because we would like to build a "modern Government 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 Government Register 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 (including assignments—not just confirmatory licenses), and manually categorized these conveyances as title or license. Putting these pieces together, we can classify transactions as (i) conveying interest to a government entity, and (ii) whether that interest is title or license.

We think 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 register, 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.[5]

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`), as well as a separate file listing government organizations extracted from these interest statements (`g_gov_interest_org.tsv`).[6] Patents in these lists only partially coincide with the historical Government Register and the USPTO's Patent Assignment Dataset, in part because interest statements are not always not consistently included in government interest patents. PatentsView data are nevertheless useful as a complement to other sources, since each includes patents which others omit (we provide statistics in Section 6).

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 register and the Patent Assignment Dataset). After probing the data and comparing them to the underlying interest statements, we observed they mostly agree, but also noticed occasional imprecisions and discrepancies.[7] 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, and prompting it to identify the U.S. federal agencies acknowledged, or the U.S. government (generically) if it the statement claims a U.S. government interest but no specific agency is named or thanked for funding, and omitting all other named entities.

After running this prompt on all ∼171,000 patents in the PatentsView file (using the OpenAI API), we received back a list of government funders. We manually processed this list to map named entities to our focal agencies—while weeding out foreign and state governments and private organizations 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,

[5] URL: https://assignment.uspto.gov/patent/index.html, last accessed February 2024.
[6] These files are available at https://patentsview.org/download/data-download-tables.
[7] For example, patents where the U.S. Geological Survey (USGS) or Air Force Office of Scientific Research (AFOSR) is acknowledged but only the "U.S. Government" is measured (rather than the Department of Interior or Defense), or a patent where the Federal Bureau of Investigation is crosswalked to DoD rather than DOJ.

under which we identify 169,360 patents with a U.S. government interest.[8] Additional advantages to this approach are that it is cheap (at a cost of about $1.50 per 1,000 interest statements processed) 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.[9,10]

4.4 A "Government Patent Register" for research

We combine these sources to create a Government Patent Register that runs from the early twentieth century (the beginning of the historical register) 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 more fully 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 Government Register, (ii) the Patent Assignment Dataset, (iii) government interest statements

[8] 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 2,321 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 7 (to DOI, DoD, and DOJ respectively).

[9] Though the authors provide a patent-level indicator for government-assigned patents, we manually reviewed assignees to correct occasional errors in their dataset—primarily cases with foreign government assignees or which are U.S. government-assigned but were not measured as such because the OCR text was garbled.

[10] To prioritize precision, we also choose to not use the Fleming et al. (2019) government acknowledgment indicator unless it is duplicated by other sources. Through extensive manual checks we have found that the occasional cases where only Fleming et al. (2019) (and not our other sources) indicate government acknowledgement tend to be false positives, caused by the presence of specific words in the patent description that get misclassified. For example, false positives occur in cases where the patent description describes the "government" as a potential user, or when it explicitly explains that the government has no legal interest.

(from PatentsView), and (iv) government assignees (from Fleming et al. 2019 and PatentsView). Triangulating across sources, we measure government interest patents as those which are identified via any of these methods.[11] 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. Figure 3 shows the agency distribution of government interest patents, pre-1976 (left panel, based on the historical Government Register 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.

[Figure 3 about here]

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. Figure 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)

[11] 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).

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.

[Figure 4 about here]

Figure 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.[12] 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. Figure 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. Figure 5(B) shows that whereas DoD, DOE, and HHS were relatively efficient at turning R&D into patents in the 1950s and 1960s, they subsequently converged to the (in)efficiency 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, Figure 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.

[Figure 5 about here]

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 Figure 6, focusing on the 1930 to 1990 period,

[12] 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/.

when policies varied across agencies (prior to the Bayh-Dole Act and its expansion), and when our measurement of title and license patents is most complete. Figure 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).

[Figure 6 about here]

5.3 Comparisons to Previous Measures

As we noted in Sections 1 and 4, previous 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 in this paper 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 Government Register and the Patent Assignment Dataset may provide more complete, and more precise, measurement than observational approaches which rely on voluntary disclosure or compliance with interest statement reporting (both of which are incomplete, especially historically) and are more subject to error (both over- and under-classification) in algorithmic measurement methods, even in complete samples.

We focus our comparisons to the data provided by Fleming et al. (2019), who have produced the longest time-series to date (extending back to 1926, using freshly-OCRed patent text). Using the authors' data (see Section 4), we count government interest patents in (i) the historical Government Register, (ii) the Fleming et al. (2019) data, and (iii) both sources over the 1930-1990 period. We do so for all patents and separately for title and license patents, interpreting the Fleming et al. (2019) 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 register (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).

Figure 7 shows how these sources compare, plotting the government interest share of patents in one, the other, or both. The historical register fills significant gaps in what is available under the Fleming et al. approach—especially in the mid-century (through roughly 1970; Panel A). Panels (B) and (C) reveal why: though these sources largely overlap in measuring government-assigned patents, the register 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 register 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.

[Figure 7 about here]

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., Figure 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" which we described in Section 4, which lists U.S. patents issued between 1836 and 2020 and merges in data from the

historical Government Register, the modern Patent Assignment Dataset, PatentsView, and Fleming et al. (2019). We include source-specific indicators of title and license patents and funding agencies, as well as overall indicators (aggregating across sources). The data and documentation are posted to the Harvard Dataverse (see Gross and Sampat 2024 for a link).

As we noted above, the dataset contains 275,000 government-funded patents through 2020. Table 1 counts how many of these patents appear in every combination of sources, showing what each of these sources contributes to measurement individually and where we have redundancy. 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 Government Register. Another 32% are in both the historical register and our assignee-based measures, and 3% in assignee-based measures only. Post-1976, our sources diversify and include significant redundancy: 56% of government interest patents are measured by two or more sources. Of post-1976 patents, 14% are in the historical register, 51% in the Patent Assignment Dataset, 82% in PatentsView, and 27% in our assignee measures.

[Table 1 about here]

With this paper we are also releasing several other data resources. Chief among these is a digitized version of the historical Government Register. We provide the original data and a polished dataset produced from them. We also provide datasets derived from other sources. We complement these with code and instructions for future extensions of the measures we have produced from the Patent Assignment Dataset and from PatentsView data files.

[Figure 8 about here]

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 sources of information on inventive activity. Beyond standard "front page" information (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 evaluate its effects. They can also be used 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 previous research examining the long-run effects of World War II R&D on the U.S. innovation system (Gross and Sampat 2023a), we used the historical Government Register data to control for patterns in postwar publicly-funded R&D at disaggregated levels. Similarly, these data may be useful for other historical exercises assessing specific R&D shocks. They are particularly useful in contexts where DoD research activity is the focus, or is a potential confounder: as discussed, standard patent assignment data would severely undercount 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 in U.S. innovation. What topics does the government specialize in versus the private sector? How often are government-funded inventions novel or disruptive, relative to private sector inventions? How do shocks to agency funding affect government-funded patenting, and what are the effects of these patents on private patenting on 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 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—and 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 1998), 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 in government interest statements today (Rai and Sampat 2012) and even in the USPTO's Patent Assignment Dataset (Graham et al. 2018), there was likely some non-compliance with Executive Order 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 nevertheless useful for research using historical or modern data on government-funded patents. Though it is possible our data are missing some government-funded patents—there may be cases not reported anywhere—we believe the data are precise (with few false positives) and significantly expand on prior sources.

References

Anderson, Gary and Francisco Moris. 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/.

De Rassenfosse, Gaétan, Adam Jaffe, and Emilio Raiteri. 2019. "The procurement of innovation by the US government," *PloS one*, Vol. 14, No. 8, p. 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.*, Vol. 82, p. 1663.

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

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

——— . 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*. Washington: Brookings Institution Press.

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

Fleming, Lee, Hillary Greene, G Li, Matt Marx, and Dennis Yao. 2019. "Government-funded research increasingly fuels innovation," *Science*, Vol. 364, No. 6446, pp. 1139–1141.

Graham, Stuart JH, Alan C Marco, and Amanda F Myers. 2018. "Patent transactions in the marketplace: Lessons from the USPTO patent assignment dataset," *Journal of Economics & Management Strategy*, Vol. 27, No. 3, pp. 343–371.

Griliches, Zvi. 1998. "Patent statistics as economic indicators: A survey," in Griliches, Zvi ed. *R&D and productivity: The econometric evidence*, Chicago: University of Chicago Press, pp. 287–343.

Gross, Daniel P and Bhaven N Sampat. 2023a. "America, jump-started: World War II R&D and the takeoff of the US innovation system," *American Economic Review*, Vol. 113, No. 12, pp. 3323–3356.

——— . 2023b. "The World War II crisis innovation model: What was it, and where does it apply?" *Research Policy*, Vol. 52, No. 9, p. 104845.

——— . 2024. *The Government Patent Register: Data and Documentation*. Harvard Dataverse, DOI: 10.7910/DVN/BDCCN9.

Jones, Christina and Sarvothaman Madhavan. 2020. "PatentsView Government Interest Extraction and Processing–Version 2.0,"Technical report, American Institutes for Research.

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*, Vol. 68, No. 1, pp. 5–26.

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

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

Masclans, Roger, Sharique Hasan, and Wesley M Cohen. 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. and Nathan Rosenberg. 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*, New York: Pergamon Press.

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

Rai, Arti K and Bhaven N Sampat. 2012. "Accountability in patenting of federally funded research," *Nature Biotechnology*, Vol. 30, No. 10, pp. 953–956.

Sampat, Bhaven N. 2020. "Whose drugs are these?" *Issues in Science and Technology*, Vol. 36, No. 4, pp. 42–48.

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

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

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

——— . 1961. *Patent Practices of the Department of Defense*. Washington: Government Printing Office.

Watson, Donald Stevenson and Mary A Holman. 1964. "Patents from Government-Financed Research and Developments," *Pat. Trademark & Copy. J. Res. & Ed.*, Vol. 8, p. 199.

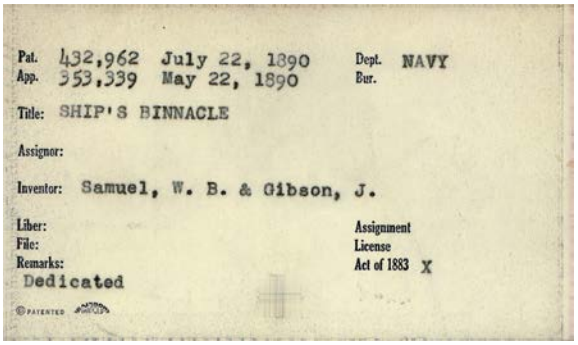
——— . 1966. "The Federal Government's Propensity to Patent," *Pat. Trademark & Copy. J. Res. & Ed.*, Vol. 10, p. 61.

——— . 1967. "Concentration of patents from government financed research in industry," *The Review of Economics and Statistics*, pp. 375–381.

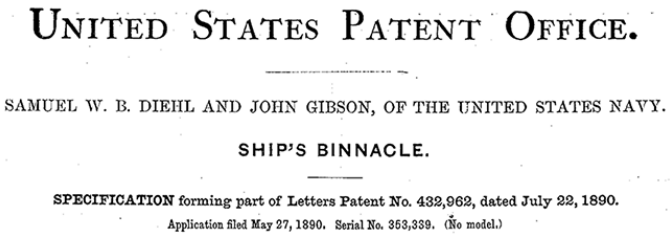
Figure 1: Example index cards in the historical Government Register

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

Figure 2: First and last patents in the historical Government Register



Pat. 432,962 July 22, 1890 Dept. NAVY
App. 353,339 May 22, 1890 Bur.
Title: SHIP'S BINNACLE
Assignor:
Inventor: Samuel, W. B. & Gibson, J.
Liber: Assignment
File: License
Remarks: Act of 1883 X
Dedicated



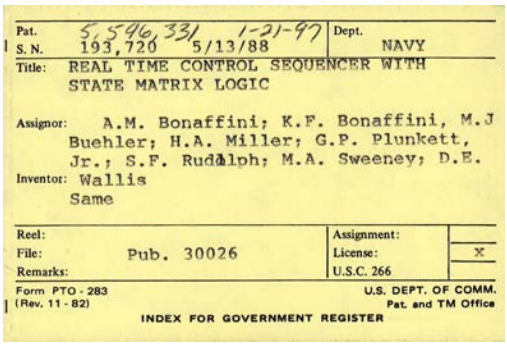
UNITED STATES PATENT OFFICE.

SAMUEL W. B. DIEHL AND JOHN GIBSON, OF THE UNITED STATES NAVY.

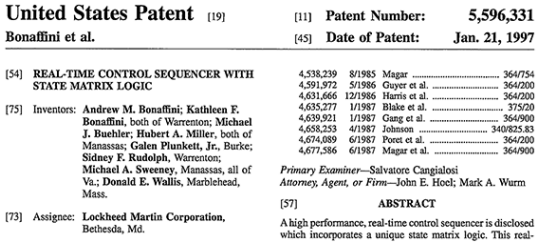
SHIP'S BINNACLE.

SPECIFICATION forming part of Letters Patent No. 432,962, dated July 22, 1890.

Application filed May 27, 1890. Serial No. 353,339. (No model.)



Pat. 5,596,331 1-21-97
S. N. 193,720 5/13/88 Dept. NAVY
Title: REAL TIME CONTROL SEQUENCER WITH STATE MATRIX LOGIC
Assignor: A.M. Bonaffini; K.F. Bonaffini, M.J Buehler; H.A. Miller; G.P. Plunkett, Jr.; S.F. Rudolph; M.A. Sweeney; D.E.
Inventor: Wallis
Same
Reel: Assignment:
File: Pub. 30026 License: X
Remarks: U.S.C. 266
Form PTO - 283 (Rev. 11 - 82) U.S. DEPT. OF COMM. Pat. and TM Office
INDEX FOR GOVERNMENT REGISTER



United States Patent [19]
Bonaffini et al.

[11] Patent Number: 5,596,331
[45] Date of Patent: Jan. 21, 1997

[54] REAL-TIME CONTROL SEQUENCER WITH STATE MATRIX LOGIC

[75] Inventors: Andrew M. Bonaffini; Kathleen F. Bonaffini, both of Warrenton; Michael J. Buehler; Hubert A. Miller, both of Manassas; Galen Plunkett, Jr., Burke; Sidney F. Rudolph, Warrenton; Michael A. Sweeney, Manassas, all of Va.; Donald E. Wallis, Marblehead, Mass.

[73] Assignee: Lockheed Martin Corporation, Bethesda, Md.

4,538,239	8/1985	Magar	364/754
4,591,972	5/1986	Guyer et al.	364/200
4,631,666	12/1986	Harris et al.	364/200
4,635,277	1/1987	Blake et al.	375/20
4,639,921	1/1987	Gang et al.	364/900
4,658,253	4/1987	Johnson	340/825.83
4,674,089	6/1987	Poret et al.	364/200
4,677,586	6/1987	Magar et al.	364/900

Primary Examiner—Salvatore Cangialosi
Attorney, Agent, or Firm—John E. Hoel; Mark A. Wurm

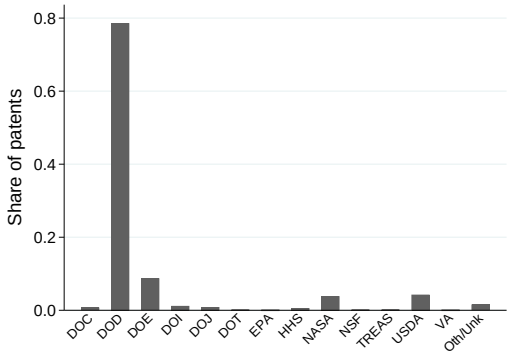
[57] ABSTRACT

A high performance, real-time control sequencer is disclosed which incorporates a unique state matrix logic. This real-

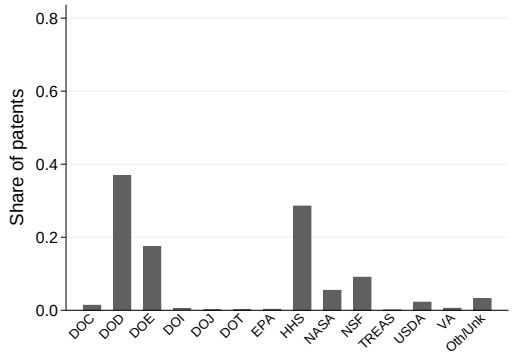
Notes: Figure shows the first and last index cards in the historical Government Patent Register (issued in 1890 and 1997, respectively), juxtaposed against the associated patent publications.

Figure 3: Agency shares of government-interest patents

Panel (A): Pre-1975
(source: historical Government Register)

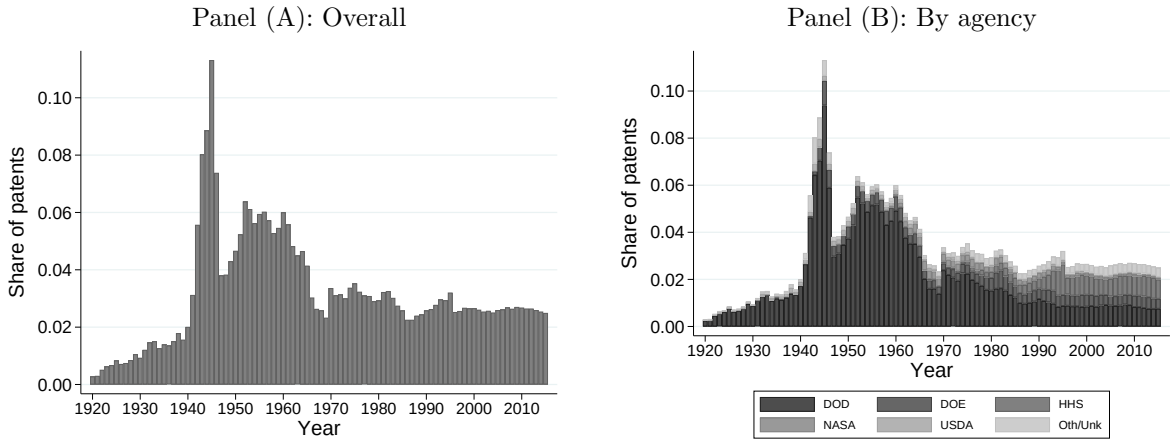


Panel (B): Post-1976
(source: same + modern data sources)



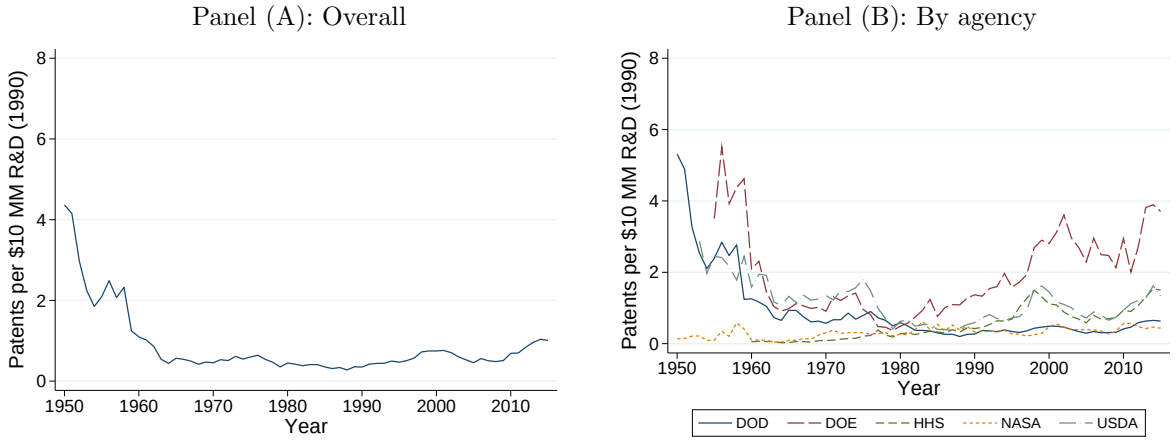
Notes: Table shows each of 13 federal agencies' share of government interest patents, pre-1975 (i.e., the era before modern electronic records, using the historical Government Register; Panel A) and post-1976 (using the historical Government Register and modern data sources such as the USPTO Patent Assignment Dataset and PatentsView; Panel B).

Figure 4: Share of U.S. patents produced with government funding, overall and by agency, shown by filing year (1920-2015)



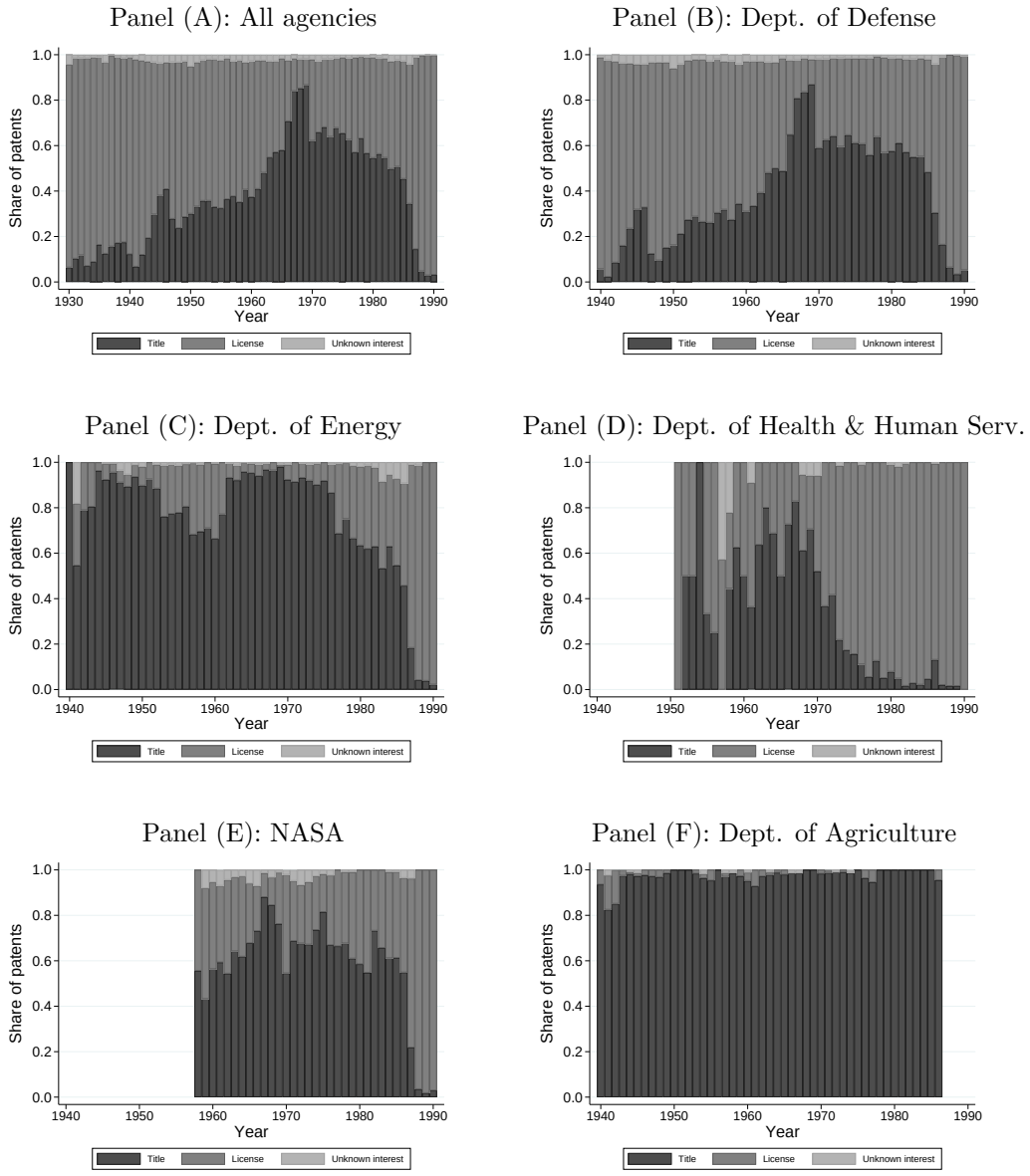
Notes: Figure shows the government-funded share of annual U.S. patenting (across filing years), overall (Panel A) and by funding agency for major R&D funders (Panel B).

Figure 5: Patents per R&D dollar, overall and by agency, shown by filing year (1920-2015)



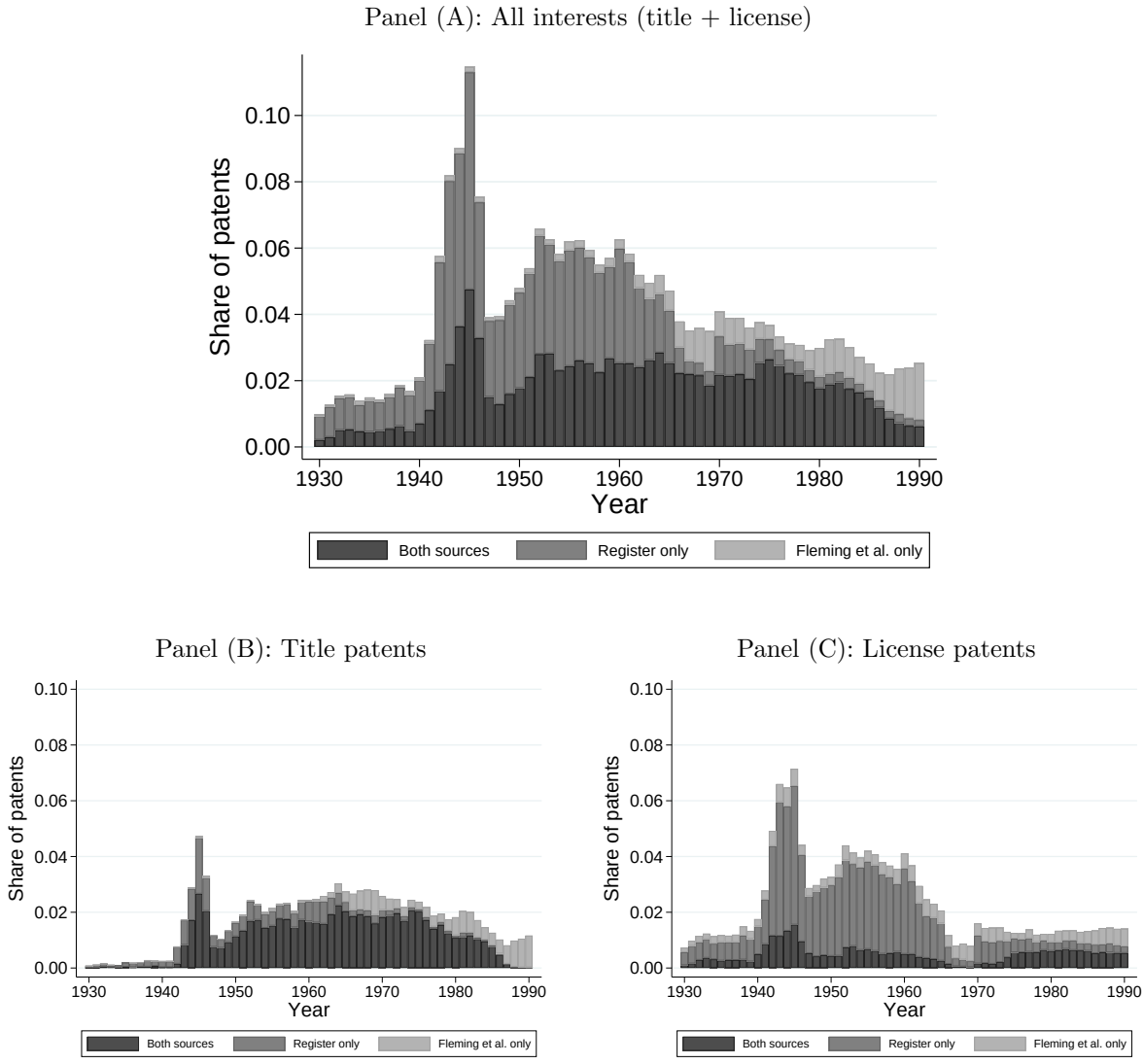
Notes: Figure plots annual patents per $10 million dollars of R&D expenditure (in 1990 dollars), overall (Panel A) and by funding agency for major R&D funders (Panel B).

Figure 6: Title vs. license shares in the historical Government Register, by filing year (1930-1990)



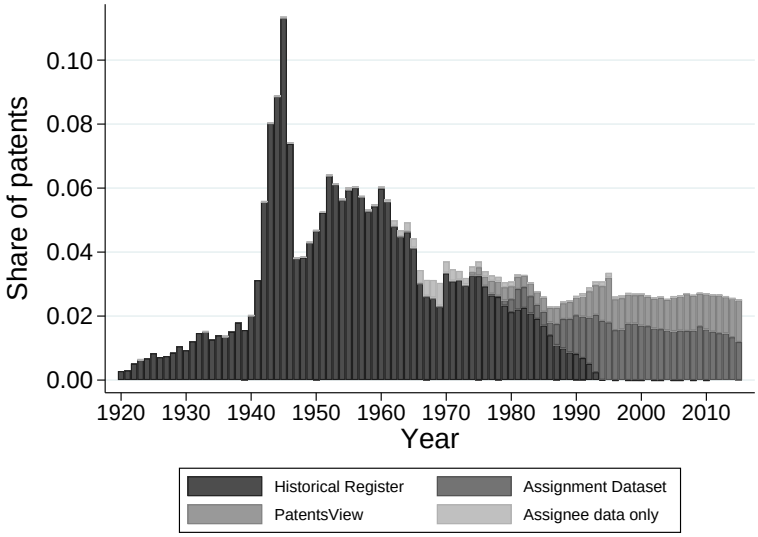
Notes: Figure shows the share of patents in the historical Government Register with title, license, and unknown government interest, overall (Panel A) and for select R&D funding agencies (Panels B to F). Sample is restricted to patents with either title or license indicated in the register (95% of all), 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 Government Register).

Figure 7: Comparison of historical Government Register and Fleming et al. data, by filing year (1930-1990)



Notes: Figure compares historical Government Register data to existing measures of government interest patents from Fleming et al. (2019), for all government interest patents (Panel A) and for title and license patents (Panels B and C, respectively).

Figure 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 Government Register, the modern USPTO Patent Assignment Dataset (UPAD), PatentsView (PV) government interest statement data, and assignee-based measures (based on Fleming et al. (2019) assignees pre-1976, and PatentsView assignees post-1976). The historical Government Register, UPAD, and PV are sufficient to capture nearly all government interest patents measured by Fleming et al. (2019), suggesting that these sources can be used to (more cheaply) extend this dataset in the future.

Table 1: Government interest patent counts through 2020, by data source

	Data source combinations				Associated patents					
	Historical	Assignment			All years		Pre-1976		Post-1976	
	Register	Database	PatentsView	Assignee data	Patents	Share	Patents	Share	Patents	Share
1.	Y	-	-	-	60,849	22.3%	57,057	65.2%	3,792	2.0%
2.	Y	-	-	Y	30,898	11.3%	27,608	31.6%	3,290	1.8%
3.	Y	-	Y	-	8,305	3.0%	0	0.0%	8,305	4.5%
4.	Y	-	Y	Y	3,795	1.4%	0	0.0%	3,795	2.0%
5.	Y	Y	-	-	361	0.1%	10	0.0%	351	0.2%
6.	Y	Y	-	Y	1,314	0.5%	0	0.0%	1,314	0.7%
7.	Y	Y	Y	-	901	0.3%	0	0.0%	901	0.5%
8.	Y	Y	Y	Y	3,406	1.2%	0	0.0%	3,406	1.8%
9.	-	-	-	Y	6,277	2.3%	2,774	3.2%	3,503	1.9%
10.	-	-	Y	-	63,977	23.4%	0	0.0%	63,977	34.4%
11.	-	-	Y	Y	3,731	1.4%	0	0.0%	3,731	2.0%
12.	-	Y	-	-	9,988	3.7%	34	0.0%	9,954	5.4%
13.	-	Y	-	Y	11,633	4.3%	0	0.0%	11,633	6.3%
14.	-	Y	Y	-	48,944	17.9%	0	0.0%	48,944	26.3%
15.	-	Y	Y	Y	18,927	6.9%	0	0.0%	18,927	10.2%
Total					*273,306*	*100.0%*	*87,483*	*100.0%*	*185,823*	*100.0%*

Notes: Table provides the number of government interest patents in our data identified by different combinations of data sources, illustrating where these data sources are overlapping versus additive. 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 government register (as seen in the first row of the table).

Web Appendix

A Data Appendix

Figure A.1: Box List for Historical Register records at the U.S. National Archives

TR-0241-2017-0042 - Patent Assignment - Government Licensees

ACCESSION #: TR-0241-2017-0042		
Patent Assignments Government Licensees (CY1890 thru CY1955) - 174 boxes		
Box #	**First Name or Serial Number**	**Last Name or Serial Number**
Box 1 of 174	A	AK
Box 2 of 174	AL	AN
Box 3 of 174	AP	AZ
Box 4 of 174	B	BEK
Box 5 of 174	BEL	BJ
Box 6 of 174	BL	BRI
Box 7 of 174	BRO	BZ
Box 8 of 174	C	CD
Box 9 of 174	CE	CL
Box 10 of 174	CM	CO
Box 11 of 174	CR	CZ
Box 12 of 174	D	DIE
Box 13 of 174	DIF	DZ
Box 14 of 174	E	EE
Box 15 of 174	EF	EZ
Box 16 of 174	F	FG
Box 17 of 174	FI	FZ
Box 18 of 174	G	GE
Box 19 of 174	GE	GH
Box 20 of 174	GI	GO
Box 21 of 174	GR	GZ
Box 22 of 174	H	HARD
Box 23 of 174	HARE	HE
Box 24 of 174	HI	HS
Box 25 of 174	HU	HZ
Box 26 of 174	I	I
Box 27 of 174	J	J
Box 28 of 174	K	KH
Box 29 of 174	KI	KZ
Box 30 of 174	L	LA
Box 31 of 174	LE	LL
Box 32 of 174	LO	LZ
Box 33 of 174	M	MAR
Box 34 of 174	MAS	McC
Box 35 of 174	McD	MG
Box 36 of 174	MI	ML
Box 37 of 174	MO	MZ
Box 38 of 174	N	NG
Box 39 of 174	NI	NZ
Box 40 of 174	O	O
Box 41 of 174	P	PH
Box 42 of 174	PI	PZ
Box 43 of 174	R	RAL

1

Figure A.1: Box List for Historical Register records at the U.S. National Archives

TR-0241-2017-0042 - Patent Assignment - Government Licensees

Box 44 of 174	RAM	REE
Box 45 of 174	REF	RI
Box 46 of 174	RO	RZ
Box 47 of 174	S	SC
Box 48 of 174	SE	SL
Box 49 of 174	SM	SP
Box 50 of 174	SQ	STO
Box 51 of 174	STR	SZ
Box 52 of 174	T	TH
Box 53 of 174	TI	UNITED - R
Box 54 of 174	UNITED - S	V
Box 55 of 174	W	WAL
Box 56 of 174	WAM	WER
Box 57 of 174	WES	WESTH
Box 58 of 174	WESTI	WEY
Box 59 of 174	WH	WZ
Box 60 of 174	X - Y - Z	Z
Box 61 of 174	DESIGN FOREIGN - REISSUE	PLANT & HOLDING
Box 62 of 174	ABANDONED	APPLICATIONS
Box 63 of 174	1948	SER
Box 64 of 174	ABANDONED. (CONT)	DEDICATED PLANT etc.
Box 65 of 174	ERDA - ASSIGNMENTS	2,000,000 - 2,999,999
Box 66 of 174	ERDA - ASSIGNMENTS	3,000,000 - 3,599,999
Box 67 of 174	ERDA - ASSIGNMENTS	3,600,000 - S.N. 802,400
Box 68 of 174	ERDA - LICENSES	1,900,000 - S.N. 293,862
Box 69 of 174	AGRICULTURE - USDA - ASSIGN	1,300,000 - S.N. 809,803
Box 70 of 174	AGRICULTURE - USDA - LICENSES	3,600,000 - APP. 419,326
Box 71 of 174	AIR FORCE - ASSIGNMENTS	2,400,000 - 3,699,999
Box 72 of 174	AIR FORCE - ASSIGNMENTS	3,700,000 - S.N. 7/669,256
Box 73 of 174	AIR FORCE - LICENSES	1,500,000 - 2,699,999
Box 74 of 174	AIR FORCE - LICENSES	2,700,000 - 2,899,999
Box 75 of 174	AIR FORCE - LICENSES	2,900,000 - 3,049,999
Box 76 of 174	AIR FORCE - LICENSES	3,050,000 - 3,299,999
Box 77 of 174	AIR FORCE - LICENSES	3,300,000 - 4,199,999
Box 78 of 174	AIR FORCE - LICENSES	4,200,000 - S.N. 08/072,605
Box 79 of 174	ARMY - ASSIGNMENTS	1,200,000 - 2,699,999
Box 80 of 174	ARMY - ASSIGNMENTS	2,700,000 - 3,167,999
Box 81 of 174	ARMY - ASSIGNMENTS	3,168,000 - 3,699,999
Box 82 of 174	ARMY - ASSIGNMENTS	3,700,000 - 4,299,999
Box 83 of 174	ARMY - ASSIGNMENTS	4,300,000 - S.N. 606,742
Box 84 of 174	ARMY - LICENSES	980,000 - 1,899,999
Box 85 of 174	ARMY - LICENSES	1,900,000 - 2,149,999
Box 86 of 174	ARMY - LICENSES	2,150,000 - 2,329,999
Box 87 of 174	ARMY - LICENSES	2,330,000 - 2,411,999
Box 88 of 174	ARMY - LICENSES	2,412,000 - 2,499,999
Box 89 of 174	ARMY - LICENSES	2,500,000 - 2,699,999
Box 90 of 174	ARMY - LICENSES	2,700,000 - 2,999,999

2

Figure A.1: Box List for Historical Register records at the U.S. National Archives

TR-0241-2017-0042 - Patent Assignment - Government Licensees

Box 91 of 174	ARMY - LICENSES	3,000,000 - 3,999,999
Box 92 of 174	ARMY - LICENSES	4,000,000 - S.N. 6/410,220
Box 93 of 174	NASA - ASSIGNMENTS	2,700,000 - S.N. 105,846
Box 94 of 174	NASA - LICENSES	2,400,000 - S.N. 08/095,930
Box 95 of 174	NAVY - ASSIGNMENTS	1,115,000 - 2,899,999
Box 96 of 174	NAVY - ASSIGNMENTS	2,900,000 - 3,149,999
Box 97 of 174	NAVY - ASSIGNMENTS	3,150,000 - 3,499,999
Box 98 of 174	NAVY - ASSIGNMENTS	3,500,000 - 3,899,999
Box 99 of 174	NAVY - ASSIGNMENTS	3,900,000 - 4,399,999
Box 100 of 174	NAVY - ASSIGNMENTS	4,400,000 - S.N. 7/501,996
Box 101 of 174	NAVY - LICENSES	432,000 - 2,409,999
Box 102 of 174	NAVY - LICENSES	2,410,000 - 2,549,999
Box 103 of 174	NAVY - LICENSES	2,550,000 - 2,749,999
Box 104 of 174	NAVY - LICENSES	2,750,000 - 2,959,999
Box 105 of 174	NAVY - LICENSES	2,960,000 - 3,149,999
Box 106 of 174	NAVY - LICENSES	3,150,000 - 3,999,999
Box 107 of 174	NAVY - LICENSES	4,000,000 - S.N. 08/128,124
Box 108 of 174	DOE - ENERGY - ASSIGNMENTS	3,000,000 - S.N. 567,679
Box 109 of 174	DOE - ENERGY - LICENSES	3,000,000 - S.N. 08/128,124
Box 110 of 174	HEW - HHS	DEPT OF HEALTH ...
Box 111 of 174	INTERIOR	DEPT OF INTERIOR
Box 112 of 174	VARIOUS GOV'T AGENCIES	A - K
Box 113 of 174	VARIOUS GOV'T AGENCIES	L - Z
Box 114 of 174	O.S.R.D ASSIGN - LICENSES	APP 527,364 - 639,139
Box 115 of 174	DESIGN	FOREIGN APPLICATIONS
Box 116 of 174	REVERSIONARY/UNID'D	NO INTEREST (NAVY)
Box 117 of 174	R.F.C & ARMY	UNIDENTIFIED / Designs
Box 118 of 174	ABANDONED (Ø)	APPLICATIONS (650,000)
Box 119 of 174	ABANDONED (650,000)	APPLICATIONS (800,000)
Box 120 of 174	ABANDONED 1948 & 1960	SERIES FORFEITED
Box 121 of 174	ABANDONED (567)	1948 SERIES (400,000)
Box 122 of 174	PATENTS	400,000 - 1,749,999
Box 123 of 174	PATENTS	1,750,000 - 1,949,999
Box 124 of 174	PATENTS	1,950,000 - 1,989,999
Box 125 of 174	PATENTS	1,990,000 - 2,259,999
Box 126 of 174	PATENTS	2,260,000 - 2,349,999
Box 127 of 174	PATENTS	2,350,000 - 2,399,999
Box 128 of 174	PATENTS	2,400,000 - 2,429,999
Box 129 of 174	PATENTS	2,430,000 - 2,457,999
Box 130 of 174	PATENTS	2,458,000 - 2,489,999
Box 131 of 174	PATENTS	2,490,000 - 2,529,999
Box 132 of 174	PATENTS	2,530,000 - 2,569,999
Box 133 of 174	PATENTS	2,570,000 - 2,614,999
Box 134 of 174	PATENTS	2,615,000 - 2,664,999
Box 135 of 174	PATENTS	2,665,000 - 2,709,999
Box 136 of 174	PATENTS	2,710,000 - 2,749,999
Box 137 of 174	PATENTS	2,750,000 - 2,789,999

3

Figure A.1: Box List for Historical Register records at the U.S. National Archives

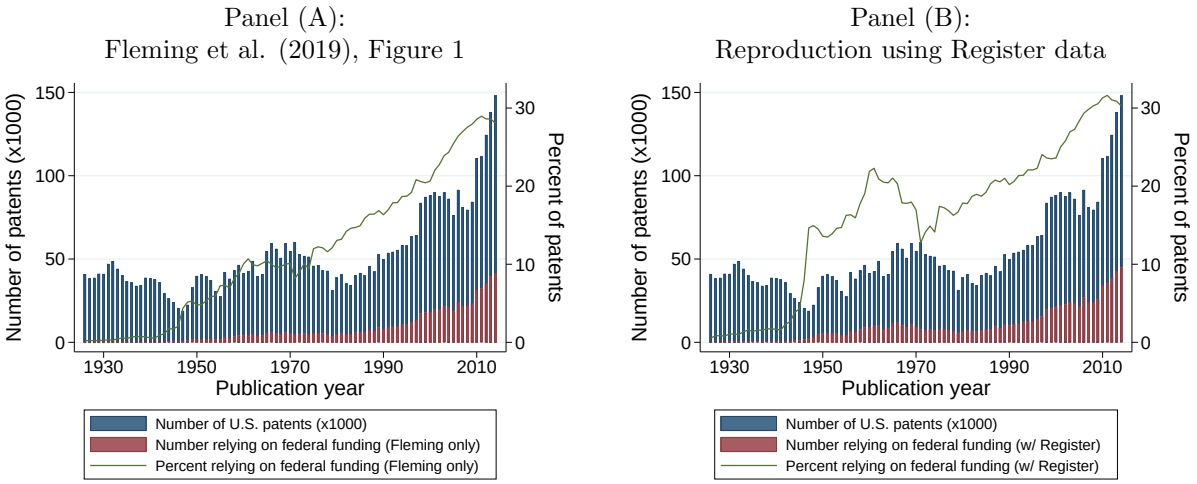
TR-0241-2017-0042 - Patent Assignment - Government Licensees

Box 138 of 174	PATENTS	2,790,000 - 2,826,999
Box 139 of 174	PATENTS	2,827,000 - 2,859,999
Box 140 of 174	PATENTS	2,860,000 - 2,896,999
Box 141 of 174	PATENTS	2,897,000 - 2,934,999
Box 142 of 174	PATENTS	2,935,000 - 2,969,999
Box 143 of 174	PATENTS	2,970,000 - 3,002,999
Box 144 of 174	PATENTS	3,003,000 - 3,039,999
Box 145 of 174	PATENTS	3,040,000 - 3,084,999
Box 146 of 174	PATENTS	3,085,000 - 3,133,999
Box 147 of 174	PATENTS	3,134,000 - 3,179,999
Box 148 of 174	PATENTS	3,180,000 - 3,229,999
Box 149 of 174	PATENTS	3,230,000 - 3,284,999
Box 150 of 174	PATENTS	3,285,000 - 3,332,999
Box 151 of 174	PATENTS	3,333,000 - 3,389,999
Box 152 of 174	PATENTS	3,390,000 - 3,449,999
Box 153 of 174	PATENTS	3,450,000 - 3,520,999
Box 154 of 174	PATENTS	3,521,000 - 3,619,999
Box 155 of 174	PATENTS	3,620,000 - 3,709,999
Box 156 of 174	PATENTS	3,710,000 - 3,779,999
Box 157 of 174	PATENTS	3,780,000 - 3,869,999
Box 158 of 174	PATENTS	3,870,000 - 3,939,999
Box 159 of 174	PATENTS	3,940,000 - 3,999,999
Box 160 of 174	PATENTS	4,000,000 - 4,069,999
Box 161 of 174	PATENTS	4,070,000 - 4,149,999
Box 162 of 174	PATENTS	4,150,000 - 4,219,999
Box 163 of 174	PATENTS	4,220,000 - 4,299,999
Box 164 of 174	PATENTS	4,300,000 - 4,399,999
Box 165 of 174	PATENTS	4,400,000 - 4,499,999
Box 166 of 174	PATENTS	4,500,000 - 4,599,999
Box 167 of 174	PATENTS	4,600,000 - 4,699,999
Box 168 of 174	PATENTS	4,700,000 - 4,799,999
Box 169 of 174	PATENTS	4,800,000 - 4,899,999
Box 170 of 174	PATENTS	4,900,000 - 5,099,999
Box 171 of 174	PATENTS	5,100,000 - 5,349,999
Box 172 of 174	PATENTS	5,350,000 - 5,573,999
Box 173 of 174	PATENTS	5,574,000 - 6,413,999
Box 174 of 174	DESIGN - H SERIES	T SERIES (Reissue)
Box ___ of ___		
Box ___ of ___		
Box ___ of ___		
Box ___ of ___		
Box ___ of ___		
Box ___ of ___		
Box ___ of ___		
Box ___ of ___		
Box ___ of ___		
Box ___ of ___		

4

B Supplementary Results

Figure B.1: Reproduction of Fleming et al. (2019), Figure 1 with Government Patent Register



Notes: Figure reproduces the main figure of Fleming et al. (2019) (Figure 1), using (i) the authors' original data (left panel), and (ii) the same data, but adding to it known government interest patents from the newly-compiled Government Patent Register which this paper introduces (right panel). We find that a substantially larger fraction of U.S. patents in the postwar era (1945-1970) were government-funded or cited government-funded invention. The driver of these differences is the data found in the Historical Register, which documents thousands more patents in the mid-century which were government-supported than can be measured from patent publications. Patterns in recent years are more consistent across data sources, likely reflecting that the stock of citable government-funded invention has grown, and patents tend to cite recent prior art.