NBER WORKING PAPER SERIES

AMERICA, JUMP-STARTED:
WORLD WAR II R&D AND THE TAKEOFF OF THE U.S. INNOVATION SYSTEM

Daniel P. Gross
Bhaven N. Sampat

Working Paper 27375
http://www.nber.org/papers/w27375

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2020, Revised September 2022

This paper is a revised version of a working paper previously circulated as “Inventing the Endless Frontier: The Long-Run Impacts of the World War II Research Effort on U.S. Innovation.” We thank Mike Andrews, Ashish Arora, Pierre Azoulay, Sharon Belenzon, Enrico Berkes, Wes Cohen, Claudia Goldin, Jon Gruber, Adam Jaffe, Simon Johnson, Shawn Kantor, David Mindell, Petra Moser, Tom Nicholas, Bitsy Perlman, Scott Stern, and Martin Watzinger for helpful comments. We also thank numerous conference and seminar audiences, and discussants Alex Whalley and Taylor Jaworski. We also thank Hayley Pallan, Greg Saldutte, and Innessa Colaiacovo for outstanding research assistance; Enrico Berkes and Gilles Duranton for sharing data; and the National Science Foundation, the Duke University Fuqua School of Business, Columbia University, Harvard Business School, and the NBER Innovation Policy grant for financial support. This material is based upon work supported by the National Science Foundation under Grant No. 1951470. 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.

America, Jump-started: World War II R&D and the Takeoff of the U.S. Innovation System
Daniel P. Gross and Bhaven N. Sampat
NBER Working Paper No. 27375
June 2020, Revised September 2022
JEL No. H56,N42,N72,O31,O32,O33,O38,R11

ABSTRACT

During World War II, the U.S. government's Office of Scientific Research and Development (OSRD) undertook one of the largest public investments in applied R&D in U.S. history, entering into thousands of contracts with firms and universities to perform research essential to the war effort. Using data on all OSRD-funded invention, we show that this shock had a formative impact on the U.S. innovation system, catalyzing technology clusters around the country with accompanying increases in high-tech entrepreneurship and employment. These effects continue growing to at least 1970 and appear to be attributable to agglomeration externalities, rather than sustained public R&D investment, which led to widening disparities in inventive output across the country. In the aggregate, wartime R&D permanently changed the trajectory of U.S. innovation in the direction of funded technologies, including electronics and communications.

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

Bhaven N. Sampat
Department of Health Policy and Management
Columbia University
722 W 168th Street, Room 486
New York, NY 10032
and NBER
bns3@columbia.edu

A large literature in economics has studied the determinants of innovation (Cohen 2010, Bryan and Williams 2021), including government funding (Bloom et al. 2019). The U.S. innovation system is especially rich in specialized, regional technology clusters which are thought to be important to overall technological progress (Chatterji et al. 2014, Carlino and Kerr 2015), while also contributing to growing gaps in regional economic performance (Gruber and Johnson 2019). Yet this literature has few examples of systemic R&D shocks, and underexplores issues such as (i) the long-run effects of public R&D investments, (ii) the impacts of large, actively-managed applied research programs (Azoulay et al. 2019a, Gross et al. 2022), and (iii) to what degree, and how, these investments affect regional economic development—issues which are central to policy today.[1]

In this paper, we study the long-run effects of the largest R&D shock in U.S. history. In World War II, the newly-created Office of Scientific Research and Development (OSRD) led an expansive effort to develop technologies and medical treatments for the Allied war effort (Gross and Sampat 2022d). From 1940 to 1945, OSRD engaged industrial and academic contractors in more than 2,200 R&D contracts worth over $9 billion (current dollars), despite no pre-war tradition of funding extramural R&D. At the height of the war, the U.S. government was funding the research behind nearly 1 of every 8 U.S. patents—more than five times pre-war and modern levels, and nearly twice the level at the peak of the Cold War in the 1950s and 1960s (Figure 1).

[Figure 1 about here]

The immediate effect of these investments was a range of technological advances which were not only instrumental to the success of the Allied campaign, but also of wide civilian value after the war ended. Its longer-run impact was to reshape the U.S. innovation system. We document four main findings. First, World War II R&D kicked off the postwar growth of technology clusters (counties x technologies) around the country: despite parallel trends prior to the war, the most heavily-treated clusters were by 1970 producing another 40% to 50% more patents per year than untreated clusters. Second, this sustained growth benefited from, but did not depend on, postwar federal R&D investment. Instead, our evidence suggests OSRD catalyzed Marshallian agglomeration, including through firm in-migration, entry, and growing spillovers across inventors and technologies. Third, we find evidence that these changes were accompanied by growth in local industrial employment and firm creation in related high-tech industries. Finally, we show that wartime R&D had permanent

---

[1] These gaps have recently become relevant: in August 2022, the U.S. initiated its largest public investment in applied, use-oriented R&D since the Cold War (via the CHIPS and Science Act). Among its provisions, it adds a $20 billion technology directorate to the National Science Foundation (NSF) and a $10 billion investment in regional technology hubs, aiming to develop new domestic capabilities in frontier technologies and to create new capacity in regions which have not previously been major R&D centers (Gruber and Johnson 2019).

effects on the direction of U.S. innovation, which pivoted towards electronics and communications. By (rapidly) extending the frontier of key emerging technologies while stimulating agglomeration, the impacts of this shock were thus to enhance competitiveness but widen differences in inventive productivity, and in turn economic performance, across the country.

The OSRD project was radical for its time, representing the first serious government funding of extramural research in the U.S. and marking a major turning point in research policy. OSRD itself existed only for the duration of the war, but in that span it was responsible for many important technological developments, including radar, electronic communication and early computing, underwater detection (sonar), rockets and jet propulsion, and atomic fission, as well as advances in the medical and pharmaceutical fields such as mass-produced penicillin, influenza and other vaccines, new malaria treatments, new approaches to managing myriad human hardships from sleep and oxygen deprivation to nutrient deficiencies, and dozens more.

To study the effects of World War II R&D, we have collected the universe of OSRD contracts from archival records, including detailed information on the contractors, contracts, and all inventions, patents, and scientific publications they produced. We merge these records with the complete U.S. patent record, new administrative data on postwar government-funded patenting and research institutions, and SIC-level measures of local firm creation and industrial employment. We use these to study the effects of the OSRD shock on invention, the growth of local research institutions, and downstream economic outcomes from 1930 to 1970. Our empirical design will compare pre- and post-war patenting in clusters shocked by the war effort, which we measure as the OSRD share of cluster patents in the 1940s, and we take a similar logic in evaluating other outcomes at the cluster, regional, and national levels—where, for example, we compare U.S. to foreign patenting over time in more- and less-intensively treated technology areas.

Across this paper, a single pattern consistently presents itself: parallel pre-war trends, a wartime spike in invention in OSRD-funded technologies, and a postwar take-off that continues through the end of our analysis window. This is what we observe, for example, in comparing invention over time in more- and less-intensively shocked clusters. The magnitudes are large: a doubling of the OSRD share of 1940s patents in a given cluster is associated with 20% higher patenting by 1960 and 30% by 1970, relative to pre-war levels. In a subset of clusters, these magnitudes were off the charts: Middlesex, MA (the locus of World War II radar R&D; see Gross et al. 2022) experienced a nearly *30-fold* increase in electronics invention during the war, a postwar return to nearly pre-war levels, and then a sustained take-off, with patenting in 1960 already ten times pre-war levels. We also find evidence of localized spillovers: OSRD investment in proximate technologies in a given county had positive, albeit attenuated, effects on cluster growth.

The puzzle of this paper, however, is not only what effect the OSRD R&D shock had, but why its impacts were so persistent. We first establish that the post-war take-off in patenting is not driven by direct follow-on to OSRD invention, nor by patents of firms and inventors involved in the war effort. Having ruled out these explanations, we consider two other possibilities: (i) continued government R&D investment in the same locations and technologies, building off the foundation OSRD created, or (ii) Marshallian agglomeration dynamics. Our evidence is consistent with the latter: it appears entire local research ecosystems sprung up in many locations and technology areas where OSRD activity was concentrated. In more heavily-shocked clusters, we see increases in both public and private patenting, and increases from a wide variety of entities, including in-migrating firms and de novo entry. We show that concentration of patenting accordingly declines, as these clusters develop and inventive activity grows more diffuse across the local economy. Beyond patents, we provide evidence that downstream industrial activity follows: using historical Dun & Bradstreet (D&B) and County Business Patterns (CBP) data, and linking patent classes and industries via a USPTO concordance, we show that postwar firm creation and employment were higher in counties and industries which were the targets of OSRD-funded research.

Finally, we show that the collective impact of the OSRD shock was permanent change in the direction of aggregate U.S. innovation. In a triple-differences design comparing U.S. and foreign patenting over time in more- versus less-intensively treated technology classes, we see a clear U.S. divergence in the direction of OSRD-funded technologies after the war, without pre-trends. In technologies areas in the top quartile of treatment intensity, U.S. patenting had grown >50% more than in foreign countries by 1970. These differences are present whether comparing patents at USPTO to foreign patent authorities, or among domestic and foreign inventors on U.S. patents. Taken together, the results of this paper demonstrate that World War II laid the groundwork for the growth of postwar technology hubs and high-tech industries.

There is widespread recognition that World War II was a sea-change event in government-science relations and in science and technology policy.[2] Although policymakers and scholars from Bush (1945) to Gruber and Johnson (2019) have appealed to the wartime effort as a paradigmatic example of the benefits of federal research funding, there has been limited empirical assessment of OSRD itself. A sizable literature has studied the impacts of other public R&D investments on innovation

---

[2] Impressed by its immediate results, President Roosevelt asked Bush to draw lessons from this "unique experiment" to harness science in peacetime to increase economic growth, improve national security, and develop new medical treatments (Roosevelt 1944). Bush's response, *Science: The Endless Frontier* (Bush 1945) is often considered the ideological blueprint for postwar science policy (Mowery 1997, Nelson 1997). While many of the specific institutional recommendations in the Bush Report were not adopted, the enthusiasm it generated and the public perception that science won the war helped launch a massive postwar expansion of U.S. federal funding for research.

(Bloom et al. 2013, Azoulay et al. 2019b, Myers 2020, Myers and Lanahan 2022) and other outcomes (Howell 2017).[3] Most existing evidence, however, is drawn from studies of marginal changes in funding, and often for basic science. As a result, there is, to-date, limited evidence as to what effects a systemic shock to R&D funding of this scope and scale may have, where, over what horizons, and whether these investments can trigger virtuous cycles of endogenous growth or require continued investment to have sustained effects—the main questions we tackle in this paper. Crucially, the passage of time allows us to evaluate these long-run effects. With the postwar era often viewed as a golden age of economic growth (Marglin and Schor 1991), the impacts of the war on innovation seem central to understanding how this growth was achieved, especially given the evidence in this paper of a postwar "innovation takeoff" (Appendix Figure C.1).

Our results contribute to research in the geography of innovation (Feldman 1994, Audretsch and Feldman 2004), especially agglomeration and innovation (Carlino and Kerr 2015, Kerr and Robert-Nicoud 2020). This literature frequently documents the localization of inventive activity and R&D spillovers (Jaffe et al. 1993, Audretsch and Feldman 1996), explores the features of specific technology clusters (e.g., Saxenian 1996), and relates agglomeration to R&D productivity (e.g., Buzard et al. 2017, Berkes and Gaetani 2021, Moretti 2021, Gruber et al. 2022), identifying numerous reasons why innovation experiences locally increasing returns to scale.[4] The literature has made less progress on the inverse question—whether discrete R&D shocks trigger agglomeration (Duranton 2007, Kerr 2010), and more generally what catalyzes change (Chattergoon and Kerr 2022).[5] Because innovation is frequently tied to population and industrial activity, our results also have implications for the broader literature on industrial agglomeration (Marshall (1890), Krugman (1991); see Duranton and Puga (2004) for a review), including place-based policy (e.g., Chatterji et al. 2014, Kline and Moretti 2014a,b). We use this literature to help frame our analysis as we explore why our transitory shock had such long-lived effects.[6]

One implication of the clustering of innovation which our paper documents is growing regional differences in inventive output, which may contribute to growing polarization in economic outcomes across the U.S. These regional disparities are a key question in policy debates today that was

---

[3] Also see Santoleri et al. (2022) and Bergeaud et al. (2022), among others.

[4] Also see Almeida and Kogut (1999), Thompson and Fox-Kean (2005), Thompson (2006), Arora et al. (2018), Kantor and Whalley (2014, 2019), and Andrews (2021), among numerous others.

[5] On this theme, Schweiger et al. (2022) study the long-run impacts of Soviet investments in "Science Cities" before and during the Cold War on local economic development, finding that the workforce in these cities is now higher-income, more educated, and produce more patents than matched controls.

[6] In some ways, the most closely related research to this part of our paper may be Buenstorf and Klepper (2009) and Klepper (2010), who also study the emergence of high-tech U.S. clusters, attributing their growth to organizational reproduction (e.g., spinoffs) from industry pioneers, as well as Arthur (1990), who explores the effects of historical accidents on clustering via path dependence in an evolutionary framework.

anticipated in the debates near the end of the war (the so-called "Bush-Kilgore debates") about the best way to allocate postwar research funding—which included controversy over whether to use science funding to promote the best science (which may be localized), or to prioritize a broad geographic and institutional distribution and spread the wealth (Kevles 1977a). The evidence in this paper suggests such controversy may be an inescapable fallout of concentrated public R&D investment if research activity increasingly agglomerates in funded clusters.

Beyond these themes, this paper relates to research on the social returns to R&D (Jones and Williams 1998, Jones and Summers 2021, Myers and Lanahan 2022), and in turn the wider literature on endogenous growth (Romer 1986, 1990), where innovation features increasing returns to scale but which has few examples of discrete shocks and takeoffs. Most recently, Kantor and Whalley (2022) have examined the impacts of the Cold War expansion of aerospace R&D on local manufacturing and used this context to estimate a large fiscal multiplier on public R&D. Our paper complements this literature, highlighting the growth initiated by the World War II shock, and the long-lasting changes to the U.S. innovation system this event brought about.

As such, we bring a renewed perspective to the origins of the modern U.S. innovation system (e.g., Kevles (1977b), Geiger (1993); also see Arora et al. (2021)), while building on research that studies the role of defense R&D (Mowery 2010, Howell et al. 2021, Moretti et al. 2022, Belenzon and Cioaca 2021) and the impacts of war (e.g., Ruttan 2006), which has been a laboratory for several questions in the economics of innovation.[7] Beyond wars, the COVID-19 pandemic has raised new questions around crisis R&D policy and its long-term effects, where a closer examination of World War II and OSRD may be instructive (Gross and Sampat 2021, 2022b,c,d).

We proceed as follows. In Section 1 we review the World War II research effort, describing OSRD's work and legacy. Section 2 introduces our data and characterizes the World War II shock. Section 3 documents the effects of World War II R&D on local, postwar invention, and Section 4 explores the myriad changes we see taking places in treated clusters. In Section 5, we examine downstream impacts on employment and firm creation. Section 6 then evaluates impacts of World War II R&D on the direction of innovation at the national level. Section 7 offers concluding remarks, including insights for open and long-running policy debates today.

---

[7] For example, Moser and Voena (2012), Iaria et al. (2018), Biasi and Moser (2021), and Gross (2022).

1 Historical and Policy Background

1.1 Historical Background

World War II had myriad impacts on the U.S. economy (Fishback and Cullen 2013, Jaworski 2017, Field 2022), but among them was a shock to innovation and innovation policy. Before World War II, there was very little federal funding of research outside of agriculture. Most academic research was funded by philanthropic foundations (Rockefeller and Carnegie, in particular) and industry. There was, if anything, an aversion among academics to public funding, reflecting concerns that it may distort the direction of research and restrict scientific freedom.

World War II changed this. Even before the attack on Pearl Harbor and the United States' official entry into the conflict, scientists, the military, and politicians anticipated that the development and application of technology would be critical for an Allied victory, that existing U.S. military R&D was inadequate, and that coordination would be required to mobilize the scientific and technological capabilities that had developed during the interwar era.[8]

In June 1940, Vannevar Bush (the former vice president and dean of engineering at MIT, president of the Carnegie Institution of Washington, and chairman of the National Advisory Committee for Aeronautics) together with other members of the U.S. scientific and technological establishment[9] convinced President Roosevelt to establish the National Research Defense Committee (NRDC) to "correlate and support scientific research on the mechanisms and devices of warfare" (Pursell 1979). The NRDC was to supplement existing military research (by the Departments of War and Navy) related to warfare "by extending the research base and enlisting the co-operation of institutions and scientists" (James Conant, quoted in Stewart 1948).

Perhaps as important as any of the technologies it helped to develop, the wartime research effort was a major innovation in the way science was supported and conducted.[10] While the first World War disrupted universities and firms by drawing scientists out of laboratories, and previously U.S. government agencies themselves had some done research internally, the NDRC effort primarily funded research "extramurally" through contracts, engaging both firms and universities, from individual investigators to larger laboratories. Impressed by its early successes, NDRC was expanded by a 1941 Executive Order to emphasize more development work (beyond just research), to solidify

---

[8] Most famously, in 1939 Albert Einstein warned President Roosevelt about Nazi advances in uranium research, and urged a U.S. atomic R&D program. Development of radar technology was another top priority.

[9] Bush was supported by Karl Compton, president of MIT; James Conant, president of Harvard; and Frank Jewett, president of the National Academies, vice president of AT&T, and director of Bell Labs.

[10] Scholars have since described the wartime arrangement as having "portended the beginning of a new relationship between the federal government and the nation's universities" (Geiger 1993).

links with military agencies conducting research, and to take over wartime medical research and development. The new organization, the Office of Scientific Research and Development (OSRD), was also eligible for regular Congressional budget appropriations. As *The New York Times* wrote, this effectively made Vannevar Bush "the czar of research."[11]

OSRD itself comprised three constituent bodies: NDRC, which continued to contract for R&D on instruments of war; the newly-added Committee on Medical Research (CMR), which supported research in military medicine; and an Advisory Council, which coordinated research activities across OSRD, the military, and other civilian agencies. OSRD also provided a range other administrative functions to support an end-to-end R&D effort, including through an Engineering and Transition Office (which advised on the production and use of devices developed under OSRD research) and an Office of Field Service (which assisted in getting OSRD-funded technology into the battlefield). NDRC remained the chief organ of OSRD, with 19 major divisions (e.g., Division 1, "Ballistics Research"; Division 11, "Chemical Engineering", Division 14, "Radar") and several special sections, including the top secret S-1 section, which organized early research on controlled nuclear reactions and was subsequently spun out into the Manhattan Project atomic weapon development program. OSRD entered into >2,200 R&D contracts with industrial and academic contractors over the course of the war with a cumulative value of $9 billion current dollars, which, as indicated, was one or two orders of magnitude larger than pre-war government R&D spending.

This effort helped develop a range of technologies that were crucial to the Allied victory. Radar, mass-produced penicillin, and the atomic bomb are among its most memorable achievements, but OSRD also produced significant advances in rocketry, jet propulsion, radio communications, and electronic computing, plus treatments for malaria, pesticides like DDT, and more—all of which found commercial applications after the war. Much of this research was concentrated in a network of research labs hastily organized at the start of the war, in part through Vannevar Bush's personal network and that of his NDRC deputies. Each had a distinct technological focus and worked in tandem with firms and corporate R&D labs which, provided inputs and helped develop prototypes, whereas basic research was typically contracted to academic researchers around the country.[12] These early "national labs" attracted scientists and engineers from around the country, many of

---

[11] In addition to Bush, Compton, Conant, and Jewett, OSRD's civilian leaders included Richard Tolman (President of the California Institute of Technology) and Conway Coe (U.S. Patent Commissioner).

[12] For example, radar development was centered at the MIT Radiation Laboratory (the "Rad Lab"), and radar countermeasures at the nearby Harvard Radio Research Lab (RRL). Rocket and jet propulsion research was based at the CalTech Jet Propulsion Lab (JPL), and proximity fuze development at the Johns Hopkins Applied Research Lab (APL). Early, NDRC-supported research on uranium fission took place at academic labs at the University of Chicago, UC Berkeley, and Columbia University before spinning out into the Manhattan Project, which was based in Los Alamos, New Mexico, supported by project sites around the country. These labs were the predecessors of postwar national labs in these locations, most of which are still operating today.

whom dispersed at the end of the war—though some also stayed.[13] In parallel research, we and others have documented how these coordinated R&D programs laid foundations of new industries which emerged after the war (e.g., Klepper 2016, Agarwal et al. 2021, Gross et al. 2022), potentially to the benefit of regions where these industries were based.

Even before the war was over, there was broad agreement that the government should be involved in funding research at universities after the war. Perhaps ironically, the initial attempts to create a structure for postwar funding came from a critic of OSRD, Senator Harley Kilgore (D-W.Va.). Kilgore, a New Deal Democrat, was concerned about the concentration of OSRD funding in big business and a handful of universities (Kevles 1977a).[14,15] Kilgore had other concerns about OSRD model, including that many of the contracts allowed the recipients to retain patent rights—making the intellectual output of government-supported research private property—and that there was a lack of representation from small business, independent inventors, and non-elite universities in the wartime effort. He believed each of these feautures of OSRD hurt the rate of technological development during the war and also led to concentration of the benefits of federal funding in a few research fields, institutions and regions (Kevles 1977a, Kleinman 1995). In a series of bills introduced during the war, culminating in a 1944 proposal of a new "National Science Foundation", Kilgore attempted to forge a peacetime research policy that would fund basic and applied research in response to specific socio-economic problems, with a mandate for broad geographical and institutional distribution of funds, wide dissemination of research results (including public ownership of resulting patents), and political accountability of researchers.

Vannevar Bush's seminal report *Science, The Endless Frontier* (Bush 1945), written at the request of President Roosevelt and published near the end of the war, was, in many ways, a rejoinder to Kilgore's arguments and proposal.[16] Like Kilgore, Bush recommended a single agency (a "National Research Foundation"), but with a focus on basic research, run by scientists, with broad scientific autonomy, and aimed at stimulating high-quality research by the best institutions and scientists. In making the case for federal funding of fundamental research at universities, the Bush Report

---

[13] By the fall of 1941, OSRD research had already involved 78% of top American physicists and 52% of top chemists, as measured by the publication *American Men of Science* (Stewart 1948). That many of these researchers migrated and stayed is apparent from, for example, the records of the Rad Lab (Gross et al. 2022).

[14] Kilgore was more generally concerned about cartels and concentration of economic power in the U.S., and became interested in the science policy issues as a member of the Senate Military Affairs Committee. Various witnesses testified before the Committee that the OSRD-led research effort was not taking full advantage of the nation's research capabilities beyond elite institutions and large firms (Maddox 1981).

[15] As Appendix Table A.3 shows, 60% of OSRD funding went to just 10 institutions, and nearly 50% to MIT, Harvard, and CalTech alone—all universities represented by OSRD leadership.

[16] Kevles (1977a) documents that although the Report was framed as a response to President Roosevelt's November 1944 letter asking Bush to draw lessons from OSRD experiment for peacetime science policy, Bush himself was instrumental in drafting the letter and coordinating the Presidential request.

also anticipated the "market failure" theory of federal funding[17] (e.g., Arrow 1962) and the "linear" model of science and innovation[18] (Mowery 1997, Nelson 1997).

Though the Bush Report had a strong ideological impact on U.S. policy, many of its specific proposals met a cool reception, including from Kilgore and other liberals, who preferred a more egalitarian peacetime approach, and from President Truman, who insisted on a politically-appointed director. In the five years after the war, NSF legislation reflecting the Bush, Kilgore, and compromise visions was introduced and debated. By the time the National Science Foundation Act was passed in 1950, many of OSRD's remaining research contracts had been transferred to mission agencies (e.g., the Office of Naval Research, the Atomic Energy Commission, and the National Institutes of Health), precluding the single-agency approach Bush (and Kilgore) had envisioned. Though the NSF was in large part "a triumph for Bush" (Kevles 1977a)—primarily focused on basic research, administered by scientists—its budget was small, and it was a "puny partner" in the overall enterprise (a position which, under recent legislation, may be changing, as NSF's scope expands into funding downstream research and applied, use-oriented technology development).

While each of the other major postwar R&D funding agencies had their own rules and procedures, a striking feature of federal research funding in the decades that followed was its continued geographic and institutional concentration. Though a variety of legislative initiatives and programs, historical and recent, have attempt to widen the distribution of funding—channeling Kilgore's criticism of OSRD and concerns about extending the OSRD model in peacetime—opponents of these programs typically argue that funding should be directed to the best researchers and institutions, as determined by the scientific community, echoing Bush. One reason for this tension is disagreement over what the goals of R&D policy are, or should be. But a key and complementary gap in this debate is evidence on the impacts of these choices: whether the geographic distribution of R&D funding matters for local economic development, and the degree to which returns accrue locally versus more broadly. One goal of this paper is to speak to these questions.

2 Data

To assess the effects of the World War II shock, we have collected, transcribed, and harmonized a complete record of all 2254 OSRD contracts (to 461 distinct contractors), and all 7910 inventions,

---

[17] "[W]e cannot expect industry adequately to fill the gap ... basic research is essentially noncommercial in nature ... it will not receive the attention it requires if left to industry."

[18] "[B]asic research is the pacemaker of technological progress."

2637 patents, and 2463 scientific publications reported under them.[19] Through additional sources not included in OSRD's public records (e.g., the Manhattan Project), we are able to identify a total of 3,137 OSRD-funded, patented inventions, which we use to measure the OSRD shock, preferring these to OSRD's prime contract spending because they represent outputs, merge to other sources, and bring us closest to the level where the work was performed.[20]

We link these data to the U.S. patent record. To do so, we have collected data on all U.S. patents granted between 1920 and 1979, merging a USPTO master file of patents with patent number, patent class (USPC), and issue date (Marco et al. 2015) with data on (i) serial numbers and filing dates; (ii) the full network of front-page (historically, final-page) citations; (iii) harmonized assignee names and types; and (iv) inventor locations, which we measure using data from Petralia et al. (2016), Berkes (2018), and Bergeaud and Cyril (2022) (see Appendix B.1).[21] We supplement these data with newly-collected administrative, archival data on government-funded patents since the early 1900s, which we introduce in Appendix B.2 and discuss in depth in concurrent writing (Gross and Sampat 2022a), and which comprises a significantly larger set than can be measured from published patents alone (Fleming et al. 2019). For our cross-country comparisons, we add data from the European Patent Office (EPO) PATSTAT database on granted patents in the U.S., Great Britain, and France over the same period, which includes similar information to that of our USPTO base layer: patent number, patent class (IPC), and grant date.

In Section 5, we measure county-level employment and firm creation by industry using the U.S. Census Bureau's County Business Patterns (CBP) and Dun & Bradstreet's (D&B) historical data files. We use CBP and D&B data from 1980 (the latter lists over 4.5 million U.S. establishments, including their 4-digit SIC and founding year) to study long-run outcomes, and we apply a USPTO crosswalk to map SIC codes and patent subclasses to a common, SIC-derived (but USPTO-generated) clas-

---

[19] We observe detailed data on each contract, including the contractor, subject matter (OSRD division which wrote the contract), total value, security classification, patent policy, and termination date.

[20] Beyond the 3382 patent applications (2637 issued patents) identified in OSRD records, we measure an additional 461 OSRD-funded serials (388 patents) associated with the Manhattan Project through a public records request (Streifer 2017) and 36 serials (8 patents) from records of the Army's Judge Advocate General's office (see Gross 2022). We also supplement these records with an automated, text-based search for continuations and divisions of these patent applications, which identifies another 104 OSRD-supported patents.

[21] We are grateful to all three sets of authors for sharing the data.

sification, enabling us to perform analysis that links our treatment to industry outcomes.[22,23,24] We restrict the D&B sample to single-location firms and headquarters establishments, and to firms which we can accurately geocode using address information (89% of the sample). We then aggregate up firm counts to the county x industry x founding decade level, to smooth over bunched rounding in founding years. From the CBP, we thus obtain a 1980 cross section of county-industries, and from D&B, we build a 1920-1980 panel of county-industries.

Distribution of OSRD activity across space and subject matter

OSRD contracted for research in a wide range of subject areas, and with an array of contractors Table 1 lists the top 10 OSRD patent classes and their share of OSRD patents, contrasting these frequencies to the share of patents these classes comprised in the recent pre-war era. Together with Figure 1, the table brings into relief how large a shock World War II was to the U.S. innovation system, both in scale and in the technologies OSRD was pushing.

[Table 1 about here]

In the appendix we provide additional context. Of particular note are Appendix Tables C.1 and C.2, which report the top (i) technology areas and (ii) and patent classes with OSRD patents in the 1940s, ranked by share of 1940s patents that were OSRD funded—measuring the size of the shock. Atop Table C.1 is nuclear energy, but most other high ranking subjects are in the domain of electronics and communications, including radar and microwave engineering, semiconductors, electrical computing, and cryptography, highlighting the role that World War II research made in advancing these fields, with potential applications beyond warfighting.

Figure 2 maps locations in the continental U.S. with OSRD-funded patents, demonstrating OSRD's geographic scope: although a handful of states received a large majority of its funding (Appendix Table A.2), and particular programs were concentrated in specific locations, OSRD-funded R&D spanned the country. Table 2 weaves these threads together, listing the top five counties with the

---

[22] Our choice to use 1980 data files has several motivations. Earlier CBP editions which we have experimented with (e.g., 1956, 1959, 1970) report 3- and 4-digit SICs with much lower frequency, undermining the patent-industry crosswalk and limiting power, whereas the CBP from the late 1970s onwards provides finer disaggregation. Earlier D&B files are significantly smaller, and we believe only a partial accounting. Additionally, we prefer data produced under the same SIC edition as the USPTO crosswalk (1972). We lose relatively little by limiting the CBP-based analysis to 1980, as the CBP only exists post-1947, precluding pre-/postwar comparisons.

[23] The D&B data cover a large sample of U.S. establishments, approximating the universe (4.531 million establishments in 1980, versus 4.543 million in the CBP and 4.533 million in the Census Bureau's Longitudinal Business Database). Note that the D&B firm counts are by construction conditioned on survival to 1980. We will use industry x founding year fixed effects to account for differential survival rates across firm birth cohorts.

[24] Data available at https://www.uspto.gov/web/offices/ac/ido/oeip/taf/data/sic_conc/.

most OSRD patents in select technology areas, and the OSRD-funded share of local patenting in the 1940s—i.e., the shock whose effects we will examine next.

[Figure 2 and Table 2 about here]

3 Postwar Take-off of World War II Technology Clusters

To understand the impacts of World War II on the U.S. innovation system, our starting point is to examine its constituent parts: the growth of regional innovation hubs. Descriptive evidence hints at the effect that OSRD-supported R&D may have had on local innovation, foreshadowing the results we systematically find. Figure 3, Panel (A) shows the time series of filed patents in the 12 largest Massachusetts counties from 1935 to 1965. Prior to the war, Middlesex County—home to Harvard and MIT, which hosted large, central laboratories for wartime radar research—produced annually more patents in levels but was not on a noticeably different time trend than other counties. During the war, invention spiked, driven by OSRD-funded R&D (see Table 2), and after the war returned to pre-war levels, before taking off in the 1950s. By the mid-1960s, Middlesex was producing twice its number of pre-war patents, as this modern cluster was taking shape.

[Figure 3 about here]

In Panel (B) we look within-county, comparing the trajectories of high-level technology areas (1-digit NBER categories) around the war. We plot indexed time series for six technology categories (chemical, communications, pharmaceutical, electrical and electronic, mechanical, other), indexing to 1935 levels, and find similar patterns of even larger magnitude. Communications patenting was shocked nearly 30-fold in the war, returned to pre-war levels, and was by 1965 over 10 times pre-war levels. Electronics patenting followed a similar, if attenuated, pattern.

We set out to perform these comparisons in a more systematic way across the country. To do so, we compare patenting over time in clusters (counties x technology areas) with higher versus lower levels of OSRD investment. We measure technologies at the level of 2-digit NBER patent categories Hall et al. (2001), which aggregate patent classes (our emphasis on the historical period makes the U.S. patent classification, and others derived from it, preferable to more modern classification systems). Our baseline estimating equation throughout this section is as follows:

$$Ln(Patents)_{ict} = \sum_{t=1931}^{1970} \beta_t \cdot Ln(OSRD\ Rate)_{ic} \cdot Year_t + \alpha_{ic} + \delta_t + \varepsilon_{ict} \tag{1}$$

where i indexes counties, c indexes patent categories, and t indexes years, and the sample runs from 1930 to 1970, with standard errors clustered at the county level. Our preferred treatment measure is what we henceforth call the "OSRD rate": the fraction of patents filed in a given cluster between 1941 and 1948 which were OSRD-funded. Our primary specification uses a continuous measure of the logged OSRD rate, which mechanically restricts the sample to clusters with at least one OSRD patent. We at times present results from specifications with treatment quartiles, which allows us to compare segments of the treatment distribution in a more flexible way, against each both other and clusters with no OSRD patents (the reference group):

$$Ln(Patents)_{ict} = \sum_{q=1}^{4} \sum_{t=1931}^{1970} \beta_{qt} \cdot \mathbb{1}(\text{Treatment quartile } q)_{ict} \cdot Year_t + \alpha_{ic} + \delta_t + \varepsilon_{ict} \tag{2}$$

We note that these specifications will not necessarily estimate the effects of the OSRD shock on local invention in isolation, because in equilibrium our units may be interdependent: each cluster's outcomes are co-determined with others' (e.g., a migration response would implicate both treated and untreated clusters). What we can estimate is the effects of the shock on agglomeration and on widening gaps between clusters that by implication follow.

3.1 Identification

A natural concern is the endogeneity of the locations and subjects of OSRD research, and the possibility that funding choices may correlate with other determinants of innovation. Contractors and R&D priorities were not randomly chosen—though to a first order, they were more product of short-run necessity than long-run technological promise.[25] OSRD's priorities were set in close collaboration with the military, though which it identified needs that could potentially be met by new technology. In some cases, it engaged in new problems (e.g., engineering controlled nuclear reactions). In other cases, it took existing problems that were stuck and pushed them forward (e.g., microwave radar). Its portfolio included projects with high uncertainty, some unsuccessful despite early enthusiasm (e.g., synthetic penicillin), and others with long odds that succeeded. Technical feasibility was a criterion in deciding how to allocate scarce inputs (especially research talent, more than funding), but OSRD's first condition for any project was that it would help win *this* war—a condition which, for example, led to the atomic bomb being prioritized over advanced rocketry, which was viewed as a weapon of future wars (Zachary 1997). In a postwar retrospective, OSRD Secretary Irvin Stewart summarized the nature of this shock:

---

[25] In concurrent writing (Gross and Sampat 2022d), we examine in detail how OSRD managed wartime R&D, including selecting priorities and R&D performers. We refer readers to this paper for a closer reading.

> The shift in emphasis and even in direction was enormous ... subjects of minor importance in peacetime become of controlling importance in war. Some subjects are born of war. (Stewart 1948, p. 102)

The qualitative evidence that short-run, military demand and the potential for immediate payoffs drove funding choices—and hence, that resource allocations were not structurally endogenous to the outcomes we study—does not exclude a possible confounding effect if these correlate with long-run demand or technological promise. The reality, however, is that OSRD discontinuously pushed out the frontier for most of the technologies it funded.[26] Many of the technologies that existed at the end of the war were barely conceived or considered impossible before it, and others were conceived or known but not commercially pursued until OSRD developed them—bridging the technological valley of death where commercial development often stalls.

Research performers were explicitly chosen on their ability to deliver high-quality results, as fast as possible (Stewart 1948). Often these were new and non-obvious: many World War II R&D problems were novel, and the U.S. lacked a deep bench of researchers with direct experience (it is telling that academic physicists led most of the major OSRD programs, rather than firms or engineers; see Kevles 1977b). Features of each R&D problem also shaped OSRD's operational choices: complex, systems engineering problems like radar were not easily divisible, and thus concentrated, whereas others were more dispersed.[27] Most importantly for our purposes, insofar as contract placement was selected, this was the case more on levels than trends.

Similar to the logic of a shift-share design, technology-specific shocks will be enough to deliver identification, even when contract placement is endogenous. Nevertheless, in the absence of a true experiment, the data are revealing. Figure 3, for example, reveals pre-war differences in levels but not in trends. We will see shortly that this pattern generalizes: at the cluster level, we find that pre-war invention in more- and less-intensively treated clusters followed parallel trends and only diverged after the OSRD shock arrived. This pattern is, broadly, consistent with our understanding of the history and the nature of the shock we have discussed.

---

[26] The radar project, for example, has been described as "five years of furious technology... [that] advanced knowledge in its field by 25 years" (Massachusetts Institute of Technology 1946).

[27] The radar example is instructive: though Boston was not an electronics hub, the Rad Lab was sited at MIT due to the presence of a handful of scientists with experience in microwaves, its ability to attract more scientists to work on the radar problem, and its proximity to an airport and to the ocean (for testing). NDRC's close connections to MIT, through Bush and Compton, may have also contributed to this choice.

3.2 Baseline effects

Figure 4, Panel (A) presents our main result, displaying β_t estimates from Equation (1) with 95% confidence intervals. Clusters with a larger OSRD shock: (i) did not grow statistically differently than in clusters with a smaller shock prior to 1940, (ii) experienced a relative surge during the war, (iii) briefly contracted from their mid-war peak when the war ended, and then (iv) experienced a sustained takeoff. The magnitudes indicate that a doubling of a cluster's OSRD rate was associated with 20% greater cluster patenting by 1960 and 30% by 1970.

[Figure 4 about here]

In Panels (B) and (C), we re-estimate Equation (1) for non-OSRD and non-government interest patents, where we see similar long-run patterns, but a smoothing out of the 1940s, indicating that the mid-1940s "bump" in Panel (A) was the OSRD shock itself.

Appendix C provides several supporting results. Appendix Figure C.2 shows similar effects for citation-weighted patents and per-capita patenting, which rules out that the results are driven by population changes—and thus indicates real impacts on local inventive productivity. Appendix Figure C.3 re-estimates Equation (1), omitting states which were top producers of OSRD invention (e.g., California, Massachusetts, New Jersey, and New York; see Appendix Table A.2), establishing that the result is not driven by any one state, county, or cluster which was prominent both in World War II and afterwards (such as Middlesex, MA). Appendix Figure C.5 reproduces Figure 4, but with estimates from Equation (2), plotting the β_t parameters for clusters in the top quartile of the OSRD shock. Patenting in these clusters is 60% higher by 1970.

In Appendix D, we show that our results are similar—if anything, more precise—when estimated for inverse hyperbolic sine (IHS) patents, which approximates the log transformation but is defined at zero, and thus includes cluster-years with no patents. In Appendix E we show that our results are the same for more aggregated geographic units such as CBSAs (core-based statistical areas). Given that our analysis window spans the pre-war to postwar era, which saw a dispersion of population and economic activity from urban centers, it is ambiguous whether a more appropriate geographic unit of analysis would be counties (for the earlier era) or CBSAs (for the later era), but it is reassuring that results are not sensitive to the choice.

3.3 Heterogeneity

The most striking implication of the results thus far is that World War II was a formative event setting in motion increasing agglomeration of inventive activity around the country, and ostensibly the takeoff of technology clusters persisting to this day. A corollary question is whether it was an equalizing force, or merely deepened existing geographic differences.

To further explore this question, we partition counties into the top 5% versus bottom 95% of 1930s patenting (by patent count). When Equation 1 is estimated for each group, it becomes apparent that the effects are entirely driven by counties which were already among the most inventive before World War II (Figure 5). Yet even in these clusters, the OSRD treatment does not coincide with any differential growth leading up to the war: the entirety of the OSRD effect takes place with the wartime surge in patenting and the postwar takeoff. The evidence thus supports an interpretation of both continuity and change, like that seen in our Massachusetts example (in Figure 3): pre-war differences persisted, but the war caused a trend shift. In simpler terms, the OSRD's effect was to catalyze long-run growth in existing geographic centers of invention.

[Figure 5 about here]

A second question is whether the OSRD effect was general across all technologies whose development it funded, or stronger for some fields over others. We evaluate this question by partitioning the sample by 1-digit NBER categories (Chemicals, Computers & Communications, Drugs & Medical, Electrical & Electronics, Mechanical, and Other). Appendix Figure C.3 re-estimates Equation (1) for each of these categories. Consistent with the history, we find that our main result is primarily (although not exclusively) driven by the electrical and electronics field, where the long-run impact of OSRD was a 40% increase in cluster patenting by 1970.

3.4 Spillovers

Thus far, our analysis presumes and estimates localized impacts of the OSRD shock in the counties and technology areas where R&D investments were made. Yet investments in specific technologies may filter down to others, including via direct linkages or shared inputs and customers. Given that spillovers may be a means through which the effects of the OSRD shock compounded for specific cities and regions, we seek to more closely examine their role.

Our focus will be on within-county spillovers across technology areas. We estimate an augmented version of Equation (2), where we include measures not only of a given cluster's treatment quartile, but also measures of whether (i) whether a “nearby” technology area was in each treatment quartile, and (ii) whether a more distant technology area was in each treatment quartile, where proximity is measured vis-à-vis 1-digit NBER categories: two technologies under the same parent category are considered proximate. We will effectively estimate a horserace regression, pitting localized shocks against nearby and more distant shocks (in technology space).

Figure 6 plots estimates (and 95% confidence intervals) for all treatment quartiles (columns) across all three types of shocks: local, nearby, and distant (rows). Standard errors increase somewhat due to reduced degrees of freedom, as we are estimating more than ten times as many parameters as in Equation (1)—but several patterns are nevertheless apparent.

[Figure 6 about here]

First, our baseline effects are largest for the top treatment quartile and attenuate at lower quartiles (matching Appendix Figure C.6, which shows the full set of parameters from Equation 2). Second, these effects are largest for localized shocks (in the same technology area). Third, we find evidence of spillovers that attenuate with technological distance. Fourth, low-treatment clusters experience *declining* invention post-World War II, suggesting that the widening regional differences we observe may have been accelerated by (or even driven by) invention migrating to heavily-treated clusters. Although the evidence thus far suggests this migration was not a result of population movements per se (Appendix Figure C.2), a postwar relocation of R&D activity may have one means through which agglomeration took place. More fully understanding how these agglomerative clusters took shape after World War II is the task we take on next.

4 Emergent Local Innovation Systems

Despite the fact that World War II was, on its own, an inherently temporary shock to the innovation system, the evidence thus far indicates that its effects not only persisted but compounded for several decades. Our question for this section is why. We structure our analysis as follows. First, we consider OSRD’s direct impacts in the form of growing postwar invention building directly on OSRD-funded research, which our evidence rules out as a driving force behind the divergence we found in Section 3. We are left with two possibilities. One is postwar government R&D investment in the same regions and technology areas which OSRD funded—i.e., a sustained push—driven by continuity in defense R&D funding structures and military need. The other is an organic growth takeoff, powered by Marshallian increasing returns to scale.

4.1 Direct follow-on to OSRD invention

We begin by exploring the more direct channels through which OSRD-funded R&D might affect local invention. The examples we have in mind include direct follow-on invention, as well as increased invention by firms or inventors who participated in the OSRD effort and developed capabilities and expertise that it could harness after the war ended. We measure follow-on invention in the form of patents which cite OSRD patents, and we measure OSRD firms as firm assignees which produced an OSRD patent. Given the challenges of longitudinal inventor linking and disambiguation, and our own hesitations in the resulting links, we do not attempt to link all OSRD inventors to their pre- and postwar patents, but rather focus on researchers at two of the largest OSRD-funded research labs (the MIT Radiation Laboratory and Harvard Radio Research Lab), which we have hand-matched to patents in concurrent research (Gross et al. 2022).

In Figure 7, we estimate Equation (1) over these categories. Panels (A1) and (A2) estimate the effects of the OSRD shock on patents that do and do not cite OSRD patents (respectively), where it is apparent that the effect is entirely driven by the latter. In Panels (B1) and (B2) we estimate the effect on patents of OSRD firms and other assignees, again finding that the effect is primarily driven by the latter. For brevity, we do not show the inventor results, though the patterns are the same—which is consistent with our priors, given the magnitudes of our effects and that many of these individuals' careers had waned by the end of our sample. Collectively, the evidence suggests against an interpretation of long-lasting direct impacts.

[Figure 7 about here]

4.2 Postwar government R&D investment

Our second hypothesis is that the postwar takeoffs we find were powered by continued government R&D investment in the same subjects and regions. In the context of the Cold War expansion of federal R&D, and the continuity in many military R&D priorities (e.g., aerospace, missiles, radar, nuclear arms), this is a natural conjecture.[28] To evaluate this question, we take two approaches. First, we re-estimate Equation (1), controlling for clusters' government-funded share of postwar patents (henceforth, the "USG rate"), crossed by year—in effect, accounting for the fact that many of these clusters remained "defense R&D places" (in the language of Kantor and Whalley 2022), which may have grown differentially in the postwar era. The effects of the OSRD shock are unchanged in size and significance. Second, we partition the sample into clusters with zero, below

---

[28] Indeed, our data indicate substantial path dependence in Cold War government R&D (Appendix Table C.3).

median (conditional on non-zero), and above median postwar USG rates, and re-estimate Equation (1) for each group. Figure 8, Panels (A) and (B) show that the OSRD shock had similar effects in clusters with higher and lower postwar government R&D invention.

[Figure 8 about here]

4.3 Increasing returns to scale

The remaining possibility we see is a Marshallian takeoff, springing from wartime R&D investments that established a collection of firms, inventors, and institutions with experience in new, frontier technologies developed in the war around which a cluster could grow.

In this case, we would expect a wide range of changes to take place. With our data, we are able to examine if—and show evidence that—the OSRD shock led to an expanding set of local, R&D-performing firms and institutions; increasing private and public invention; growth of incumbent firms, in-migration, and de novo entry; deepening linkages between local invention; and, ultimately, deconcentration of invention, as these local innovation systems grew.

We begin by examining the growth in patenting across a range of actors. In doing so, we transition from a specification with annual parameters to estimating quinquennial parameters, to simplify the presentation. Table 3 estimates the effect of the OSRD shock on firm patenting, which Column (1) shows grew significantly over the postwar period. To understand the source of these patterns, we then divide our sample into patents by incumbent firms (i.e., those with prior patents) and new firms (without prior patents), and further subdivide the incumbent firms into those with prior patents in the given cluster versus those whose prior patents were in different counties or technology areas, as we look for evidence of firms crowding into treated clusters.

[Table 3 about here]

We find growing firm patenting from multiple directions, including by cluster incumbents (Column 2), but also by local firms migrating into the cluster from other technology areas (Column 3) and more geographically-distant firms in the same technology area reallocating R&D activity to treated clusters (Column 4), as well as by new patenting firms (Column 6). We do not find a comparable effect for patenting by geographically and technologically distant firms (Column 5), suggesting that the agglomerative impacts of OSRD shock had some limits in scope.

We find similar results when outcomes are measured as the number of unique firms filing patents in the given cluster and year, rather than the number of firm patents—reflecting the broadening inventive base. Complementing, and in some cases even feeding, this firm growth was government-funded invention. The well-known history of the Silicon Valley and Boston-area clusters, for example, are rich in stories of both industry and military-led research in this era. Though government invention does not explain the effects of the OSRD shock, in Table 4 we examine to what degree it followed. Column (1) shows that government-funded invention grew rapidly after the war in clusters which OSRD itself funded—but as Column (2) conveys, non-government funded R&D grew at a similar rate. The growth in government-funded invention was entirely driven by defense R&D (Columns 3 to 6), which dominated the federal R&D budget in the postwar era.

[Table 4 about here]

In Appendix Table C.4 we examine effects on other local research-related institutions, including universities and federally-funded R&D centers (FFRDCs; i.e., independent research organizations operated by a university or firm, but funded by a mission agency), which proliferated after World War II and remain an important part of the U.S. innovation system today (e.g., Jaffe and Lerner 2001). Here our analysis is more coarse, because our treatment and outcomes are measured at the county level, when the true shock was at the cluster (county x technology area) level. A few results nevertheless stand out. We do not find an effect on university growth vis-à-vis PhD graduate production, neither overall or for broad fields like the physical sciences—though it is possible our measurement is too coarse to pick up underlying changes, particularly in fields close to OSRD research like electrical engineering. On the other hand, counties with larger OSRD investment were more likely to host postwar FFRDCs, many of which have their origins in OSRD research programs (most notably the Department of Energy's national labs, but also several electronics and aerospace research labs, many of which were borne out of OSRD research programs).

In Table 5 we examine patent citation flows, which has traditionally been applied as a proxy for knowledge spillovers—a tradition we continue, despite known limitations. We estimate, in parallel, the share of backward (forward) citations made by (accruing to) patents in a given cluster and year that are to prior (from future) patents in the same county and/or technology area, as a function of the OSRD shock. Because citations were only included in patent publications beginning in 1947, our analysis of backward citations applies to post-1947 patents. Since these citations point to earlier patents, forward citations can be measured for the full sample.

[Table 5 about here]

Broadly, Table 5 provides evidence of growing local citation flows following the OSRD shock. As a benchmark, the share of backward and forward citations that occur in the immediate neighborhood of a given patent (same county and technology area) is low, at roughly 2% and 4%, respectively—reflecting the literal definition of citations as references to prior art against which the novelty of a patent's claims are evaluated by examiners, which exists widely. However, we note a few patterns. By the late 1960s, patents in more heavily shocked clusters have a higher fraction of backward citations to others in the same cluster (an increase of roughly 15% of the mean), as well as a higher fraction to patents in the same county but a different area (up 10% of the mean), or the same area but a different county (up 2% of the mean). These patents are likewise accruing a higher fraction of their forward citations within-cluster (up 25% of the mean).

The final result we present may be viewed as a summary statistic for the collective evidence. Table 6 estimates the concentration of cluster patenting across filers as a function of the OSRD shock. In Column (1) we measure a Herfindahl index, and in Columns (2) to (5) we measure concentration ratios for the top 1, 5, 10, and 20 filers. The results suggest that the shocked clusters experienced a significant broadening of their inventive base over the postwar era, with much of this effect driven by deconcentration away from the single dominant filer. In effect, what used to be company towns became significantly more diverse in their R&D performers.

[Table 6 about here]

The evidence is broadly consistent with prior research on industrial agglomeration. An expansive literature in economic geography has consolidated around three sources of agglomeration economies (Marshall 1890, Krugman 1991), which Duranton and Puga (2004) have characterized as "sharing" (of fixed, indivisible local assets, like universities or infrastructure), "matching" (of buyers and sellers of goods and labor), and "learning" (knowledge spillovers). Though as Duranton and Puga (2004) point out, these mechanisms are typically observationally equivalent, they provide useful structure for interpreting our results. The growth of local R&D-performing firms and institutions (insofar as we can measure them), including (i) firms migrating into the treated clusters from other locations and technology areas, and (ii) government agencies locating labs and contracting with firms in these clusters, is consistent with the local advantages borne out of assets like large talent pools, financial capital, and research facilities. This density also supports more efficient matching, particularly through labor mobility.[29] Insofar as patent citations may reflect intellectual linkages, we explicitly find evidence of growing knowledge spillovers.

---

[29] We do not document labor mobility directly, due to data limitations: linked employee panels are difficult to construct for this period. We forgo building a linked inventor panel across the universe of inventors due to hesitations with the quality of the links we can make, including selection into linking, disambiguation challenges, and the sensitivity

5 Downstream: Entrepreneurship and Employment

Did the growth of these postwar innovation clusters have broader impacts on local economies? The downstream effects of local and regional R&D investments is a first-order question not only for research on agglomeration but also for policy to improve local economic performance through place-based public R&D investments (Glaeser and Hausman 2020), including implementation of recently-enacted U.S. legislation that aims to develop a more geographically-distributed innovation system. What are the downstream effects of big, applied push R&D investments on local population, firm creation, employment, wages or other economic outcomes?

Research has increasingly begun to speak to these questions, especially in the context of Cold War-era R&D shocks. Schweiger et al. (2022), for example, show that Soviet "Science Cities"—R&D centers created by the Soviet government in the mid-twentieth century to support strategic R&D in then-critical technologies—today have higher educational attainment, higher skilled employment, higher incomes, and more patents. Kantor and Whalley (2022) show that in U.S. counties which were target locations for Space Race R&D, manufacturing employment, output, and productivity grew more quickly in and after the Space Race era. Here we have an opportunity to complement this literature by examining the impacts of public investments in specific technologies on the industries that produce them, and with a distinct, broad, primordial shock that provides measurable variation both within and across counties, technologies, and industries.

We use CBP data to measure local employment in select industries in 1980, and D&B data to measure local business creation in these industries from 1920 to 1980. Crucially, we link industries to the patent data (where we observe the OSRD shock) using a USPTO-produced crosswalk, which concords both SIC industries and patent subclasses to a common set of 41 unique industry codes. Several of these codes group up into an "Electrical and Electronic Equipment and Supplies" category—including the industry code with the most associated OSRD patents ("Electronic components and accessories and communications equipment"). This, together with prior evidence that the electrical and electronics area is where the OSRD shock had bite (Appendix Figure C.3) and the broader growth of this area in the postwar period, motivates our focus on this category, and the analysis below will be performed across counties and industries in the electrical field. Appendix B lists the complete set of industries included in this sample.

We first explore OSRD's long-run, downstream employment effects. For this analysis, we estimate the effects of both extensive and intensive treatment measures on industry employment. Where

---

of mobility measurement to the standardization of assignee names, firm reorganization, and name changes—as well as limited power, given that patenting is a rare event for most individuals, and the median inventor has one patent. We instead note that improved matching is a corollary of thick labor markets.

employment counts are suppressed by the CBP (e.g., for small county-industry cells which pose a risk of disclosure), we impute employment from the establishment size distribution (in the spirit of Duranton et al. 2014). To accommodate sparse samples, we replace log transformations with inverse hyperbolic sine transformations, which retains zeros in the explanatory and outcome variables but otherwise resembles the shape of our standard approach, and in successive specifications we control for counties' manufacturing employment (across all manufacturing SICs) and total employment (across all SICs). Formally, we estimate the regression below:

$$Ln(Employment)_{ci} = \beta \cdot OSRD\ Treatment_{ci} + \alpha_c + \gamma_i + X_{ci}\phi + \varepsilon_{it} \tag{3}$$

where c and i index counties and industries, α_c and γ_i are fixed effects, X_{it} are controls, and standard errors are clustered at the county level. Because employment in a given county and industry is determined in equilibrium with others, the results we obtain under this approach should not be interpreted as a multiplier on R&D (as in Kantor and Whalley 2022), but rather as divergence: without more structure, we are unable to distinguish net job growth from share-stealing, either of which could increase differences between counties. As with the rest of this paper, however, our goal is to evaluate the degree to which the OSRD shock led economic activity to agglomerate in the treated clusters and widened gaps in economic performance.

Table 7 presents the results. Columns (1) to (3) show that county-industries which engaged in OSRD R&D have roughly 90% greater employment in associated manufacturing industries in 1980 than those which did not, while Columns (4) to (6) suggest that a doubling of the OSRD rate is associated with a more than doubling of manufacturing employment. These results should be interpreted with some caution, however, given that we do not observe the pre-war period and the relationship could potentially be endogenous—though it is helpful that many of these industries (e.g., electronic components) did not exist until the postwar era.

[Table 7 about here]

In Table 8 we repeat this analysis for firm creation. Here we replace county and industry fixed effects with county-industry fixed effects, exploiting the longer panel, and estimate differences relative to 1920 (the omitted category). Year fixed effects also serve an important role in this context, given that the sample of firms is conditioned on survival to 1980, and earlier decades have fewer firms in 1980 due to intervening exits. Columns (1) to (3) indicate that counties that produced OSRD patents were increasingly likely to produce firms in associated industries, particularly during and after the 1940s, though we see a (quantitatively modest) pre-trend in the 1930s. Columns (4) to (6) indicate that a doubling of the OSRD rate is associated with a steady increase in manufacturing employment over the postwar era, from 15% to 25% in the 1940s and 1950s to a more than doubling by the 1970s, with no visible pre-trend on the intensive margin.

[Table 8 about here]

6 Aggregate U.S. Invention

We are also interested to know whether the effects we have found in this paper roll up to an impact on aggregate invention. We take two closely-related approaches to answering this question: we first compare the composition of patent filings at the USPTO and at other Allied countries (UK and France) across IPC classes (131 in total) before and after the war. Using our data on the location of inventors on U.S. patents, we also aggregate USPTO filings to USPCs (429 in total) and compare filings by domestic versus foreign inventors across classes, over time.

As before, our treatment measure is a technology area's 1940s OSRD rate (fraction of U.S. patents filed between 1941 and 1948 that were OSRD-funded). Our first specification estimates a triple-difference, comparing U.S. and foreign annual log patent filings, before and after war, in classes in different quartiles of the OSRD rate (conditional on ≥1 OSRD patent) and an omitted category for classes with no OSRD patents. This specification is estimated over a sample of country-class-years from 1930 to 1970, omitting the years 1940 to 1945, as below:

$$\begin{aligned} Ln(Patents)_{ict} = &\sum_{q=1}^{4} \beta_q \cdot (\text{Country } i = \text{US}) \cdot (\text{Class } c \in \text{quartile } q) \cdot (t > 1945) \\ &+ Country_i \times Class_c + Country_i \times Post_t + Class_c \times Post_t + \varepsilon_{ict} \end{aligned} \tag{4}$$

where i, c, and t index countries (grouped up to U.S. vs. foreign), technology classes, and years, and standard errors are clustered at the country-class level.[30]

Figure 9, Panel (A) plots the estimates (with 95% confidence intervals). U.S. filings in the most heavily-treated classes are over 40 percent higher after the war than in comparison countries, and this effect attenuates in both magnitude and significance as treatment intensity declines. Panel (B) illustrates qualitatively similar results for USPTO patents of domestic and foreign inventors, but

---

[30] Because historical PATSTAT data only provide grant (not filing) dates, t indexes grant years for the U.S. versus foreign patent authority comparisons, where we also restrict to patents with a family size of one (to ensure we are measuring the primary location), although the results are generally not sensitive to this restriction. For domestic versus foreign USPTO patents, we measure filing dates, and t indexes filing years.

with much larger magnitudes, and suggests that U.S. invention in low treatment intensity classes may have even declined slightly relative to that of foreign inventors.

[Figure 9 about here]

In Appendix C we estimate variant of Equation 4 with annual coefficients. Appendix Figure C.7 presents the estimates for patenting at USPTO versus foreign patent authorities, and Appendix Figure C.8 for USPTO patents with domestic and foreign inventors. We find similar patterns to those throughout the paper: pre-war patenting in the U.S. (or by U.S. inventors, respectively) in the most-heavily treated classes is on similar trends to foreign patenting but jumps after the war. These results are similar when estimated with continuous rather than binned treatment measures and for other transformations of the outcome variable.

7 Concluding Remarks

> "Although its part in the winning of the war was its greatest contribution... the full impact of its work must await the judgment of the future..."
>
> — Irvin Stewart, Secretary of the OSRD, 1948

Despite a large historical and science policy literature on the run effects of OSRD on the institutions of postwar science policy, we believe this paper to be the first quantitative empirical assessment of the long-run effects of what President Roosevelt called a "unique experiment" in research policy (Roosevelt 1944) on innovation and other economic outcomes.

With newly digitized archival data on OSRD contracts, linked to data on postwar patenting, firm creation, and employment, we found persistent effects of the World War II R&D shock on technology clusters. Treated clusters were producing another 40% to 50% more patents annually than untreated clusters by 1970, despite parallel trends in patenting before the war. In exploring mechanisms, we rule out that this was due to patenting by OSRD contractors themselves, or to patents citing OSRD patents. We also used newly digitized data on the history of government patenting to show that the effects are not driven by follow-on government research investment in the same technology clusters. Instead, our evidence suggests that the effects are due to Marshallian agglomeration: with growing patenting by new and older firms, public and private, in-migrant firms and established firms, and with innovation becoming increasingly dispersed over time. Beyond patents, we also find evidence that postwar firm creation and employment were higher in OSRD treated counties, decades out.

Finally, there was also an aggregate shift in the trajectory of U.S. innovation towards the most heavily treated classes: electronics and communications technologies.

The results provide new evidence on the persistent impact of a large and broad applied R&D shock on innovation, complementing a large and growing body of evidence on the returns to publicly funded research. With a historical lens, we are able to observe these long-run effects. Moreover, rich OSRD records and new patent data allowed us to examine the mechanisms of persistence. The finding that that discrete R&D shocks can spur cycles of agglomeration is, we believe, new to both the economics of innovation and agglomeration literatures.

There are also implications for practice and policy. At a high level, our results support Vannevar Bush's argument that federally-funded research can fuel innovation and improve economic performance. But whereas Bush argued for funding "basic" research in *Science, The Endless Frontier*, OSRD's funding was primarily for applied R&D. Our results suggest that large-scale federal investments in applied research can also have large returns, or at least did in this case, potentially important given a resurgence of interest in "mission-oriented" R&D.

Our results also suggest that on concentration and inequality, Bush's nemesis Harley Kilgore was right to be concerned. Much of the OSRD support was directed to researchers at elite institutions and research labs (Appendix Table A.3). Lacking a counterfactual, it is difficult to know whether the elite funding model of OSRD was the most efficient one for wartime—were there, literally, any lost Einsteins?—but it is also hard to argue with the results. Nevertheless, Kilgore's concerns about persistent concentration of innovative activity and economic power generated by such an approach seems prescient, given the results of this paper suggesting they fueled agglomeration. Gruber and Johnson (2019) have argued that broader funding, even if it reduces efficiency, could promote not only equity but also geographically-diffuse public and political buy-in for increasing federal research spending, a question we hope to explore in future research.

References

Agarwal, Rajshree, Seojin Kim, and Mahka Moeen. 2021. “Leveraging Private enterprise: incubation of new industries to address the public sector’s mission-oriented grand challenges,” *Strategy Science*, Vol. 6, No. 4, pp. 385–411.

Almeida, Paul and Bruce Kogut. 1999. “Localization of knowledge and the mobility of engineers in regional networks,” *Management Science*, Vol. 45, No. 7, pp. 905–917.

Andrews, Michael J. 2021. “How Do Institutions of Higher Education Affect Local Invention? Evidence from the Establishment of US Colleges,” *American Economic Journal: Economic Policy*.

Arora, Ashish, Sharon Belenzon, Konstantin Kosenko, Jungkyu Suh, and Yishay Yafeh. 2021. *The Rise of Scientific Research in Corporate America*. NBER Working Paper No. 29260.

Arora, Ashish, Sharon Belenzon, and Honggi Lee. 2018. “Reversed citations and the localization of knowledge spillovers,” *Journal of Economic Geography*, Vol. 18, No. 3, pp. 495–521.

Arrow, Kenneth. 1962. “Economic Welfare and the Allocation of Resources for Invention,” in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton: Princeton University Press, pp. 609–626.

Arthur, W. Brian. 1990. “‘Silicon Valley’ locational clusters: when do increasing returns imply monopoly?” *Mathematical Social Sciences*, Vol. 19, No. 3, pp. 235–251.

Audretsch, David B. and Maryann P. Feldman. 1996. “R&D spillovers and the geography of innovation and production,” *American Economic Review*, Vol. 86, No. 3, pp. 630–640.

———. 2004. “Knowledge spillovers and the geography of innovation,” in *Handbook of regional and urban economics*, Vol. 4, pp. 2713–2739.

Azoulay, Pierre, Erica Fuchs, Anna P Goldstein, and Michael Kearney. 2019a. “Funding breakthrough research: promises and challenges of the “ARPA Model”,” *Innovation Policy and the Economy*, Vol. 19, No. 1, pp. 69–96.

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat. 2019b. “Public R&D investments and private-sector patenting: Evidence from NIH funding rules,” *Review of Economic Studies*, Vol. 86, No. 1, pp. 117–152.

Belenzon, Sharon and Larisa C. Cioaca. 2021. *Guaranteed Markets and Corporate Scientific Research*. NBER Working Paper No. 28644.

Bergeaud, Antonin and Verluise Cyril. 2022. “PatentCity: a dataset to study the location of patents since the 19th century,” URL: https://doi.org/10.7910/DVN/PG6THV.

Bergeaud, Antonin, Arthur Guillouzouic Le Corff, Emeric Henry, and Clement Malgouyres. 2022. *From Public to Private: Magnitude and Channels of R&D Spillovers*. Working paper.

Berkes, Enrico. 2018. *Comprehensive Universe of U.S. Patents (CUSP): Data and Facts*. Working paper.

Berkes, Enrico and Ruben Gaetani. 2021. “The geography of unconventional innovation,” *Economic Journal*, Vol. 131, No. 636, pp. 1466–1514.

Biasi, Barbara and Petra Moser. 2021. “Effects of copyrights on science: Evidence from the WWII book republication program,” *American Economic Journal: Microeconomics*, Vol. 13, No. 4, pp. 218–60.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. “Identifying technology spillovers and product market rivalry,” *Econometrica*, Vol. 81, No. 4, pp. 1347–1393.

Bloom, Nicholas, John Van Reenen, and Heidi Williams. 2019. “A toolkit of policies to promote innovation,” *Journal of Economic Perspectives*, Vol. 33, No. 3, pp. 163–84.

Bryan, Kevin A. and Heidi L. Williams. 2021. “Innovation: market failures and public policies,” in *Handbook of Industrial Organization*, Vol. 5, pp. 281–388.

Buenstorf, Guido and Steven Klepper. 2009. “Heritage and agglomeration: the Akron tyre cluster revisited,” *Economic Journal*, Vol. 119, No. 537, pp. 705–733.

Bush, Vannevar. 1945. *Science, the Endless Frontier: A report to the President*. Washington: Government Printing Office.

Buzard, Kristy, Gerald A. Carlino, Robert M. Hunt, Jake K. Carr, and Tony E. Smith. 2017. “The agglomeration of American R&D labs,” *Journal of Urban Economics*, Vol. 101, pp. 14–26.

Carlino, Gerald and William R. Kerr. 2015. “Agglomeration and innovation,” *Handbook of Regional and Urban Economics*, Vol. 5, pp. 349–404.

Chattergoon, Brad and William R. Kerr. 2022. “Winner takes all? Tech clusters, population centers, and the spatial transformation of US invention,” *Research Policy*, Vol. 51, No. 2, p. 104418.

Chatterji, Aaron, Edward Glaeser, and William Kerr. 2014. “Clusters of entrepreneurship and innovation,” *Innovation Policy and the Economy*, Vol. 14, No. 1, pp. 129–166.

Cohen, Wesley M. 2010. “Fifty years of empirical studies of innovative activity and performance,” in *Handbook of the Economics of Innovation*, Vol. 1, pp. 129–213.

Duranton, Gilles. 2007. “Urban evolutions: The fast, the slow, and the still,” *American Economic Review*, Vol. 97, No. 1, pp. 197–221.

Duranton, Gilles, Peter M Morrow, and Matthew A Turner. 2014. “Roads and Trade: Evidence from the US,” *Review of Economic Studies*, Vol. 81, No. 2, pp. 681–724.

Duranton, Gilles and Diego Puga. 2004. “Micro-foundations of urban agglomeration economies,” in *Handbook of regional and urban economics*, Vol. 4, pp. 2063–2117.

Feldman, Maryann P. 1994. *The Geography of Innovation*. Dordrecht: Kluwer Academic Publishers.

Field, Alexander J. 2022. *The Economic Consequences of US Mobilization for the Second World War*. New Haven: Yale University Press.

Fishback, Price and Joseph A Cullen. 2013. “Second World War Spending and Local Economic Activity in US Counties, 1939-58,” *The Economic History Review*, Vol. 66, No. 4, pp. 975–992.

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.

Geiger, Roger L. 1993. *Research and relevant knowledge: American research universities since World War II*. Oxford: Oxford University Press.

Glaeser, Edward L. and Naomi Hausman. 2020. “The spatial mismatch between innovation and joblessness,” *Innovation Policy and the Economy*, Vol. 20, No. 1, pp. 233–299.

Gross, Daniel P. 2022. “The Hidden Costs of Securing Innovation: The Manifold Impacts of Compulsory Invention Secrecy,” *Management Science*.

Gross, Daniel P., Maria P. Roche, and Bhaven N. Sampat. 2022. *Coordinated R&D Programs and the Creation of New Industries*. Working paper.

Gross, Daniel P. and Bhaven N. Sampat. 2021. "The economics of crisis innovation policy: A historical perspective," *AEA Papers & Proceedings*, Vol. 111, pp. 346–450.

——— . 2022a. *A Century of Government-funded Invention*. Working paper.

——— . 2022b. "Crisis innovation policy from World War II to COVID-19," *Entrepreneurship and Innovation Policy and the Economy*, Vol. 1, No. 1, pp. 135–181.

——— . 2022c. *A Novel Experiment: The Long-Run Effects of the World War II Medical Research Effort on Science, Technology, and Practice*. Working paper.

——— . 2022d. *The World War II crisis innovation model: What was it, and where does it apply?*. NBER Working Paper No. 27909.

Gruber, Jon and Simon Johnson. 2019. *Jump-starting America: How Breakthrough Science Can Revive Economic Growth and the America Dream*. New York: PublicAffairs Press.

Gruber, Jonathan, Simon Johnson, and Enrico Moretti. 2022. *Place-Based Productivity and Costs in Science*. NBER Working Paper No. 30416.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2001. *The NBER patent citation data file: Lessons, insights and methodological tools*. NBER Working Paper No. 8498.

Howell, Sabrina T. 2017. "Financing innovation: Evidence from R&D grants," *American Economic Review*, Vol. 107, No. 4, pp. 1136–64.

Howell, Sabrina T., Jason Rathje, John Van Reenen, and Jun Wong. 2021. *Opening up military innovation: Causal effects of 'bottom-up'reforms to US defense research*. SSRN working paper, available at https://ssrn.com/abstract=3825034.

Iaria, Alessandro, Carlo Schwarz, and Fabian Waldinger. 2018. "Frontier knowledge and scientific production: Evidence from the collapse of international science," *Quarterly Journal of Economics*, Vol. 133, No. 2, pp. 927–991.

Jaffe, Adam B and Josh Lerner. 2001. "Reinventing public R&D: Patent policy and the commercialization of national laboratory technologies," *RAND Journal of Economics*, pp. 167–198.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic localization of knowledge spillovers as evidenced by patent citations," *Quarterly Journal of Economics*, Vol. 108, No. 3, pp. 577–598.

Jaworski, Taylor. 2017. "World War II and the industrialization of the American South," *The Journal of Economic History*, Vol. 77, No. 4, pp. 1048–1082.

Jones, Benjamin F. and Lawrence H. Summers. 2021. "A calculation of the social returns to innovation," in Goolsbee, Austan and Benjamin F. Jones eds. *Innovation and Public Policy*, Chicago: University of Chicago Press.

Jones, Charles I. and John C. Williams. 1998. "Measuring the social return to R&D," *Quarterly Journal of Economics*, Vol. 113, No. 4, pp. 1119–1135.

Kantor, Shawn and Alexander Whalley. 2014. "Knowledge spillovers from research universities: evidence from endowment value shocks," *Review of Economics and Statistics*, Vol. 96, No. 1, pp. 171–188.

——— . 2019. "Research proximity and productivity: long-term evidence from agriculture," *Journal of Political Economy*, Vol. 127, No. 2, pp. 819–854.

——— . 2022. *Moonshot: Public R&D and Growth*. Working paper.

Kerr, William R. 2010. "Breakthrough inventions and migrating clusters of innovation," *Journal of Urban Economics*, Vol. 67, No. 1, pp. 46–60.

Kerr, William R. and Frédéric Robert-Nicoud. 2020. “Tech clusters,” *Journal of Economic Perspectives*, Vol. 34, No. 3, pp. 50–76.

Kevles, Daniel J. 1977a. “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.

——— . 1977b. *The physicists: The history of a scientific community in modern America*. New York: Alfred A. Knopf.

Kleinman, Daniel Lee. 1995. *Politics on the endless frontier: Postwar research policy in the United States*. Durham: Duke University Press.

Klepper, Steven. 2010. “The origin and growth of industry clusters: The making of Silicon Valley and Detroit,” *Journal of Urban Economics*, Vol. 67, No. 1, pp. 15–32.

——— . 2016. *Experimental Capitalism: The Nanoeconomics of American High-Tech Industries*. Princeton: Princeton University Press.

Kline, Patrick and Enrico Moretti. 2014a. “Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority,” *Quarterly Journal of Economics*, Vol. 129, No. 1, pp. 275–331.

——— . 2014b. “People, places, and public policy: Some simple welfare economics of local economic development programs,” *Annual Review of Economics*, Vol. 6, No. 1, pp. 629–662.

Krugman, Paul. 1991. “Increasing returns and economic geography,” *Journal of political economy*, Vol. 99, No. 3, pp. 483–499.

Maddox, Robert Franklin. 1981. *The Senatorial Career of Harley Martin Kilgore.*: Taylor & Francis.

Marco, Alan C., Michael Carley, Steven Jackson, and Amanda Myers. 2015. *The USPTO Historical Patent Data Files: Two Centuries of Innovation*. SSRN working paper, available at https://ssrn.com/abstract=2616724.

Marglin, Stephen A. and Juliet B. Schor. 1991. *The golden age of capitalism: reinterpreting the postwar experience*. Oxford: Oxford University Press.

Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan.

Massachusetts Institute of Technology. 1946. *Five Years at the Radiation Laboratory*.

Moretti, Enrico. 2021. “The effect of high-tech clusters on the productivity of top inventors,” *American Economic Review*, Vol. 111, No. 10, pp. 3328–75.

Moretti, Enrico, Claudia Steinwender, and John Van Reenen. 2022. “The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers,” *Review of Economics and Statistics*.

Moser, Petra and Alessandra Voena. 2012. “Compulsory licensing: Evidence from the Trading with the Enemy Act,” *American Economic Review*, Vol. 102, pp. 396–427.

Mowery, David C. 1997. “The Bush report after 50 years: Blueprint or relic?” in Barfield, Claude E. ed. *Science for the 21st century: The Bush Report revisited*, Washington: American Enterprise Institute.

——— . 2010. “Military R&D and innovation,” in *Handbook of the Economics of Innovation*, Vol. 2, pp. 1219–1256.

Myers, Kyle. 2020. “The elasticity of science,” *American Economic Journal: Applied Economics*, Vol. 12, No. 4, pp. 103–34.

Myers, Kyle R. and Lauren Lanahan. 2022. "Estimating spillovers from publicly funded R&D: Evidence from the US Department of Energy," *American Economic Review*, Vol. 112, No. 7, pp. 2393–2423.

Nelson, Richard R. 1997. "Why the Bush Report has hindered and effective civilian technology policy," in Barfield, Claude E. ed. *Science for the 21st century: The Bush Report revisited*, Washington: American Enterprise Institute.

Petralia, Sergio, Pierre-Alexandre Balland, and David L. Rigby. 2016. "Unveiling the geography of historical patents in the United States from 1836 to 1975," *Scientific data*, Vol. 3, No. 1, pp. 1–14.

Pursell, Carroll. 1979. "Science agencies in World War II: The OSRD and its challengers," *The Sciences in the American Context: New Perspectives*, pp. 359–78.

Romer, Paul M. 1986. "Increasing returns and long-run growth," *Journal of Political Economy*, Vol. 94, No. 5, pp. 1002–1037.

——— . 1990. "Endogenous technological change," *Journal of Political Economy*, Vol. 98, No. 5, Part 2, pp. S71–S102.

Roosevelt, Franklin D. 1944. *Letter to Vannevar Bush*. Dated November 17.

Ruttan, Vernon W. 2006. *Is war necessary for economic growth? Military procurement and technology development*. Oxford: Oxford University Press.

Santoleri, Pietro, Andrea Mina, Alberto Di Minin, and Irene Martelli. 2022. "The causal effects of R&D grants: Evidence from a regression discontinuity," *Review of Economics and Statistics*.

Saxenian, AnnaLee. 1996. *Regional advantage: Culture and competition in silicon valley and route 128, with a new preface by the author*. Cambridge: Harvard University Press.

Schweiger, Helena, Alexander Stepanov, and Paolo Zacchia. 2022. "The long-run effects of R&D place-based policies: evidence from Russian science cities," *American Economic Journal: Economic Policy*, Vol. 14, No. 3, pp. 322–51.

Stewart, Irvin. 1948. *Organizing scientific research for war: The administrative history of the Office of Scientific Research and Development*. Boston: Little, Brown, and Company.

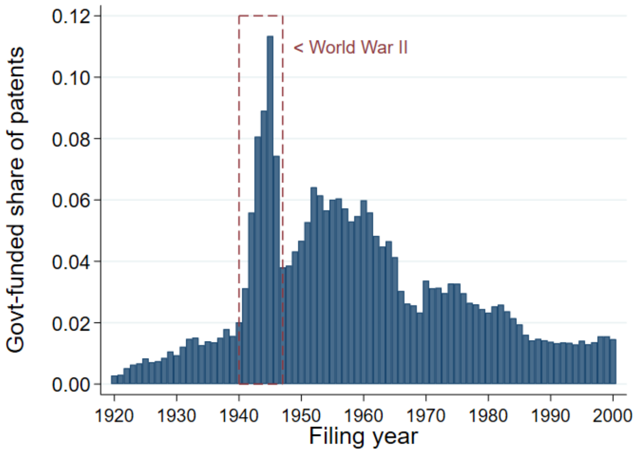
Streifer, Bill. 2017. *U.S. Department of Energy List of U.S. Patents Related to the Manhattan Project*. Obtained via FOIA request by Bill Streifer and available at the Wilson Center History and Public Policy Program Digital Archive, https://digitalarchive.wilsoncenter.org/document/165247.

Thompson, Peter. 2006. "Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations," *Review of Economics and Statistics*, Vol. 88, No. 2, pp. 383–388.

Thompson, Peter and Melanie Fox-Kean. 2005. "Patent citations and the geography of knowledge spillovers: A reassessment," *American Economic Review*, Vol. 95, No. 1, pp. 450–460.

Zachary, G. Pascal. 1997. *Endless frontier: Vannevar Bush, engineer of the American Century*. Cambridge: MIT Press.

Figure 1: Government-funded share of U.S. patent applications, 1920 to 2000

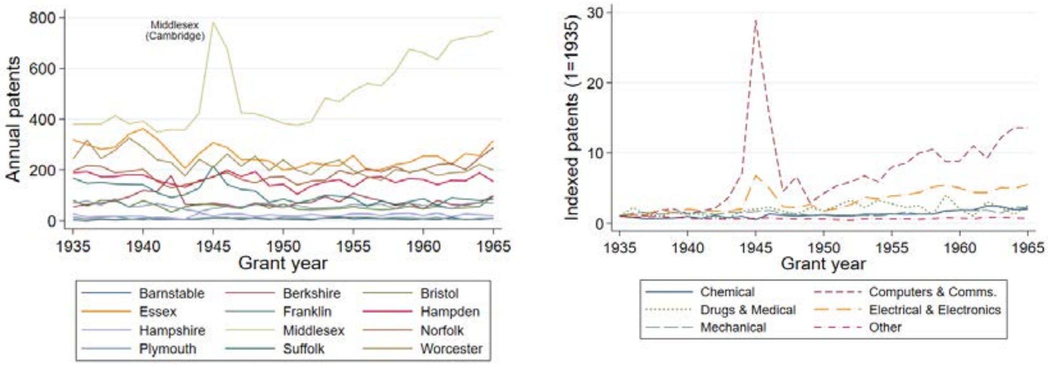


Notes: Figure plots the government-funded share of annual U.S. patent applications, using administrative data. World War II was the peak intensity of government-funded invention in U.S. history. See appendix for data details.

Figure 2: Geography of OSRD-funded invention in World War II

Notes: Figure maps counties with OSRD-funded patents. Bubble sizes proportional to each county's total number of OSRD patents.

Figure 3: Annual patenting in 12 Massachusetts counties, 1935 to 1965



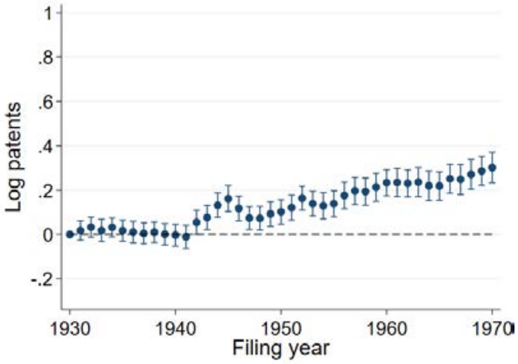
Panel (A): Twelve largest counties (all technology areas)

Panel (B): Middlesex, MA, (by technology area, indexed)

Notes: Figure shows total annual patents filed in the top 12 Massachusetts counties. The figure illustrates, for Middlesex County (location of Cambridge, home to Harvard and MIT): (i) relatively constant, pre-1940 level differences in patenting; (ii) a mid-1940s spike (doubling) of patenting, driven by war-related research; (iii) a return to approximately pre-war levels; and (iv) a take-off in the early 1950s. The raw data illustrate the general pattern that we find throughout the paper.

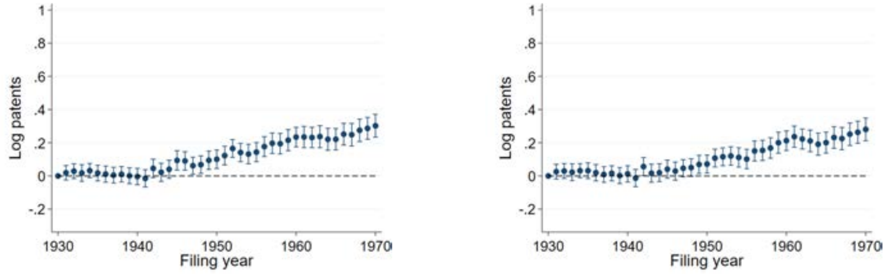
Figure 4: Effects of OSRD on cluster patenting, 1930-1970

Panel (A): All patents



Panel (A): Non-OSRD patents

Panel (B): Non-govt. interest patents

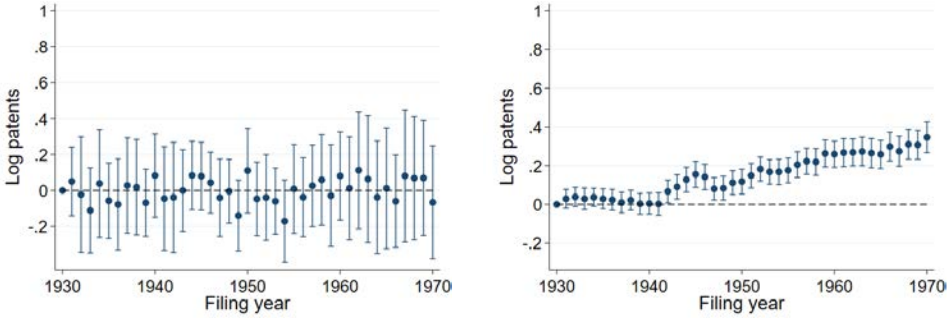


Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting. The independent variable measures the log fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure 5: Effects of OSRD on cluster patenting, for clusters in counties in the bottom 95% versus top 5% of 1930s patenting

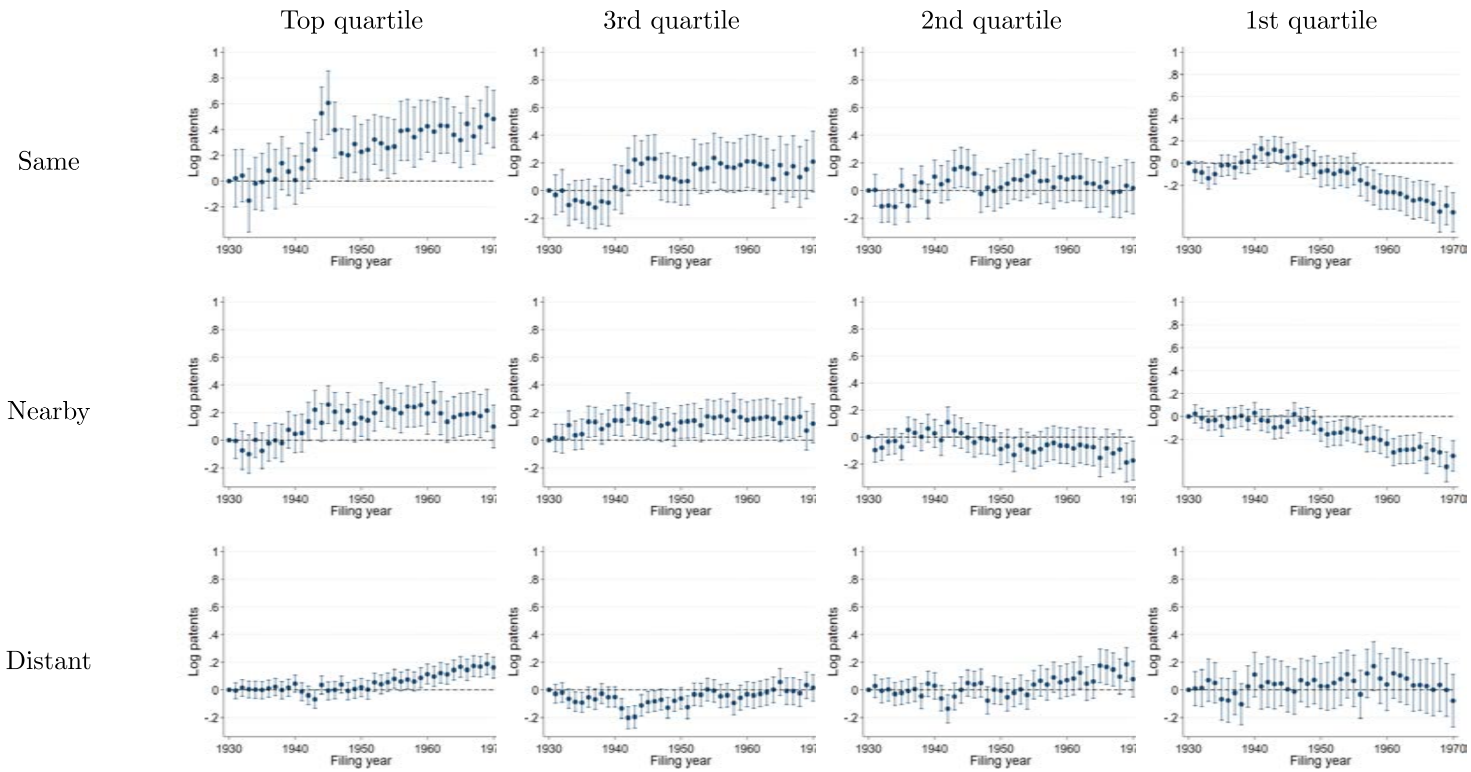
Panel (A): Bottom 95%

Panel (B): Top 5%



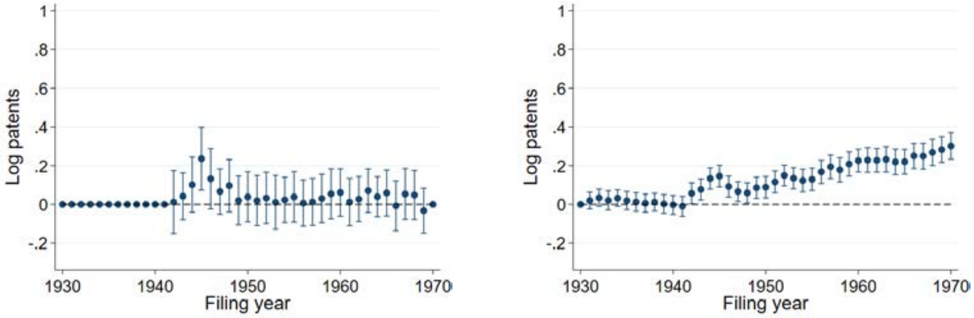
Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting, for counties in the bottom 95% and top 5% of 1930s patenting (i.e., existing technology clusters). Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure 6: Effects of OSRD on cluster patenting, 1930-1970, cross-technology area spillovers
horserace regression of treatment in (i) same technology area, (ii) nearby technology areas, (iii) more distant technology areas



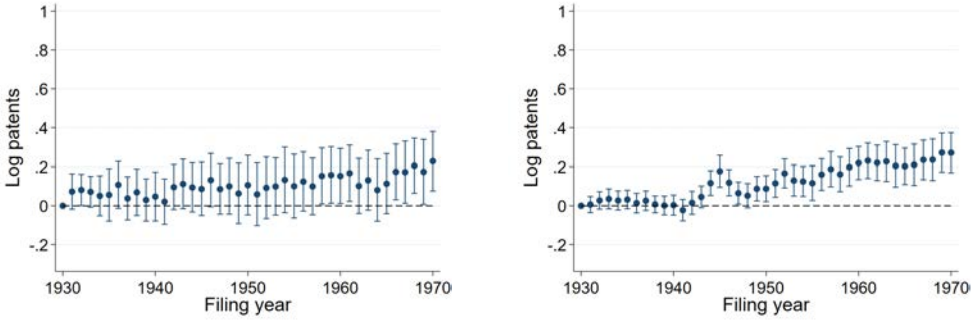
Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting. The independent variable measures the quartile of treatment intensity, conditional on treatment (the fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded, conditional on any), of three types: (i) in the given county-category (top row); (ii) in other, similar county-categories (same 1-digit NBER category, per Hall et al. (2001); middle row); and (iii) in other, more distant county-categories (other 1-digit NBER categories; bottom row). Parameters across all panels are estimated jointly (in one regression) relative to a reference group of county-categories without any OSRD patents. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure 7: Effects not explained by OSRD's direct impacts on local invention



Panel (A1):
Patents citing OSRD patents

Panel (A2):
Patents not citing OSRD patents

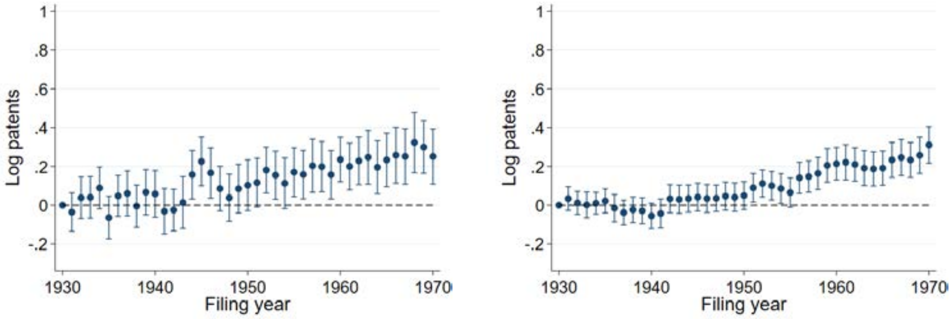


Panel (B1):
Patents assigned to OSRD contractors

Panel (B2):
Patents assigned to others

Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category (i) patents citing versus not citing OSRD patents, and (ii) patents assigned to OSRD contractors versus others, as an exploration of the direct impacts of OSRD on postwar invention in the treated clusters. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure 8: Effects not explained by sustained government investment in local invention

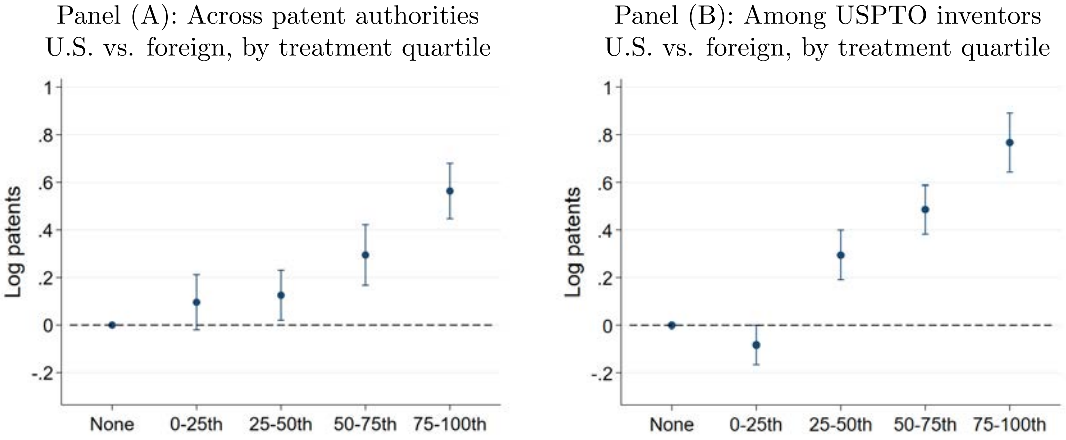


Panel (A):
High postwar govt. patent clusters

Panel (B):
Low postwar govt. patent clusters

Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patents in county-categories with above and below median postwar (1950-1969) government-funded patent rates, as an exploration of the role of sustained public R&D investment as an explanation for persistence. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure 9: U.S. versus foreign patenting in each quartile of treated patent classes (difference-in-differences, pre-1940 vs. post-1945)


Panel (A): Across patent authorities U.S. vs. foreign, by treatment quartile

Panel (B): Among USPTO inventors U.S. vs. foreign, by treatment quartile

Notes: Figure shows difference-in-difference estimates of the effects of the OSRD shock on U.S. versus foreign patenting, in technology classes (IPC classes in Panel A, USPCs in Panel B) at different quartiles of OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-funded. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-class level.

Table 1: Top 10 patent classes of OSRD patents (denominator: OSRD patents)

		OSRD patents		1933-40 patents	
USPC	Description	Percent	Rank	Percent	Rank
342	Directive radio wave systems/devices (radar)	6.6%	1	0.2%	167
102	Ammunition and explosives	5.8%	2	0.2%	170
315	Electric lamp and discharge devices: systems	4.8%	3	0.6%	302
250	Nuclear energy	4.0%	4	0.1%	117
333	Wave transmission lines and networks	3.6%	5	0.2%	164
343	Radio wave antennas	3.4%	6	0.2%	141
423	Inorganic chemistry	3.2%	7	0.7%	309
367	Acoustic wave systems/devices	3.1%	8	0.1%	79
324	Electricity: measuring and testing	3.0%	9	0.5%	284
327	Misc. electrical devices, circuits, and systems	2.9%	10	0.1%	85

Notes: Table lists the top patent classes of OSRD patents, alongside their share of OSRD patents and of post-Depression 1930s patents for comparison.

Table 2: Top clusters with OSRD patents, 1941-1948 (select technology areas)

Technology area: All

		OSRD patents, 1941-1948	
Rank	County	Number	Share of cluster
1	Middlesex, MA	446	12.3%
2	Essex, NJ	139	2.5%
3	Mercer, NJ	129	10.2%
4	Cook, IL	121	0.7%
5	Alameda, CA	98	3.7%

Technology area: Communications (21)

		OSRD patents, 1941-1948	
Rank	County	Number	Share of cluster
1	Middlesex, MA	216	45.5%
2	Mercer, NJ	37	14.3%
3	Suffolk, MA	35	35.7%
4	Essex, NJ	31	6.8%
5	Suffolk, NY	27	9.2%

Technology area: Electrical lighting (41)

		OSRD patents, 1941-1948	
Rank	County	Number	Share of cluster
1	Essex, NJ	39	10.1%
2	Middlesex, MA	38	14.4%
3	Mercer, NJ	25	17.5%
4	Schenectady, NY	17	8.5%
5	Allen, IN	9	22.5%

Technology area: Electrical devices (42)

		OSRD patents, 1941-1948	
Rank	County	Number	Share of cluster
1	Middlesex, MA	61	21.7%
2	Nassau, NY	25	10.4%
3	Washington, DC	13	7.4%
4	Suffolk, NY	12	5.7%
5	Suffolk, MA	11	15.1%

Technology area: Measuring, testing (43)

		OSRD patents, 1941-1948	
Rank	County	Number	Share of cluster
1	Monroe, NY	22	20.0%
2	Middlesex, MA	20	16.9%
3	Nassau, NY	18	13.0%
4	Harris, TX	9	7.1%
5	Los Angeles, CA	9	3.5%

Technology area: Nuclear, X-rays (44)

		OSRD patents, 1941-1948	
Rank	County	Number	Share of cluster
1	Alameda, CA	56	68.3%
2	Cook, IL	41	28.9%
3	Santa Fe, NM	14	66.7%
4	Anderson, TN	8	17.0%
5	Mercer, NJ	7	24.1%

Notes: Table lists the top clusters in select technology areas by number of OSRD patents and the share of local patents which were OSRD funded, Displayed technology areas are shown alongside their NBER technology subcategory (Hall et al. 2001) and selected due to their prominence or importance to OSRD's agenda.

Table 3: Firm patents: All, incumbents, and entrants
(incumbents by geographic and technological proximity)

		Same county		Diff. county		
	(1) All	(2) Same field	(3) Diff. field	(4) Same field	(5) Diff. field	(6) New
Ln(OSRD rate) * 1(1935-1939)	-0.007	-0.006	0.016	0.013	-0.013	0.023
	(0.015)	(0.021)	(0.015)	(0.021)	(0.026)	(0.020)
Ln(OSRD rate) * 1(1940-1944)	0.032	0.019	0.063***	0.058***	0.024	0.066***
	(0.024)	(0.030)	(0.016)	(0.019)	(0.027)	(0.022)
Ln(OSRD rate) * 1(1945-1949)	0.050	0.041	0.077***	0.076***	0.017	0.060***
	(0.034)	(0.044)	(0.016)	(0.017)	(0.021)	(0.019)
Ln(OSRD rate) * 1(1950-1954)	0.080**	0.063	0.078***	0.069***	0.023	0.092***
	(0.040)	(0.049)	(0.024)	(0.023)	(0.016)	(0.027)
Ln(OSRD rate) * 1(1955-1959)	0.124***	0.127**	0.054*	0.046*	0.037**	0.100***
	(0.045)	(0.055)	(0.030)	(0.026)	(0.017)	(0.028)
Ln(OSRD rate) * 1(1960-1964)	0.159***	0.135**	0.107***	0.095***	0.015	0.088**
	(0.049)	(0.059)	(0.033)	(0.029)	(0.025)	(0.039)
Ln(OSRD rate) * 1(1965-1970)	0.188***	0.182***	0.101***	0.041	0.002	0.088**
	(0.054)	(0.062)	(0.033)	(0.034)	(0.016)	(0.040)
N	20403	17697	10828	10478	3197	9128
R^2	0.75	0.71	0.47	0.37	0.19	0.59
Y mean	1.89	1.79	0.67	0.61	0.21	0.66
County-cat FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

Notes: Table estimates the effect of the OSRD shock on the growth of firm patenting. Observations are at the county-category-year level, and outcome variables are measured in logs. Column (1) measures all firm patents; Column (2), patents invented by firms with prior inventions in the same county and technology area; Columns (3) to (5), firms with prior patents only in a different county and/or technology area; and Column (6), firms with no prior patents. All columns include county-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county-category in parentheses.

Table 4: All patents vs. government-funded patents (total and by agency)

	(1) All	(2) Non govt.	(3) Govt.	(4) DOD	(5) DOE	(6) NASA
Ln(OSRD rate) * 1(1935-1939)	0.007	-0.006	0.003			0.127
	(0.031)	(0.012)	(0.036)			(0.098)
Ln(OSRD rate) * 1(1940-1944)	0.089	-0.003	0.060	-0.052		0.288
	(0.058)	(0.018)	(0.056)	(0.062)		(0.234)
Ln(OSRD rate) * 1(1945-1949)	0.145**	0.026	0.144**	0.009		0.208
	(0.072)	(0.026)	(0.073)	(0.056)		(0.227)
Ln(OSRD rate) * 1(1950-1954)	0.131*	0.084***	0.145**	-0.064		0.270
	(0.073)	(0.030)	(0.071)	(0.043)		(0.232)
Ln(OSRD rate) * 1(1955-1959)	0.153*	0.133***	0.173**	-0.029		0.220
	(0.081)	(0.036)	(0.079)	(0.057)		(0.221)
Ln(OSRD rate) * 1(1960-1964)	0.181**	0.193***	0.197***	0.017	0.022	0.353
	(0.075)	(0.042)	(0.072)	(0.046)	(0.074)	(0.222)
Ln(OSRD rate) * 1(1965-1970)	0.212***	0.221***	0.219***	-0.004	0.087	0.334
	(0.072)	(0.046)	(0.066)	(0.036)	(0.080)	(0.222)
N	9571	22253	8057	1344	254	279
R^2	0.44	0.79	0.44	0.44	0.44	0.48
Y mean	0.73	2.01	0.68	0.44	0.42	0.51
County-cat FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

Notes: Table estimates the effect of the OSRD shock on the growth of government-funded patenting. Observations are at the county-category-year level, and outcome variables are measured in logs. Column (1) measures all government-funded patents; Column (2), non-government funded patents; and Columns (3) to (6), patents by agency: Department of Defense (DOD), Department of Energy (DOE), National Aeronautics and Space Administration (NASA), and Department of Agriculture (USDA). Columns are labeled with modern agencies, but the DOD and DOE categories include predecessor agencies before they were established in 1947 and 1977, respectively. All columns include county-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county-category in parentheses.

Table 5: Share of forward and backward citations to local patents

	Backward citations			Forward citations		
	(1) Same county and field	(2) Same county, diff. field	(3) Diff. county, same field	(4) Same county and field	(5) Same county, diff. field	(6) Diff. county, same field
Ln(OSRD rate) * 1(1935-1939)				-0.001	-0.000	0.004
				(0.001)	(0.000)	(0.004)
Ln(OSRD rate) * 1(1940-1944)				-0.001	-0.001	0.010**
				(0.001)	(0.000)	(0.004)
Ln(OSRD rate) * 1(1945-1949)				-0.000	0.001	0.014***
				(0.001)	(0.001)	(0.004)
Ln(OSRD rate) * 1(1950-1954)	0.002	-0.000	0.003	0.001	0.001*	0.016***
	(0.002)	(0.001)	(0.005)	(0.001)	(0.001)	(0.004)
Ln(OSRD rate) * 1(1955-1959)	0.002	0.001	0.000	0.002***	0.001***	0.013***
	(0.001)	(0.001)	(0.005)	(0.001)	(0.000)	(0.004)
Ln(OSRD rate) * 1(1960-1964)	0.005***	-0.000	0.011**	0.003***	0.002***	0.011***
	(0.002)	(0.001)	(0.005)	(0.001)	(0.000)	(0.004)
Ln(OSRD rate) * 1(1965-1970)	0.006***	0.002*	0.009*	0.005***	0.002***	0.006*
	(0.002)	(0.001)	(0.005)	(0.001)	(0.000)	(0.004)
N	13887	13887	13887	22723	22723	22723
R^2	0.29	0.17	0.17	0.23	0.13	0.29
Y mean	0.04	0.02	0.46	0.02	0.01	0.28
County-cat FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

Notes: Table estimates the effect of the OSRD shock on local vs. distant citation flows. Observations are at the county-category-year level. Outcome variable in Columns (1) to (3) is the share of backward citations made by patents filed in a given cell that are to prior patents in the same county and technology category (Column 1), the same county but different technology category (Column 2), and a different county but same technology category (Column 3). Outcome variable in Columns (4) to (6) is the analogue for forward citations accruing to patents in the given cell from future patents of each column type. The front-page patent citation record begins in February 1947. All columns include county-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county-category in parentheses.

Table 6: Concentration of cluster patenting (Assignee HHI and patent shares)

		Share held by top X assignees			
	(1) HHI	(2) Top 1	(3) Top 5	(4) Top 10	(5) Top 20
Ln(OSRD rate) * 1(1935-1939)	-0.003	-0.005	-0.006	-0.001	-0.005
	(0.005)	(0.006)	(0.004)	(0.004)	(0.004)
Ln(OSRD rate) * 1(1940-1944)	-0.009*	-0.010	-0.008	-0.003	-0.001
	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
Ln(OSRD rate) * 1(1945-1949)	-0.029***	-0.009	-0.003	0.003	0.004
	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)
Ln(OSRD rate) * 1(1950-1954)	-0.030***	-0.017**	-0.010	-0.007	-0.001
	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)
Ln(OSRD rate) * 1(1955-1959)	-0.041***	-0.027***	-0.008	-0.004	-0.000
	(0.006)	(0.008)	(0.007)	(0.006)	(0.006)
Ln(OSRD rate) * 1(1960-1964)	-0.058***	-0.039***	-0.006	-0.001	0.006
	(0.006)	(0.009)	(0.008)	(0.007)	(0.006)
Ln(OSRD rate) * 1(1965-1970)	-0.054***	-0.065***	-0.030***	-0.025***	-0.018***
	(0.006)	(0.009)	(0.008)	(0.006)	(0.005)
N	23051	23051	23051	23051	23051
R^2	0.63	0.51	0.46	0.33	0.20
Y mean	0.38	0.66	0.98	1.03	1.05
County-cat FEs	X	X	X	X	X
Year FEs	X	X	X	X	X

Notes: Table estimates the effect of the OSRD shock on the concentration of local patenting. Observations are at the county-category-year level. Column (1) measures an assignee Herfindahl index, and Columns (2) to (5) measure concentration ratios. All columns include county-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county-category in parentheses.

Table 7: Effects on 1980 county employment in high-tech manufacturing industries

	Extensive			Intensive		
	(1)	(2)	(3)	(4)	(5)	(6)
1(Any OSRD patents, 1941-1948)	0.876***	0.895***	0.901***			
	(0.225)	(0.166)	(0.165)			
IHS(OSRD rate, 1941-1948)				1.485*	1.045*	1.063*
				(0.812)	(0.580)	(0.579)
N	3791	3791	3791	2027	2027	2027
R^2	0.55	0.77	0.77	0.62	0.86	0.86
Y mean	4.35	4.35	4.35	4.04	4.04	4.04
County FEs	X	X	X	X	X	X
Industry FEs	X	X	X	X	X	X
IHS mfg. empl.		X	X		X	X
IHS all empl.			X			X

Notes: Table estimates the relationship between counties' postwar employment in select industries and OSRD patenting in classes which crosswalk to these industries. Observations are at the county-industry level, with the sample restricted to industries in the broader domain of "Electrical and Electronic Equipment and Supplies" (see text). Industrial employment measured from the 1980 U.S. County Business Patterns (CBP). The outcome in all columns is the inverse hyperbolic sine (IHS) of industry employment. All columns include county and industry fixed effects, and successive columns add controls for IHS manufacturing employment and IHS total employment. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county in parentheses.

Table 8: Effects on firm creation in high-tech manufacturing industries

	Extensive			Intensive		
	(1)	(2)	(3)	(4)	(5)	(6)
1(Any OSRD patents, 1941-1948) * Decade==1930s	0.067**	0.066**	0.067**			
	(0.033)	(0.033)	(0.033)			
1(Any OSRD patents, 1941-1948) * Decade==1940s	0.413***	0.412***	0.415***			
	(0.077)	(0.077)	(0.077)			
1(Any OSRD patents, 1941-1948) * Decade==1950s	0.642***	0.640***	0.642***			
	(0.094)	(0.093)	(0.093)			
1(Any OSRD patents, 1941-1948) * Decade==1960s	1.125***	1.127***	1.128***			
	(0.111)	(0.110)	(0.110)			
1(Any OSRD patents, 1941-1948) * Decade==1970s	1.425***	1.419***	1.421***			
	(0.137)	(0.135)	(0.135)			
IHS(OSRD rate, 1941-1948) * Decade==1930s				-0.021	-0.050	-0.072
				(0.039)	(0.039)	(0.048)
IHS(OSRD rate, 1941-1948) * Decade==1940s				0.180*	0.164	0.178
				(0.098)	(0.102)	(0.108)
IHS(OSRD rate, 1941-1948) * Decade==1950s				0.251*	0.228	0.240
				(0.147)	(0.143)	(0.148)
IHS(OSRD rate, 1941-1948) * Decade==1960s				0.762***	0.748***	0.740***
				(0.273)	(0.276)	(0.279)
IHS(OSRD rate, 1941-1948) * Decade==1970s				1.056***	1.024***	1.018***
				(0.378)	(0.363)	(0.360)
N	127584	127584	127584	14610	14610	14610
R^2	0.56	0.57	0.57	0.57	0.58	0.58
Y mean	0.05	0.05	0.05	0.29	0.29	0.29
County-Ind FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
IHS mfg. firms		X	X		X	X
IHS all firms			X			X

Notes: Table estimates the relationship between counties' firm creation in select industries and OSRD patenting in classes which crosswalk to these industries. Observations are at the county-industry-decade level, with the sample restricted to industries in the broader domain of "Electrical and Electronic Equipment and Supplies" (see text). Firm creation measured from the 1980 issue of the Dun & Bradstreet establishment listings, which reports founding year for all firms in its data. Sample is restricted to headquarters and single-branch establishments, and by construction conditioned on survival to 1980. The outcome in all columns is the inverse hyperbolic sine (IHS) of industry firm creation. All columns include county-industry fixed effects, and successive columns add controls for IHS manufacturing firms and IHS total firms in the given year. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county in parentheses.

Web Appendix

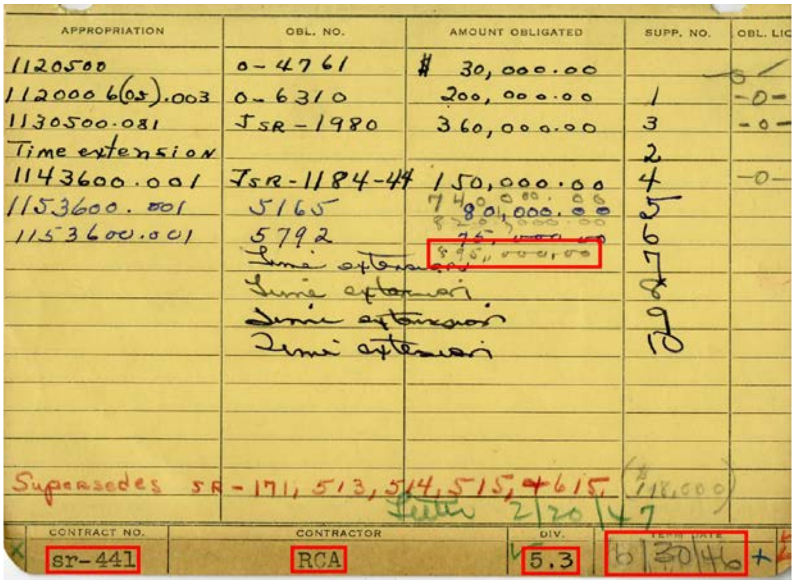
A OSRD History and Data

A.1 Data collected from OSRD archival records

Data on OSRD contractors and contracts, and OSRD-funded inventions, patents, and publications were obtained from archival records at the U.S. National Archives and Records Administration (NARA). Complete records from the OSRD have been preserved at NARA (Record Group 227), including narrative records (such as correspondence between Vannevar Bush and other OSRD administrators, contractors, and government agencies) that provide rich background on the OSRD's day-to-day operation and precedent-setting policy choices throughout the war.

For this paper, we make use of several sets of records from the OSRD collection. The three key data sources on contracts, contractors, and inventions are (i) a collection of contract index cards, which identify each contract with the contract ID, contractor, OSRD division managing the contract, termination date, and obligated value through the end of life; (ii) a directory of contractors, by contract, with the contractor's principal headquarters location; and (iii) a set of invention disclosure cards, which document inventions generated in the course of OSRD-supported research, with the associated contract, applicable patent clause (short form vs. long form), contractor, inventor(s), invention title, date reported, and if a patent application was filed, the application date and serial number. Examples of each are provided in Figures A.1 to A.3.

Figure A.1: Example Contract Index Card for OSRD Contract OEMsr-441



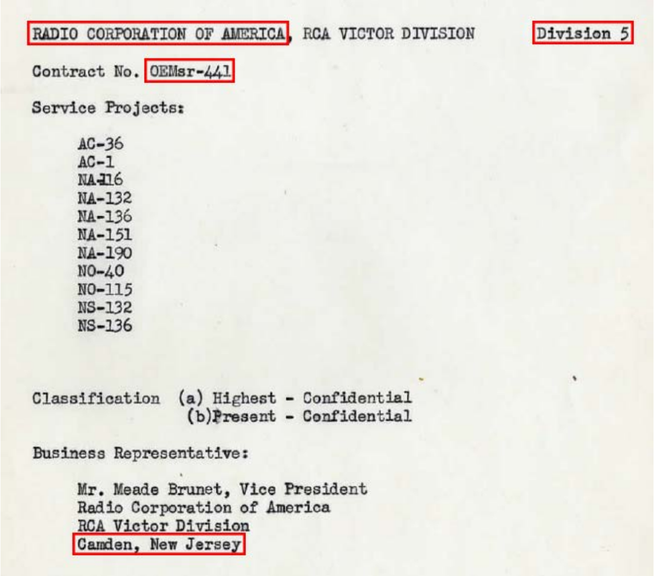
APPROPRIATION	OBL. NO.	AMOUNT OBLIGATED	SUPP. NO.	OBL. LIC
1120500	0-4761	$ 30,000.00		-0-
1120006(05).003	0-6310	200,000.00	1	-0-
1130500.081	JSR-1980	360,000.00	3	-0-
Time extension			2	
1143600.001	JSR-1184-44	150,000.00	4	-0-
1153600.001	5165	740,000.00 801,000.00 820,000.00	5	
1153600.001	5792	75,000.00	6	
	Time extension	895,000.00	7	
	Time extension		8	
	Time extension		9	
	Time extension		10	

Supersedes SR-171, 513, 514, 515, & 615 ($118,000)

Letter 2/20/47

CONTRACT NO.	CONTRACTOR	DIV.	TERM DATE
sr-441	RCA	5.3	6/30/46

Figure A.2: Example Contractor Directory Page for OEMsr-441



RADIO CORPORATION OF AMERICA, RCA VICTOR DIVISION Division 5

Contract No. OEMsr-441

Service Projects:

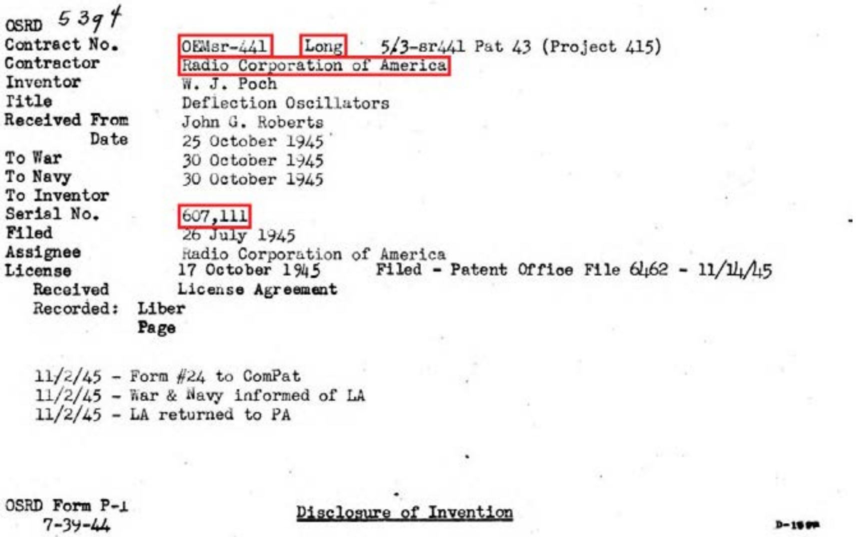
- AC-36
- AC-1
- NA-116
- NA-132
- NA-136
- NA-151
- NA-190
- NO-40
- NO-115
- NS-132
- NS-136

Classification (a) Highest - Confidential
(b)Present - Confidential

Business Representative:

Mr. Meade Brunet, Vice President
Radio Corporation of America
RCA Victor Division
Camden, New Jersey

Figure A.3: Example Invention Disclosure Card for OEMsr-441



OSRD 5394
Contract No. OEMsr-441 Long 5/3-sr441 Pat 43 (Project 415)
Contractor Radio Corporation of America
Inventor W. J. Poch
Title Deflection Oscillators
Received From John G. Roberts
Date 25 October 1945
To War 30 October 1945
To Navy 30 October 1945
To Inventor
Serial No. 607,111
Filed 26 July 1945
Assignee Radio Corporation of America
License 17 October 1945 Filed - Patent Office File 6462 - 11/14/45
Received License Agreement
Recorded: Liber
Page

11/2/45 - Form #24 to ComPat
11/2/45 - War & Navy informed of LA
11/2/45 - LA returned to PA

OSRD Form P-1
7-39-44

Disclosure of Invention

D-[illegible]

In addition to these records, we also collected three supplementary, independent lists of contractors and contracts which can be used to validate these data and fill in gaps. The first is a list of contractors, which lists contracts for each contractor provides the total obligations and counts the number of contracts with the short form vs. long form patent clause. The second is a list of contracts by contractor, and provides the associated OSRD division and patent clause. The third is a list of contracts by OSRD division. Using the collective set of records, we compile a master list of 2,288 OSRD contracts, 573 of which were CMR contracts.[1] For comparison, internal

---

[1] This list of contracts comprises the union of all record sets. Although the vast majority of contracts appear across all sources, there are gaps (missing contracts) in each data source: the contract index cards contain 2,275 of the

correspondence from the OSRD's patent division dated October 30, 1945 counts 2,266 research contracts, with the remaining 22 contracts appearing in our data as administrative contracts for the purchase/lease of office space, equipment rentals, and publication services. As another point of comparison, the administrative history of the CMR published in 1948 (Andrus 1948) lists 571 of the 573 CMR contracts in our data (of the remaining two, one was last contract the OSRD entered into and likely simply did not make it into this published list).[2]

Given that for a few contract-level variables we have multiple sources (namely, the OSRD division and patent clause), we make an additional effort to cross-validate these variables and reconcile differences as we describe below. We also harmonize contractor names and match them to assignees in the patent record (see discussion of patent datasets in the next subsection).

- The OSRD division of each contract indicates the subject matter and can be reported in up to four sources: the contract index cards, the contractor directory, and two supplemental contract lists. For 2,117 out of 2,288 contracts in the master list, at least one source reports an OSRD division and all reporting sources agree. In 152 cases, they disagree, and for these cases we take the most commonly-reported division as the true division (in the nine cases where there is no single most commonly-reported division, we prioritize the division reported in the contract index cards and contractor directory). In the 19 remaining cases, the contract's OSRD division is not reported in any of our sources and is unknown.
- The patent clause of each contract is typically either "short form" or "long form", which gives the government or the contractor (respectively) the first rights to patent inventions developed under the contract. The short form clause was used predominantly with academic contractors, and the long form clause with industrial contractors (as a means to incentivize their participation in the war effort). The patent clause of each contract is reported in one of the supplemental contract lists as well as on invention disclosure cards – but only known for 2,020 contracts in the supplemental contract list, and for an additional five contracts via the invention disclosures. Determining the applicable patent clause is further complicated by the fact that some contracts (i) had their patent clauses changed mid-contract, and/or (ii) had custom patent clauses (which were usually variants on the standard short form and long form clauses).[3] Custom patent clauses and patent clause amendments, as well as the

---

2,288 total known contracts; the contractor directory contains 2,259 contracts (the 573 CMR contracts were not included in this directory per se, but the contractor information was available in separately-maintained CMR contract ledgers); the supplementary contractor list contains 2,192 contracts; and the supplementary contract list contains 2,112 contracts. These latter two lists were compiled before the end of the war, and most of contracts missing from these lists are missing because they post-date the sample. For the 13 contracts missing from the contract index cards, the contractor and OSRD division are available from other sources, but the contract value and termination date are unknown. For the 29 contracts missing from the contractor directory, we infer contractor locations from other contracts with the same contractor where available, and otherwise through manual research.

[2] As a final piece of validation: the OSRD contract IDs were written with three different prefixes (NDCrc, OEMsr, OEMcmr) followed by an identification number. The maximum number of each series in the data is NDCrc-208, OEMsr-1507, and OEMcmr-573, and these numbers add to a total of 2,288.

[3] For example, all long form atomic energy research contracts (let under the OSRD's S-1 division) were converted to

> date of amendment, are noted in this list and thus measured in our data. In addition, several contracts are tagged as having a patent clause of “Purchase Contract” or ”Overhead” (these are typically administrative contracts, previously described), or even ”None”. The information on patent clauses from this contract list and in the invention disclosure cards overwhelmingly agrees, but where it conflicts, we use the information from the contract list, which appears to have been created specifically for this purpose.

After dropping administrative and cancelled contracts, there are 2,254 contracts in our data, made to 461 unique contractors, with a total value of $462 million in 1940-45 dollars (equivalent to roughly $7.4 billion today). For comparison, in the official administrative history of the OSRD, Stewart (1948) claims that OSRD contractors had spent “appoximately 457 million dollars through November 30, 1945” – a difference of less than one percent. We observe the contractor and OSRD division for every contract, the patent clause for 2,006 of the contracts (89%), and the obligated value and termination date for 2,208 of the contracts (98%).

We also observe the inventions produced under these contracts, which contractors were contractually required to disclose and the OSRD subsequently catalogued. As mentioned above, these index cards list the contract that the invention was developed under and also include information on patent applications, which we use to link OSRD contracts to granted patents – these patents are what we denote in the paper as “OSRD patents”. We discuss how we do so below.

Using these index cards alone, we identify 7,879 reported inventions. Each card has an identifying number in the top left corner, numbering from 1 to 8040 (e.g., see Figure A.3), with some also suffixed with letters. In all, the numbering suggests there may have been as many as 8,056 inventions, and the 177 (=8056-7879) missing index cards could either be unobserved inventions or numbers which were skipped or discarded. To fill in these potential gaps, we do a wider search for data on OSRD-funded inventions across other archival collections at NARA, and find additional lists of OSRD-supported inventions in the records of the U.S. Army Judge Advocate General’s Office (Record Group 153). Using these records, we recover data for five of the 177 missing index cards, and identify an additional 26 unnumbered inventions which were reported after the index card file was no longer maintained. For the inventions on which patent applications were filed, we use the serial number to link these to granted patents, as we explain in more detail in the next subsection. In all, we have 7,910 inventions, on which 3,382 patent applications were filed, and 2,659 patents granted (which we identify by linking serials to granted patents). 2,657 of these patents were granted by 1980, when our sample for this paper ends.

For completeness, we also search for continuations, divisions, and continuations-in-part of OSRD patent applications. To do so, we parse the first 800 characters of the text of patents filed between 1935 and 1969 (application series 2, 3, and 4) to identify patents which mention an OSRD serial, and manually check these cases to determine whether they were continued or divided from earlier

---

short form in 1942, when the research program was spun off into a weapons development program (the Manhattan Project), to ensure that the government controlled the intellectual property.

OSRD-supported filings. Through this process we find an additional 104 OSRD-supported patents, bringing our total to 2,763 patents (2,761 granted by 1980).

Our final source of data on OSRD-supported invention is a distinct list of patent applications and granted patents related to the Manhattan Project, obtained by FOIA (from the U.S. Department of Energy) by researchers at the Woodrow Wilson Center and available at its digital archives.[4] This document lists over 1,000 U.S. government-supported, nuclear energy-related patent applications (and >850 grants) from the World War II period, along with the contracts under which they were produced, including many OSRD contracts. Most of these inventions were placed into secrecy at the time of filing (see Gross 2019) and were not in the NARA records, likely due to their sensitive nature, and through this list we identify another 374 OSRD-supported patents (373 granted by 1980), bringing our total to 3,137 patents (3,134 granted by 1980).

A.2 Descriptive characterization of OSRD R&D investment

Administrative records from OSRD offer additional insight into the nature of the work it funded, and the firms and institutions it mobilized into the war effort. In prior work we have used these data to study how OSRD organized research for war and managed the World War II research effort (Gross and Sampat 2022d). Here we reproduce some of this evidence, while also adding more, to bring color to the specific features of the OSRD-led research effort.

Borrowing from Stewart (1948), Table A.1 shows OSRD's organizational structure in the form it eventually evolved into as the scope of its work grew. As we explain in the paper, OSRD was the parent agency of two organizations, NDRC (which managed the technological research effort—e.g. radar, atomic fission, and more) and CMR (which managed the medical research effort—e.g., penicillin, antimalarials). NDRC grew to have into 19 core divisions and seven special sections and panels. These units covered a wide range of subjects and varied equally widely in scale. The table shows total contract authorizations only for 1943 onwards, when NDRC organized into this structure. The two largest divisions were Radar (14) and Rocket Ordnance (3), with the majority of funding going to MIT and CalTech, respectively, to support major research labs. NDRC also directed the atomic fission research program until it was transferred to the Army in mid-1943, which operated at a similar scale to the radar program while under OSRD.

CMR, in contrast to NDRC, was more contained with six divisions of roughly equal size (in dollars), and as we describe in (Gross and Sampat 2022d), operated very differently. Despite having less than one-tenth the budget of NDRC, CMR was nevertheless similarly important to the war effort, as infectious disease and other ailments had in previous wars killed more soldiers than battlefield wounds, making military medicine a key R&D priority.

In the paper, we characterize the content of OSRD's work in technology space, across patent classes. Another way to visualize the focus of its research is with word clouds: in Figure A.4, we measure

---

[4] See https://digitalarchive.wilsoncenter.org/document/165247.

words in the title of NDRC patents and CMR publications and use them to produce (value-weighted) word clouds for the most common words in OSRD output. Among the more notable subjects of NDRC research were electrical communication (circuit, frequency, antenna, radio, wave), atomic fission (uranium), and rockets and missiles (rocket). Notable subjects of CMR research include blood and blood substitutes, shock, penicillin, and antimalarials.

Table A.2 illustrates the geographic concentration of OSRD research, listing the top 10 states with OSRD contracts (left panel) and OSRD patents (right panel). Massachusetts, California, New York, and New Jersey rank highly on both inputs and outputs, and together comprise around 65 to 75% of both contract obligations and patents. Table A.3 lists the top 10 university and firm contractors. The major university contractors hosted large, central laboratories which led specific research programs, such as the Radiation Laboratory at MIT for radar (Gross et al. 2022) and the Jet Propulsion Laboratory at CalTech for rockets and projectiles.

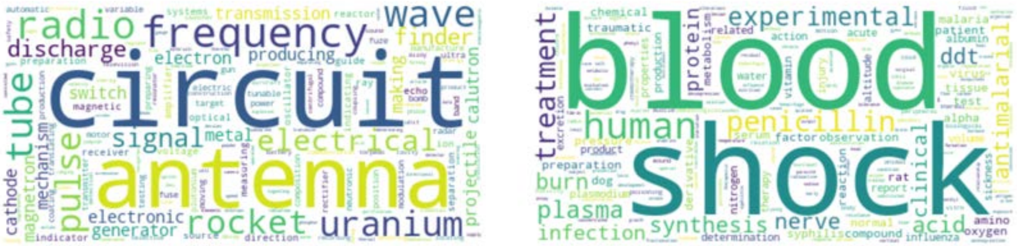
Table A.1: OSRD Divisions, Panels, and Special Sections (1941-1947)
[reproduced from Gross and Sampat (2022d)]

National Defense Research Committee (NDRC) Division/Section	Name/Description	Contract Authorizations ($, '000s) (1943-1947)	
1	Ballistics	5,327.2	
2	Effects of Impact and Explosion	2,701.4	
3	Rocket Ordnance	85,196.5	
4	Ordnance Accessories	20,014.3	
5	New Missiles	12,881.2	
6	Subsurface Warfare	33,883.5	
7	Fire Control	7,711.7	
8	Explosives	11079.9	
9	Chemistry	4,698.2	
10	Absorbents and Aerosols	3,524.2	
11	Chemical Engineering	9,216.2	
12	Transportation Development	2,199.4	
13	Electrical Communication	2,073.9	
14	Radar	104,533.4	
15	Radio Coordination	26,343.0	
16	Optics	5,923.9	
17	Physics	7,655.3	
18	War Metallurgy	3,794.4	
19	Miscellaneous Weapons	2,416.1	*
AMP	Advanced Mathematics Panel	2,522.9	
APP	Applied Psychology Panel	1,542.5	*
COP	Committee on Propagation	453.0	*
TD	Tropical Deterioration	232.4	*
SD	Sensory Devices	272.5	*
S-1	Atomic Fission	18,138.2	*
T	Proximity Fuzes	26,400.0	*
Total		400,735.1	

Committee on Medical Research (CMR) Division	Name/Description	Contract Authorizations ($, '000s) (1941-1947)
1	Medicine	3,873.3
2	Surgery	2,847.6
3	Aviation Medicine	2,466.5
4	Physiology	3,981.5
5	Chemistry	2,383.9
6	Malaria	5,501.9
–	Miscellaneous	3,635.3
Total		24,689.9

Notes: NDRC authorizations from January 1, 1943 onwards, except where noted below. CMR authorizations reported for the entire history of CMR.
*Authorizations for Division 19 from April 1, 1943; APP, from September 18, 1943; COP, from January 22, 1944; TD, from May 18, 1944; SD, from November 1, 1945. Authorizations for Sections S-1 and T are from June 27, 1940 onwards, with Section S-1 terminating in September 1943.

Figure A.4: Common words in OSRD patent and publication titles
[reproduced from Gross and Sampat (2022d)]


Panel (A): Words in patent titles

Panel (B): Words in publication titles

Notes: Figure illustrates the most common words appearing in the title of OSRD-supported patents and academic publications. Font size is proportional to number of occurrences, with larger words being more common. Patents primarily resulted from NDRC-supported technological R&D, and academic publications from CMR-supported medical research.

Table A.2: Top 10 states with OSRD contracts, by contract obligations
[reproduced from Gross and Sampat (2022d)]

Top states by share of OSRD contracts

		OSRD contracts, 1941-1948	
Rank	County	Value	Share of total
1	Massachusetts	$147.9 mil.	32.0%
2	California	$97.3	21.1%
3	New York	$91.3	19.8%
4	Illinois	$22.7	4.9%
5	District of Columbia	$17.1	3.7%
6	Pennsylvania	$14.7	3.2%
7	Maryland	$13.0	2.8%
8	New Jersey	$12.3	2.7%
9	Ohio	$8.6	1.9%
10	Michigan	$6.9	1.5%
	Total	$431.8	93.5%

Top states by share of OSRD patents

		OSRD patents, 1941-1948	
Rank	County	Value	Share of total
1	Massachusetts	597	21.4%
2	New York	516	18.5%
3	New Jersey	509	18.3%
4	California	247	8.9%
5	Illinois	159	5.7%
6	Pennsylvania	113	4.1%
7	Maryland	92	3.3%
8	Ohio	64	2.3%
9	District of Columbia	61	2.2%
10	Delaware	54	1.9%
	Total	2412	86.6%

Notes: Table lists the top states by total obligations (left panel) and patents (right panel). Percentages measure each state's percent of total OSRD obligations/patents.

Table A.3: Top OSRD contractors, by contract obligations
[reproduced from Gross and Sampat (2022d)]

Top 10 Universities

Contractor	Total oblg.	Pct. of total
Massachusetts Inst. of Tech.	$106.8 mil.	23.1%
California Inst. of Tech.	$76.6	16.6%
Harvard University	$29.1	6.3%
Columbia University	$27.1	5.9%
University of California	$14.6	3.2%
Johns Hopkins University	$10.8	2.3%
George Washington University	$6.9	1.5%
University of Chicago	$5.7	1.2%
Princeton University	$3.6	0.8%
University of Pennsylvania	$2.9	0.6%
Total	$284.0	61.5%

Top 10 Firms

Contractor	Total oblg.	Pct. of total
Western Electric Co.	$15.2 mil.	3.3%
General Electric Co.	$7.6	1.6%
Radio Corp. of America	$6.0	1.3%
E. I. Dupont De Nemours & Co.	$5.4	1.2%
Monsanto Chemical Co.	$4.5	1.0%
Eastman Kodak Co.	$4.3	0.9%
Zenith Radio Corp.	$4.2	0.9%
Westinghouse Elect. & Mfg. Co.	$3.9	0.8%
Remington Rand, Inc.	$3.7	0.8%
Sylvania Electric Products, Inc.	$3.1	0.7%
Total	$57.8	12.5%

Notes: Table lists the top firms and universities with OSRD contracts by total obligations. Percentages measure each contractor's percent of total OSRD research spending.

B Additional Data Sources

B.1 Construction of U.S. patent datasets

B.1.1 Base data

The construction of the patent datasets used in this paper begins with the USPTO historical master file (Marco et al. 2015), which provides a master list of utility patents with grant dates, patent class/subclass (USPC), and two-digit NBER category (Hall et al. 2001). In building this paper's dataset, we restrict the sample to patents granted between January 1, 1920 and December 31, 1979—although most of the paper invokes only a subset of these, emphasizing the sample filed between 1930 and 1970. For all granted patents in this set, we obtain additional patent characteristics from the following sources:

- FreePatentsOnline.com (FPO): serial numbers, filing dates, and the network of forward and backward citations (front-page citations only)
- Derwent Innovation database (DI): assignee names[5]
- Petralia et al. (2016a) HistPat, Berkes (2018) CUSP, and Bergeaud et al. (2022) PatentCity datasets: inventor country, state, county[6]
- USPTO Historical Government Register: administrative data on government-funded invention (owned or licensed), collected since 1944[7]

The DI assignee names are (mostly) standardized and were later found to match those in Google Patents data, which are freely available through Google BigQuery. These data are mostly complete, but a small number of patents are missing filing dates and assignees. Table B.1 shows the number patents with missing data, by decade of grant. For the period sampled in this paper (1930-1970), approximately 2.3% of patents are missing a filing date and 2.3% missing an assignee (note: these percentages are calculated for patents granted between 1930 and 1970, whereas the paper uses the sample of patents known to have been filed between 1930 and 1970).

---

[5] Note that serial numbers, filing dates, and the network of patent citations were also retrieved from the Derwent database for comparison against the FPO data, as a validation exercise. The two data sources overwhelmingly agreed, and where they disagreed, spot checks revealed that FPO was consistently the more accurate of the two, and when there was an error in the FPO data, it typically reflected the occasional typographical error on the printed patent publication itself, such as two flipped digits, or a digit one unit off the correct value. Given their reliability, the data for this paper thus use serial numbers, filing dates, and citations from FPO.

[6] Also see Petralia et al. (2016), as well as Andrews (2019) for additional discussion of historical patent geography data. We evaluate the quality of these data, and use all three sources to produce a composite measure of inventor locations, below. We are grateful to the authors for sharing their data.

[7] We introduce these data, and compare them to Fleming et al. (2019), below.

Table B.1: Number of patents with missing data, by decade

		No filing date		No assignee data	
Decade of grant	Patents	Number	Percent	Number	Percent
1920-1929	414901	25738	6.2%	25918	6.2%
1930-1939	442842	11102	2.5%	11221	2.5%
1940-1949	307630	5470	1.8%	5546	1.8%
1950-1959	425985	12461	2.9%	12661	3.0%
1960-1969	567761	11203	2.0%	11363	2.0%
1970-1979	689027	2	0.0%	73	0.0%
Total	2848146	65976	2.3%	66782	2.3%

Notes: Table shows counts of patents with missing data, and their fraction of all patents, by decade.

Patented, OSRD-funded inventions are identified in the OSRD archival records by the serial number of the patent application. It is thus critical to have accurate data on serial numbers. We manually reviewed and validated the application-level data (serials and filing dates) from FPO for the period around World War II by checking patents with serial numbers or filing dates which are out of sequence. The important feature of the USPTO's application numbering system for our purposes here is that applications are organized into application "series", which span several years, and identified by a serial number within that series, generally issued in the order in which patent applications arrive at the USPTO, with serial numbers never exceeding six digits. Application series increment, and serial numbers reset, at the beginning of a year in which the serial numbers from the previous series are expected to surpass 1,000,000. Series 2 begins January 1, 1935 and ends December, 1947 and is the focus of this data cleaning effort. We take all patents identified by FPO as belonging to Series 2 and sort these patents by serial. We then look for patents where the previous and next serial have the same filing date but the given patent has a different filing date, and then manually validate the serial and filing date for these patents. Out of over 370,000 patents in Series 2, corrections were made to 279 serials and 188 filing dates. Although these corrections are valuable for matching patents to serials in OSRD records, the low error rate for this sample also indicates that such errors are not widespread in the data.

B.1.2 Harmonizing assignee names

Although the assignee names from DI are largely already standardized, closer examination reveals that there are still variants on individual assignee names (e.g., BELL TELEPHONE LABOR INC with >10,000 patents, and BELL TELPHONE LAB INC, BELL TEL PHONE LAB INC, and BELL TEIEPHONE LAB INC with 1 patent each). We undertake several procedures to further harmonize assignee names. We begin by sorting a list of assignees in alphabetical order, and for each assignee recording other nearby assignees up to 9 positions before/after in the sorted list. We then calculate the edit distance between the given assignee name and each of these nearby assignee names. When this edit distance is less than 25% of the length of the longer name in each pair, We flag that pair as a candidate for manual review. We then review all such matches for several categories of assignees, and standardize names when a match is found:

- Assignees with $\geq$15 patents between 1930 and 1960
- Assignees which were OSRD contractors
- Assignees identified as government agencies (see next section)
- Assignees identified as universities or hospitals (see next section)
- Assignees which were synthetic rubber manufacturers
- Assignees which were spinouts from Standard Oil

This process is repeated (because each round of harmonization may bring new assignees into the set with $\geq$15 patents between 1930 and 1960) until no new matches are found.

This harmonization is neither perfect nor exhaustive, but it is believed to be effective for the purposes of this paper. It is also worth noting that for the vast majority of assignee names which were standardized by this procedure, there was clearly a primary spelling for that assignee in the original DI data, with hundreds or thousands of associated patents in the case of large assignees, and at worst a handful of secondary spellings with one or two associated patents—such that the actual effects of both (i) performing this harmonization for the priority assignees above, and of (ii) not performing it for non-priority assignees, are likely minimal.

B.1.3 Determining assignee types

Assignees are then classified into four categories—firms, universities and hospitals, government agencies, and individuals—through a combination of rule-based and manual classification. We begin by classifying assignees as firms when the assignee name includes any of roughly 120 words which indicate firms (e.g., CO, CORP, INC, LTD, SPA, GMBH, etc., as well as technical words such as AERO, AUTO, CHEM, ENG, MACHINE, OIL, PROD, TECH, WORKS; full list available on request). We then manually classify remaining assignees with $\geq$15 patents between 1930 and 1960, as well as assignees whose name includes any of the following strings:

- COLLEGE, INST, UNIV, HOSP, RES FOUND
- US, CANADA, UK, FRANCE, GERMANY, SWITZERLAND, AUSTRALIA, JAPAN, ISRAEL, and assorted other countries
- ATOM (to identify international atomic energy commissions)

Assignees with >200 patents in the 1920-1979 period which are thus far unclassified are then classified as firms. Any remaining unclassified assignees are classified as individuals.

This classification procedure was developed over several years, and although—like the name harmonization—it is neither perfect nor exhaustive, random spot checks suggest it is overwhelmingly effective at categorizing assignees into the right bins. In total, 60.1% of patents with an assignee in the 1920-1979 sample are assigned to a firm, 0.2% to a university, 0.8% to a government agency, and 39.1% to an individual (numbers sum to >100% because 5% of patents have multiple assignees, and 0.2% have assignees in multiple categories). Using administrative data, we will see below that the fraction we measure through names as assigned to a government entity is an undercount, primarily because the DI data sometimes undermeasure patent assignment.

B.1.4 Patent geography data

Measuring historical patents' inventor locations with completeness and accuracy is critical to our analysis. For modern, post-1976 patents, the USPTO provides inventor locations as part of the electronic record, but for pre-1976 patents, locations are only available from the text of the patent itself, presenting a formidable measurement challenge. Over the past few years, several researchers have invested in harvesting inventor and assignee locations from patent text for the universe of historical U.S. patents, typically rely on a mix of OCR with rule-based regex string parsing, machine learning, and manual correction (see Andrews 2019, for a summary). The most recent such effort has been produced by Bergeaud and Cyril (2022), which expands on prior work by measuring the location of historical patents issued by European patent authorities.

Because efforts to measure inventor locations from patent text are potentially susceptible to significant errors, we combine three sources—the Petralia et al. (2016a) HistPat, Berkes (2018) CUSP, and Bergeaud and Cyril (2022) PatentCity datasets—to develop and validate a composite measure of first inventor location in which we have high confidence, with the inventor's country, and if in the U.S., state and county. We begin by cleaning and harmonizing the inventor location data in each data source (country codes, state abbreviations, and county names). In some cases, FIPS codes are provided, but where needed, we assign FIPS codes to counties on state and name. We then restrict to first-listed inventors, calling their location the location of invention (a standard practice in research using patent data). We move forward with these data.

Our next step is to compare data sources, to evaluate how complete they are and how often they differ versus agree. Table B.2 provides summary statistics, including the rate at which they locate the patents in our sample, the rate at which they co-occur with and agree with other data sources. The simple takeaway is that CUSP and PatentCity are much more complete than HistPat, but any two sources agree on first inventor location for only 80% of patents.

Table B.2: Completeness and agreement of patent location data sources

	Overall	When HistPat located		When CUSP Located		When PatentCity located	
	Located	Located	Matches	Located	Matches	Located	Matches
Source	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HistPat	0.85	1.00	1.00	0.84	0.81	0.85	0.83
CUSP	0.93	0.93	0.81	1.00	1.00	0.94	0.85
PatentCity	0.94	0.94	0.83	0.94	0.85	1.00	1.00

Notes: Table summarizes the geolocation rates of patents by data source, and the pairwise rate of agreement across sources. For each row, the table shows the fraction of patents in our 1920-1979 sample for which it provides a first inventor location (Column 1). It then shows how often a location is provided when the column-wise data source also provides a location, and when so, the fraction of patents for which they agree (Columns 2-3, 4-5, and 6-7). Table illustrates the degree to which the data sources differ and we aim to reconcile.

These differences motivate our effort to aggregate these location measures to extract signal from the noise. To support this effort, we manually collected two ground-truth validation samples (patents for which we hand-coded their locations), against which we could benchmark the performance of different approaches to aggregation. One validation sample was generated from the USPTO's Index of Patents, an annual publication that lists all patents issued in a given year, along with inventor and assignee information. We selected two pages at random from each of the 1930, 1940, and 1950 editions and transcribed all patents therein, with approximately 350 patents in total.[8] In a second validation sample, we select 200 patents at random from each of the years 1930, 1940, 1950, 1960, and 1970 and hire workers on Mechanical Turk to extract inventor and assignee locations, which we then review for consistency. Through iterative comparisons to these validation samples, we ultimately arrived at an approach where we assign each patent in our sample a location as follows. Where a majority of our data sources agree, we assign a patent the consensus or majority location. This rule accounts for 90.6% of patents. When multiple sources provide a location but our sources disagree, we apply the following rules, in sequence:

1. If the PatentCity relevance score == 1, use PatentCity Location. Else:
2. If the HistPat accuracy score == 'High', use HistPat location. Else:
3. If the HistPat accuracy score == 'Medium', use HistPat location. else:
4. If CUSP has a county, use CUSP location. Else:
5. If PatentCity has a county, use PatentCity location. Else:
6. If CUSP has a state, use CUSP location. Else:
7. If PatentCity has a state, use PatentCity location. Else:
8. If CUSP has a country, use CUSP location. Else:
9. If PatentCity has a country, use PatentCity location. Else:
10. Leave as missing.

The end result is the assignment of a location to 99.9% of patents in our sample, whose data sources break down as shown in Table B.3. When compared against our validation samples, we find that 98% to 99% of patents where all three sources agree on a location have the correct location, and 90% to 93% of those where two of three sources agree on a location have the correct location. However, when our sources disgree, or only one sources provides a location, it is correct only 45% to 50%

---

[8] The 1960 edition did not provide inventor locations.

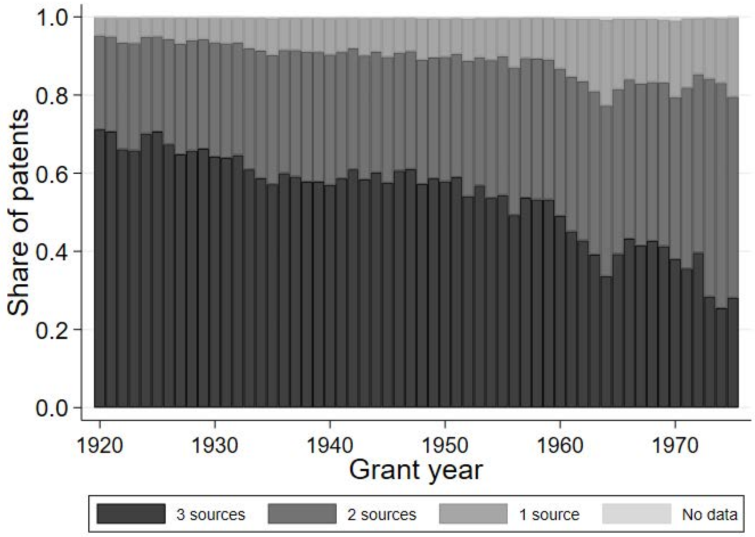
of the time. To limit potential mismeasurement, throughout the paper we restrict our analytical sample to patents where at least two sources agree on location. In Figure B.1 we show the share of patents with 3, 2, 1, and 0 data sources, by year, where we can see that between 80% and 95% of patents each year have agreement among 2+ location data sources.

Table B.3: Summary of final patent location measure

Source	Count	Percent	Cumulative
All sources (agreement)	1,600,150	56.2	56.2
2 of 3 sources (majority)	980,271	34.4	90.6
Single source (conflict)	265,040	9.3	99.9
No sources (no data)	2,685	0.1	100.0

Notes: Table summarizes the location data sources of all patents in our sample.

Figure B.1: Consensus across sources of patent geography data



Notes: Figure shows the annual fraction of granted U.S. patents from 1920 to 1975 for which we have a first inventor location where we have agreement (i) across all three data sources, (ii) across two sources, (i) no agreement across sources, and (iv) no data. Post-1976, USPTO began maintaining electronic records, such that inventor locations are broadly well-measured for this period.

In a final test, we also compare our data to administrative patent counts at the state-year level from the USPTO's 1977 Technology Assessment & Forecast report.[9] Though accurately measuring counties is most crucial for our purposes in this paper, this comparison can be insightful as to whether our measurement is broadly on track. Figure B.2 plots state-year totals from each of our data sources (including our composite measure) against the USPTO state totals for the 1930 to 1960 period. All series are broadly consistent with each other and with administrative totals, but specific series seem to face measurement challenges in specific states (e.g., HistPat in WV, CUSP in FL, and PatentCity in RI, though it appears the USPTO data also have irregularities in RI). Our composite measure minimizes the RMSE against USPTO totals, being roughly 35% lower than that of PatentCity, 40% lower than HistPat, and 85% lower than CUSP.

---

[9] We thank Seamans et al. (2018) for pointing us to this source and sharing their data.

Figure B.2: State-year patent totals, HistPat vs. Administrative data, 1930-1960

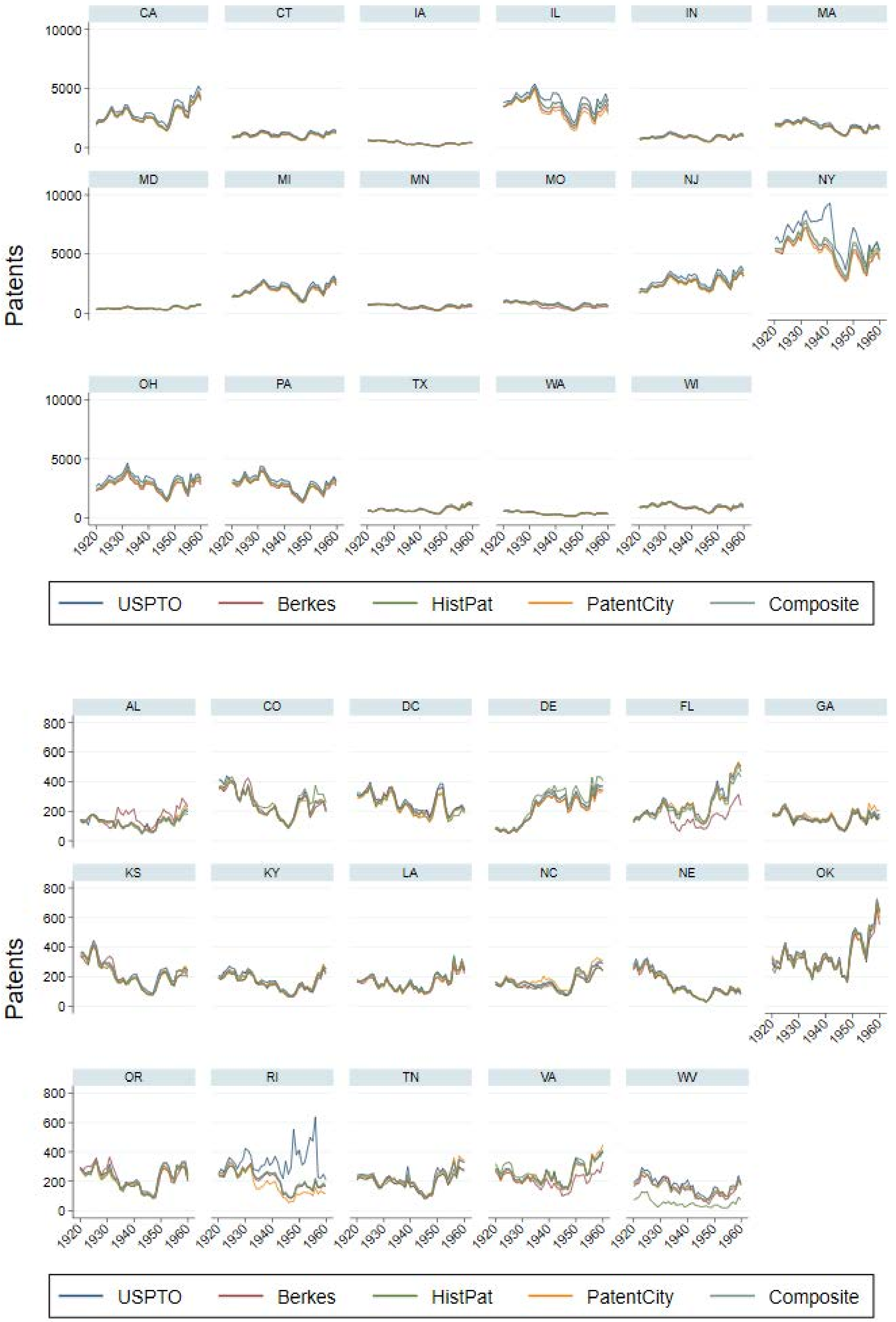
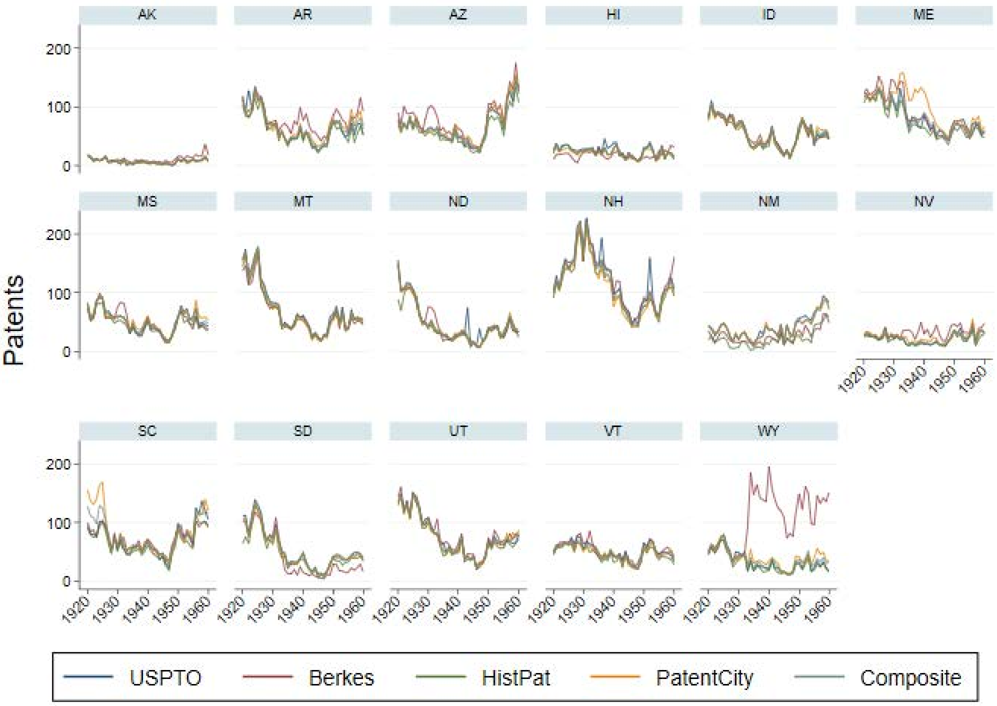


Figure B.2: State-year patent totals, HistPat vs. Administrative data, 1930-1960



Notes: Figure compares USPTO administrative totals of patent counts by state and year to those in each of our patent geography data sources (Petralia et al. 2016, Berkes 2018, Bergeaud and Cyril 2022), and to the composite measure we construct for this paper. Sample divided into the top third of states by average annual patents over the sampled period, middle third, and bottom third for ease of viewing.

Fixing county definitions

A final concern for location measurement is the fact that county definitions change over time, due to counties being combined, divided, or simply shifting borders—though most of these changes occur before 1940 (with the exception of independent cities in Virginia being carved out of, or merged into, their surrounding counties across most of the postwar period). To fix county boundaries across our sample period, we group up counties into the smallest stable units. We use two resources for tracking county boundary changes: the Atlas of Historical County Boundaries maintained by the Newberry Library, which provides a comprehensive list of county boundary changes for every state over the entire history of the U.S., and the U.S. Census Bureau's list of substantial changes to county boundaries since 1970. Note that we apply these changes to other county-level data sources as well, to ensure geographic consistency across all of our analysis.

B.2 Data on government-supported patents

A first-order question we wrestle with in the paper is to what degree the effects of World War II R&D we observe are driven by continued public investment in the same locations and technologies as those funded in the war, and (or) whether they occur organically, driven by Marshallian increasing returns to scale. To answer this question, we not only want to relate wartime government-funded R&D to postwar R&D, but also to compare places and technologies which were shocked by World War II and received more versus less postwar public research investment. Doing so requires comprehensive data on U.S. government (henceforth, USG) funded invention.

Recent efforts to comprehensively measure USG-supported patents have done so by algorithmically reading patent text for patent assignments to government entities and for government interest statements or other in-text mentions of government support (Fleming et al. 2019). With considerable effort, albeit also with the imprecisions of algorithmic measurement in a very large corpus, this work captures the set of patents where government interests can be discerned from the patent document. But this approach also bears important limitations that could result in a significant undercount, especially in the historical period, and present problems for our exercise. The first issue is that a large number of government-supported, patented inventions are produced by contractors and grantees in the course of supported work, who sometimes retain title to these inventions with the funding agency holding a paid-up license for government use. The Department of Defense, for example, has historically been a major funder of applied research and invention but has not taken title to patents, instead letting contractors have title while retaining a royalty-free license.[10] Though many such patents write government interest statements into their text, these are systematically required by agencies only after the Bayh-Dole Act of 1980, and even then, compliance and enforcement have been haphazard (Rai and Sampat 2012). Additionally, many patents by government employee-inventors have historically not been assigned nor acknowledge government employment, and are thus undetectable by reading and processing patent text. The text-based approach will thus tend to undermeasure USG-supported invention.[11]

In light of these measurement challenges, we chose to pursue administrative data on government-funded invention. Since 1944, USPTO has maintained a "Register of Government Interest in Patents", pursuant to Executive Order 9424 (18 Feb. 1944), which contains information on patents resulting from government grants and contracts. This Register has been used in research before: Watson and Holman (1964) study the so-called "Government Register" to report on the U.S. government patent portfolio, describing an index card series on which government interests were recorded. To our knowledge, it has not been revisited since.

---

[10] Watson and Holman (1964) note that individual agencies are usually either "title-policy" or "license-policy", and document that at least through the mid-1960s, more than 2/3 of federally-funded patents were licensed, not owned, by the agencies that funded the underlying R&D.

[11] As a more pedantic matter, as is we are able to identify roughly 1.5% more government-assigned patents in the (Fleming et al. 2019) data than the authors identify, using the assignee data in their replication package, though this difference is too small to meaningfully alter our work.

We located these records at the U.S. National Archives (NARA) and arranged to have them scanned and transcribed. These records comprised 127,852 index cards, of which examples are shown in Figure B.3. The cards provide several pieces of information: identifying information (patent number and issue date, serial number and filing date, title), the assignor (i.e., contractor) and inventor, the sponsoring government agency (e.g., Army, Navy, Air Force, AEC, NASA), and the government interest (title vs. license, as well as Act of 1883 or U.S.C. 266, which are legal statutes which assign the government rights to employee inventions). Though the collection of these data began in 1944 under the E.O., the records include patents filed and/or issued as early as the 1890s, and as late as the early 1990s, when USPTO completed its transition to electronic records.[12]

We undertake several steps to clean and regularize these data, including: hand-checking the values of numeric fields with non-numeric characters, correcting errors in transcription as well as on the original cards; 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, NSA all become DOD; AEC becomes DOE; HEW, PHS, NIH all become HHS; etc.). The Government Register also identifies a number of OSRD patents, as the example in Figure B.3 illustrates, albeit far fewer than the OSRD records do, as many of the OSRD-funded inventions end up in these records marked as assigned or licensed to the armed services. We thus rely on OSRD's official, archival records to measure OSRD-funded invention, and on the Government Register to measure USG-funded invention in the wider population of patents.

---

[12] Analogous, modern data are recorded in the USPTO's Patent Assignment Dataset, where all assignments of interests (title, license, and more) are recorded. See Marco et al. (2015b). Conveyances of title and license can also be searched online at https://assignment.uspto.gov/patent/index.html#/patent/search.

Figure B.3: Example Government Register index cards

Pat. 2,416,718 March 4, 1947. Dept. Army
App. 460,328 Oct. 1, 1942. Bur.
Title: Pulse Generator
Assignor: Bell Telephone Laboratories, Incorporated
Inventor: Shockley, William
Liber:
File: Dept.14112 & Pub. 5927
Remarks:
Assignment
License X
Act of 1883
PO-283(b) (6-50) PAT'D ☐
Department of Commerce U. S. Patent Office

Pat. 2,780,595 Feb. 5, 1957 Dept. ~~O.S.R.D.~~ AEC
App. 534,129 May 4, 1944 Bur.
Title: Nuclear Chain Reacting Systems
Assignor: Fermi, Enrico
Inventor: Same
Liber: U.S.12, p. 643;
File: Reel 131, Frame 082
Remarks: Outright
Assignment X
License
Act of 1883
PO-283(b) (6-50) PAT'D ☐
Department of Commerce U. S. Patent Office

Pat. 2,540,654 Feb. 6, 1951. Dept. Navy
App. 16,998 Mar.25, 1948. Bur.
Title: Data Storage System.
Assignor: Engineering Research Associates, Inc.
Inventor: Cohen, Arnold A., Keye, William R. & Tompkins, Charles B.
Liber:
File: Pub.5209
Remarks:
Assignment
License x
Act of 1883
PO-283(b) (6-50) PAT'D ☐
Department of Commerce U. S. Patent Office

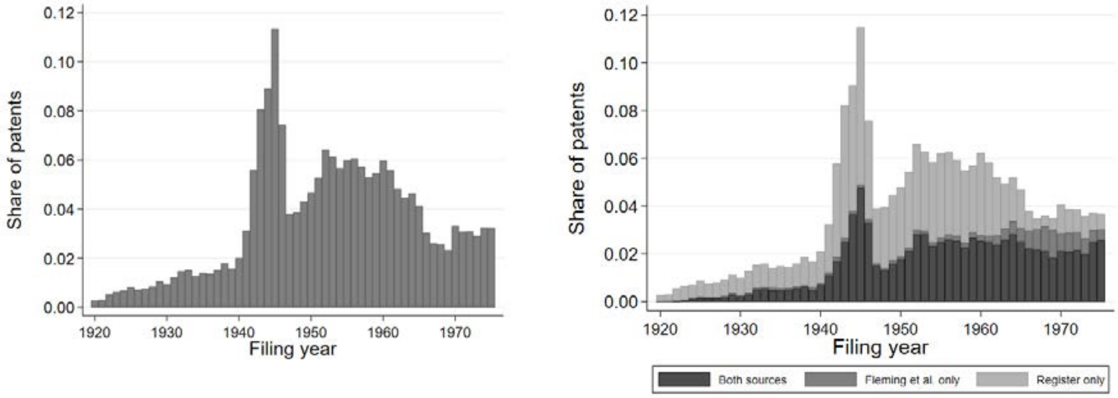
Pat. 2,744,042 - May 1, 1956 Dept. Air Force
App. 232,691 - June 21, 1951 Bur.
Title: Laminated Panels
Assignor: Goodyear Aircraft Corporation
Inventor: Pace, Henry A.
Liber:
File: Dept. 15265
Remarks:
Assignment
License X
Act of 1883
PO-283(b) (6-50) PAT'D ☐
Department of Commerce U. S. Patent Office

Pat. 3309626 - Mar. 14, 1967 Dept. Army
S.N. 415,836-Dec. 3, 1964
Title: Microwave Limiter
Assignor: Higgins, Vincent J.
Inventor: Same
Reel: 1484 F. 679
File:
Remarks: Outright
Assignment: X
License:
U.S.C. 266 x
FORM PO-283(b) (11-63)
U.S. DEPARTMENT OF COMMERCE PATENT OFFICE

Pat. 4,267,953 May 19, 1981 Dept. N.A.S.A.
S.N. 122,967 FEB. 20, 1980
Title: METHOD for ALLEVIATING THERMAL STRESS DAMAGE in LAMINATES
Assignor: CHARLES A. HOFFMAN, JOHN W. WEETON, and NORMAN W. ORTH
Inventor: SAME
Reel:
File: 3739 F. 927
Remarks: Outright
Assignment: x
License:
U.S.C. 266
PTO-283
U. S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

In the figures below we illustrate what these data have to offer. Using the Register data, in Figure B.4, Panel (A) we first measure and plot the government-funded share of all U.S. patents by filing year, from 1920 to 1975, where we can see the USG share reaching its peak in World War II (note: when the series is extended into the 2000s, using modern data, we find that this share continues dropping). In Panel (B) we compare our data against those of Fleming et al. (2019), measuring what fraction of patents are in one, the other, or both sources. For most of the period we study, the Government Register data more than doubles the known number of government interest patents, with the Fleming et al. (2019) set fully subsumed by the Register data—though in the late 1960s and early 1970s, some patents begin to appear in Fleming et al. (2019) which are not present in the Register, potentially because it is less complete for these years.

Figure B.4: Government-funded share of U.S. patent applications, 1920-1975


Panel (A):
Government Register (administrative)

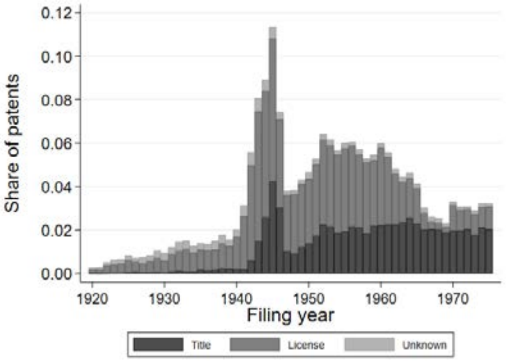
Panel (B):
Comparison to Fleming et al. (2019)

Notes: Figure shows the government-funded fraction of U.S. patents from 1920 to 1975 (Panel A) and compares the annual government share of patents in our newly-collected administrative data to Fleming et al. (2019) (Panel B).

Figure B.5 examines the share of Register patents over time where the government has title, license, or an ambiguous interest, first overall (Panel A) and then for several of the larger R&D-funding agencies (Panels B to E). Here we see that license policy patents make up the vast majority of government-funded invention until the late 1960s, explaining the gap between the Register and Fleming et al. (2019) counts. We can also see that DOD is by far the driving force behind USG-supported invention: the next-largest agencies are only a small fraction of its size. In Figure B.6, we plot the share of Register patents identified as government assigned or supported (acknowledging), overall (Panel A) and by agency (Panels B to E). This share grows over time at DOD, fluctuates between 50% and 80% at DOE and NASA for most of the period, and is consistently near 100% at USDA. The completeness of the B.6 is largely driven by the degree to which each agency applies a title policy versus a license policy, and how this changed over time.

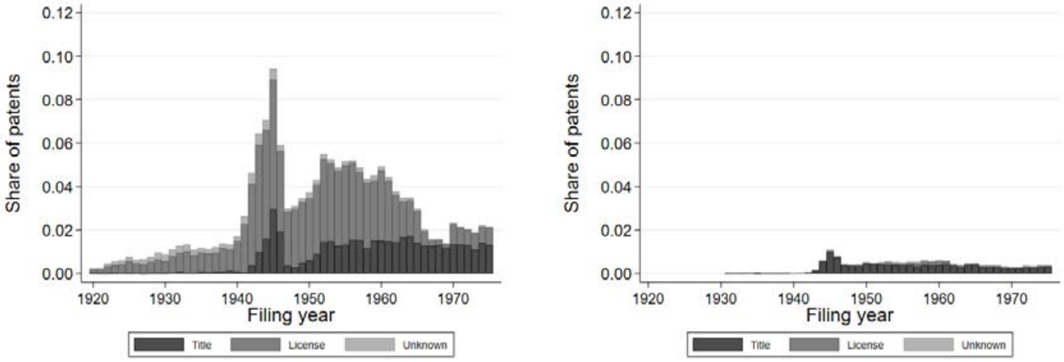
Figure B.5: Title vs. license policy share of U.S. patent applications, by agency, 1920-1975

Panel (A): All agencies



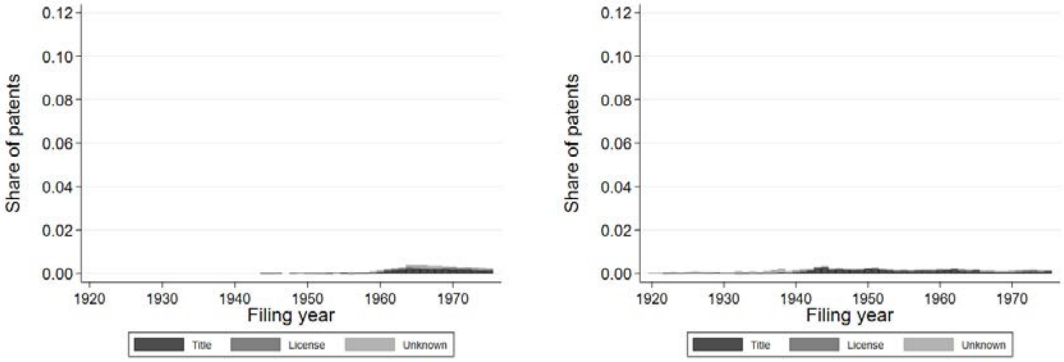
Panel (B): DOD

Panel (C): DOE



Panel (D): NASA

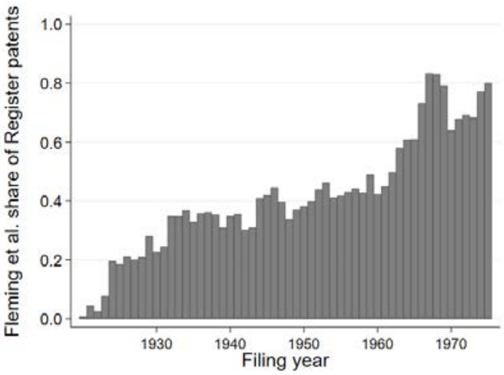
Panel (E): USDA



Notes: Figure shows the government-funded fraction of U.S. patents from 1920 to 1975, overall (Panel A) and for the four highest patenting agencies (Panels B to E). Each subfigure shows the fraction of patents in our data that are title policy (i.e., government retains title) versus license policy (inventor allowed title, government retains paid-up license). License policy patents comprise two-thirds of government-funded invention but prior to these data could not be systematically measured.

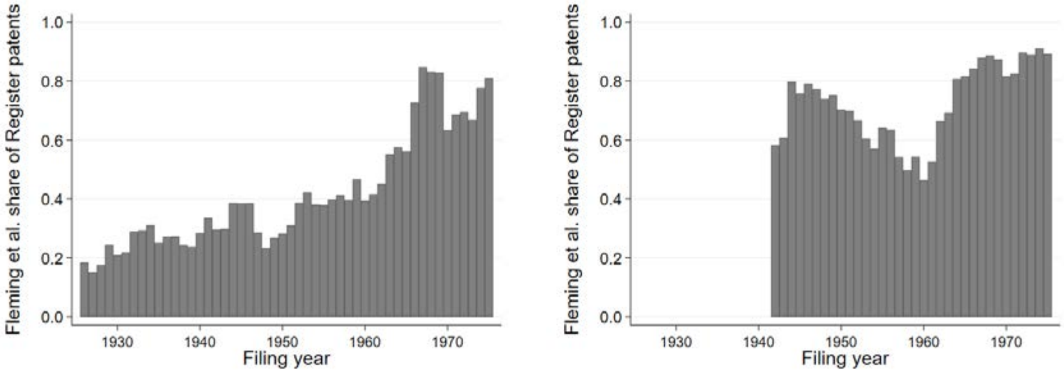
Figure B.6: Share of patents captured by Fleming et al. (2019), by agency, 1920-1975

Panel (A): All agencies



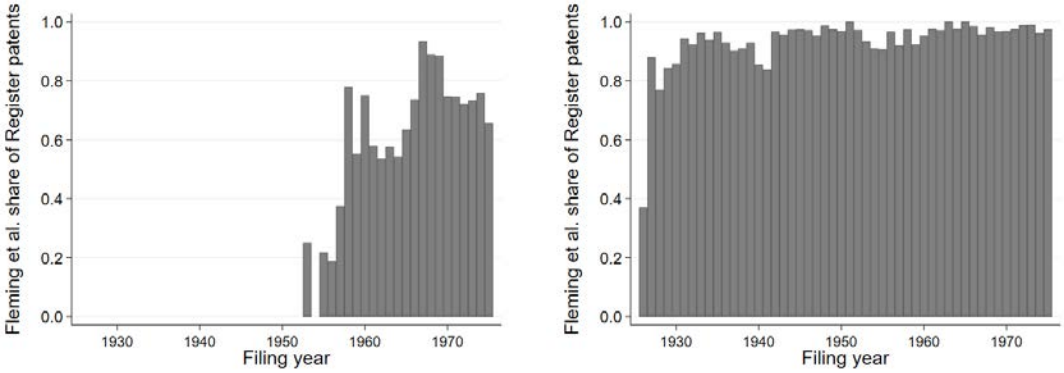
Panel (B): DOD

Panel (C): DOE



Panel (D): NASA

Panel (E): USDA



Notes: Figure shows the share of government-funded patents from 1920 to 1975 which appear in Fleming et al. (2019), overall (Panel A) and by agency (Panels B to E). These shares correlate with each agency's use of a title patent policy, both in the cross-section and over time, since title patents are assigned to the U.S. government and easily identified in modern data. DOD was historically a license policy agency but increasingly used title policy over time, while USDA has always been a title policy agency.

B.3 Construction of foreign patent dataset

We complement the U.S. patent data with foreign authority patent data. Specifically, we collect data on all utility patents granted by the U.S., British, and French patent authorities between 1920 and 1979 from the EPO PATSTAT database. Much like the U.S. historical master file, these data include grant dates and patent class (IPC), which are the main information used in this paper, as well as filing date, number of inventors and original assignees, the cross-jurisdiction patent family size, and the patent family's forward citations. In preparing these data, we restrict to patents from each patent authority (US, GB, FR) of inventions (as opposed to utility models in some jurisdictions, or design patents), and we restrict the U.S. patents in this dataset to those which are also in the USPTO dataset which we constructed and described above.

B.4 Other datasets

B.4.1 Historical U.S. County Business Patterns data

In the paper, we use several years of U.S. County Business Patterns (CBP) data to go beyond the patent record and study the effects of the OSRD on the local development and agglomeration of high-tech industries. The CBP measures establishments, employment, and wages at the level of counties and industries, and was published at irregular intervals from 1946 to 1964 and annually thereafter. PDFs of complete printed volumes are available for most years from HathiTrust, and electronic CBP data are available from NARA for 1970 onwards.[13]

For this paper, we collected data from four editions of the CBP at roughly 10-year intervals: 1947, 1959, 1970, and 1980. These editions measure industrial activity within counties down to the level of 4-digit SIC industries, along with 2- and 3-digit level totals (excepting the 1947 edition, which reaches only the 3-digit level, and only for select industries). The 1970 and 1980 CBP editions are available in electronic format, and 1956 is available from Gilles Duranton (Duranton et al. 2014), though the data turned out to be too early, and too aggregated, for our purposes: too early because SIC 367 ("Electronic Components and Accessories", a key industry in relation to the OSRD shock) was only created by the 1957 SIC classification, and too aggregated because it has few 3- and 4-digit SICs, which is the level at which more precise links between technology areas and industries are drawn. We digitized subsets of the 1959 CBP as well, in particular to support deep dives into SIC 366 ("Communications Equipment) and SIC 367, but this too reports 3- and 4-digit SICs with insufficient frequency to use widely. This is the case even up to the 1970 CBP, which is in part why in the body of the paper we focus on 1980 alone, which reports lower-level SICs more widely. An additional reasons we prefer 1980 is its later date and its consistency with the USPTO crosswalk we

---

[13] More information on the Census' CBP data program is available at https://www.census.gov/programs-surveys/cbp/technical-documentation/methodology.html. See the NARA online catalog at https://catalog.archives.gov/search?q=*:*&f.parentNaId=613576&f.level=fileUnit&sort=naIdSort%20asc for 1970-2007 data files. All of the CBP editions used in this paper report mid-March employment and first quarter payroll.

use to connect industries and patent classes, both of which are based on the 1972 SIC classification. Though we collected data from multiple CBP years, in the course of this work we concluded that we lose little by invoking only one CBP edition, because the analysis is inherently cross-sectional due to the absence of pre-war data (CBP begins in 1947).

Working with CBP data over long horizons poses three challenges: (i) changes to the SIC classification over time, (ii) data suppression in small county-industry cells, and (iii) county boundary changes and the binning of small counties in select states.

1. *Changes to SIC classification.* Between 1947 and 1980, the SIC classification underwent multiple revisions, adding new industries, combining/dividing/reclassifying existing industries, and updating industry definitions to shift some types of businesses across industries. These changes primarily occur at the level of 4-digit industries. We used the 1945, 1957, 1967, and 1972 editions of the SIC Classification Manual (the latest editions preceding each of our sampling years) to build year-to-year crosswalks between all SIC industries on the 366 and 367 branches of the classification tree. Broader harmonization is beyond our scope, for now, though we note that at these more aggregated levels, definitions were stable: all changes to 4-digit industries occurred entirely within these branches.

2. *Data suppression.* Another challenge is data suppression: the CBP historically suppressed employment and wages in county-industries with very few establishments, to avoid disclosing data on individual establishments. This is particularly true for later editions (the 1947 edition does not appear to have suppressed data) and for 4-digit SICs.

   In these cases, we impute employment using data on the distribution of establishments across size bins, which is provided unsuppressed for all county-industries. In 1980, the data suppression flags also indicate not only that the value was suppressed, but also a range it falls in. Between the two, we have enough information to approximate the suppressed employment totals. To have a consistent approach across CBP editions, we focus on imputation from the establishment size distribution, using two procedures. We first assign each establishment in the size distribution an employment level equal to the midpoint of its bin, and add them together.[14] For each CBP year, we also regress total employment (where observed) on the establishment size distribution, and use the estimated parameters to predict employment for all county-industry cells. These two approaches yield distinct, internally-consistent, imputed estimates of employment. The correlation of these imputed values with each other and with reported values (where observed) is generally 0.8 or higher, although this correlation is in part driven by the skewed distribution of county size.

3. *County binning and changes.* A third potential challenge is the fact that the 1947 and 1959 CBP report county-level data for most counties, but group up low-density rural counties in

---

[14] These bins are as follows. For 1959 and 1970: 1-3, 4-7, 8-19, 20-49, 50-99, 100-249, 250-499, 500+. For 1980: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000-1499, 1500-2499, 2500-4999, 5000+. For firms in the upper, unbounded bins, we assign them an employment level equal to that lower bound.

eight states, typically into groups of 2-3 counties (the encoding limited each state to up to only 99 counties and county groups). In Virginia, independent cities were also binned with adjoining counties, and in one instance, a newly-created county was binned with its parent (Los Alamos and Sandoval, NM). The affected states are as follows:

	Number of counties			
	1947		1959	
State	Separated	Grouped	Separated	Grouped
Georgia	66	93	66	93
Illinois	96	6	96	6
Kansas	93	12	93	12
Kentucky	84	36	84	36
Missouri	85	30	85	30
New Mexico			30	2
North Carolina	98	2	98	2
Texas	39	215	39	215
Virginia	98	2	70	62

Because high-tech industries were concentrated in populated counties with large urban centers, the binning up of small rural counties in these states does not create a problem for this paper: the only binned county groups with electronics industry employment were a half dozen counties and independent cities in Virginia, which (i) are too few in number to affect our results, and (ii) we might wish to drop anyway, because of frequent land exchanges and ambiguous geographic definitions with Virginia's independent cities.

An additional challenge is county boundary changes. Although county boundaries in the lower 48 states are mostly fixed by the 1940s, boundaries do change for a few counties post-1947. As in the patent data and other county-level data, to fix county boundaries across our sample period we group up counties into the smallest stable units. For example: for counties (or independent cities) which were eliminated between 1947 and 2000, we combine them with the absorbing county in the CBP data. For the counties which were established between 1947 and 2000, we add them back to their parent counties.

Given our use of a USPTO crosswalk that maps SICs and patent classes to common industry codes, it is useful to describe in more detail what it is. First, we note that the crosswalk files are available here: https://www.uspto.gov/web/offices/ac/ido/oeip/taf/data/sic_conc/. This classification has a total of 41 product codes (which we call industry codes, in this paper, for convenience) that roll up to the following 15, not mutually exclusive categories:

- R1: Chemicals and Allied Products
- R2: Chemicals, except Drugs and Medicines
- R3: Basic Industrial Inorganic and Organic Chemistry

- R4: All Other Chemicals
- R5: Primary Metals
- R6: Machinery, except Electrical
- R7: Other Machinery, except Electrical
- R8: Electrical and Electronic Machinery, Equipment and Supplies
- R9: Electrical Equipment, except Communications Equipment
- R10: Other Electrical Machinery, Equipment and Supplies
- R11: Communications Equipment and Electronic Components
- R12: Transportation Equipment
- R13: Motor Vehicles and Other Transportation Equipment, except Aircraft
- R14: Other Transportation Equipment
- R15: All Industries

Our focus in Section 5 is on the "Electrical and Electronic Machinery, Equipment and Supplies" category, which includes the following industries:

USPTO Code	Description	Crosswalked SICs
35	Electrical transmission and distribution equipment	361, 3825
36	Electrical industrial apparatus	362
38	Household appliances	363
39	Electrical lighting and wiring equipment	364
40	Miscellaneous electrical machinery, equipment and supplies	369
42	Radio and television receiving equipment	365
43	Electronic components, accessories and communications equip.	366-367

Industry 43 (SICs 366 and 367) was the category with the most associated OSRD patents, and the second-highest OSRD rate (behind ordnance and explosives). In light of their importance, it is also useful to describe some of the industries and products within these 3-digit SICs. The table below provides the 4-digit subindustries and examples of products they manufacture, from the 1972 SIC classification manual. Where products are not listed, it is because the products approximately match the industry description.

- 366: Communications Equipment
  - 3661: Telephone and Telegraph Apparatus
  - 3662: Radio and TV Transmitting, Signaling, and Detection Apparatus
    - Aircraft control systems, Amplifiers, Antennas, Digital encoders, Electronic control systems, Inertial guidance systems, Laser systems, Linear accelerators, Microwave communication equipment, Radar equipment, Sonar equipment, Transponders
- 367: Electronic Components and Accessories
  - 3671: Radio and TV Receiving Type Electron Tubes, except Cathode Ray
  - 3672: Cathode Ray TV Picture Tubes

- 3673: Transmitting, Industrial and Special Purpose Electron Tubes
- 3674: Semiconductors and Related Devices
- 3675: Electronic Capacitors
- 3676: Resistors for Electronic Applications
- 3677: Electronic Coils, Transformers, and Other Inductors
- 3678: Connectors for Electronic Applications
- 3679: Electronic Components (n.e.c.)
    - Antennas, Circuit boards, Magnetic recording tape, Oscillators, Relays

B.4.2 Historical firm creation from Dun & Bradstreet

We supplement these data with distinct data from Dun & Bradstreet (D&B), which has for two centuries collected data on U.S. firms for the purpose of assessing creditworthiness, making these data commercially available. Electronic records begin in roughly 1970, and modern D&B data cover millions of U.S. firms. These data are useful for our purposes because they measure firms' locations, 4-digit SIC industries, and founding years, enabling us to produce a measure of business creation by county, SIC, and year. An important caveat of these data are that our measurement is conditioned on both inclusion in the D&B sample (though this is very large) and survival to a given data year (this is more limiting, especially with the passage of time). Because it appears the D&B sample grew significantly over the 1970s, and in order to be able to include the 1970s in our analysis, we use the 1980 D&B data, balancing data quality against the survival of older cohorts. As with the CBP data, we focus our attention on firms in SIC 366 and 367, measuring the number of firms created in these SICs by county and decade.

B.4.3 Rosters of technical staff at OSRD-funded R&D labs

We also collect data on individuals who were at two major OSRD-funded R&D labs during the war: the MIT Radiation Laboratory (Rad Lab or RL), which was the epicenter of the Allied radar research effort during WW2, and the Harvard Radio Research Laboratory (RRL), which worked on radar countermeasures (i.e., stealth movement and enemy radar jamming). Both labs were large, employing thousands of scientists and engineers over the course of the war, and the Rad Lab was considered so successful that it became a model for post-war federally-funded research labs and is celebrated as an important part of the history of MIT and the broader region.

We use records from the Harvard and MIT university archives to compile rosters of technical staff from each of these labs. The MIT archive's collection of Rad Lab records contains three documents useful towards this end: (i) a list of former RL staff members published in June 1946, which lists the name, field, highest degree, year of degree, and years spent at the RL; (ii) the most reliable known address of former staff members, as of March 1946; and (iii) the post-war job (or grad school) placement of former staff members, as of January 1946. We digitize these data sources, harmonizing name formats in the process so that we can successfully link individuals across them. Although some individuals appear in only a subset of these records, for most we have complete records. From the Harvard archive, we observe only RRL staff member addresses, as of March 1946, without the ancillary information on educational background or post-war job placement. We supplement this information by collecting data on all pre-, mid-, and post-war patents by these individuals, which we can use to study the effects that these individuals might have had on local inventive activity after the war in the fields where they were active.

B.4.4 Universities and PhD student production across the 20th century

We also collect data on historical PhD production at U.S. universities, by field and decade. Our source for these data is the 1963 report *Doctorate Production in United States Universities, 1920-1962*, published jointly by the National Academy of Sciences (NAS) and National Research Council (NRC). This report provides a wide range of information on PhD graduates and the baccalaurate institutions that feed PhD programs, collected from PhD-granting universities. We use this source to count up the number of PhD graduates in each county and decade, by field (fields include the physical sciences, biological sciences, social sciences, arts and professions, and education). Using these data we can test whether the agglomerative effects of World War II R&D may have operated through the postwar growth of local research universities.

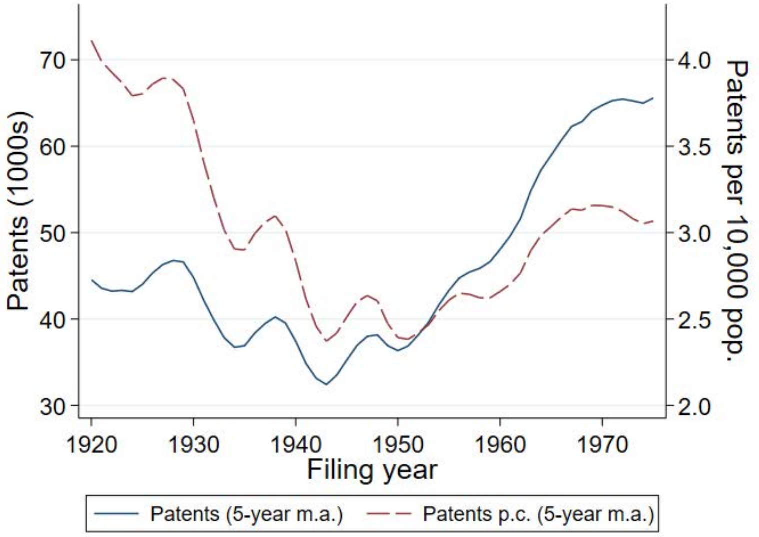
B.4.5 Identities and locations of FFRCs (FFRDCs)

To investigate another channel through which the OSRD shock may have persisted, we collect data on post-war Federally-funded Research Centers (FFRCs), a category of government-funded research labs which proliferated in the post-war era, now known as Federally-funded Research and Development Centers (FFRDCs). Our source is the NSF's annual *Federal Funds for Science* publication (later titled *Federal Funds for Research, Development, and Other Scientific Activities*, NSF (1957)), which lists active FFRDCs in the 1958-1977 editions identifying: (i) the name, (ii) the funding agency (e.g., Department of Defense, Atomic Energy Commission, NASA), and (iii) the organization administering the lab (typically a firm or university, although these often had only an arms-length relationship and were a conduit for staff and/or funding). An example is Lincoln Lab, which was funded by the Air Force and managed by MIT. After compiling a list of FFRDCs by year, we then manually research each FFRDC's (i) type (research lab, study and analysis center, test facility, or systems/technical direction), (ii) location, and (iii) subject. Using these data we can test whether World War II R&D was associated with the establishment (or, in some cases, continuation) of local federally-funded research centers.

C Supplementary Results

C.1 U.S. patenting in the 20th century

Figure C.1: Aggregate U.S. patenting, 1920 to 1975



Notes: Figure shows the time series of U.S. patents, by filing year, 1920 to 1975.

C.2 Top OSRD-supported technology fields

Table C.1: Top 10 technology areas with OSRD patents (denominator: all patents)

NBER	Description	Patents from OSRD contracts	Pct. of patents from OSRD, 1941-48	Max pct. OSRD in any year, 1941-48
44	Nuclear, X-rays	194	12.5%	24.8%
21	Communications	671	6.9%	16.6%
46	Semiconductor devices	15	5.4%	12.1%
42	Electrical lighting	241	4.2%	10.0%
22	Computer hardware/software	65	4.1%	8.5%
43	Measuring, testing	187	3.1%	6.8%
41	Electrical devices	308	2.5%	6.5%
45	Power systems	163	1.7%	4.2%
31	Drugs	27	1.7%	6.4%
49	Misc. (elec)	93	1.6%	3.7%

Notes: Table lists the top 2-digit NBER technology subfields (Hall et al. 2001), ranked by the fraction of 1941-1948 patents which were OSRD-funded, and the maximal fraction of OSRD-funded patents in any year.

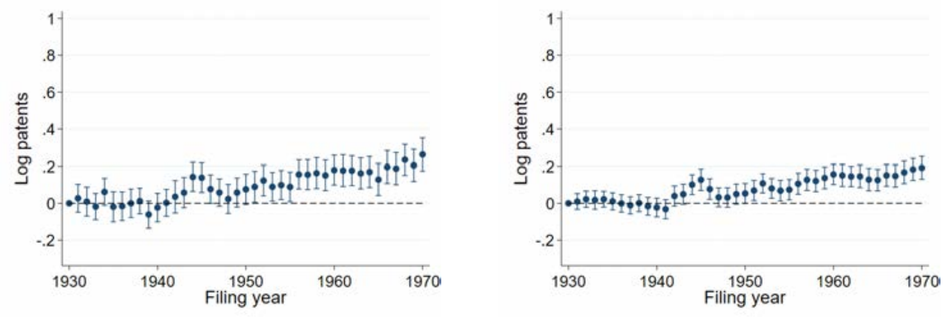
Table C.2: Top 10 patent classes with OSRD patents (denominator: all patents)

USPC	Description	Patents from OSRD contracts	Pct. of patents from OSRD, 1941-48	Max pct. OSRD in any year, 1941-48
376	Induced nuclear reactions	62	33.9%	81.8%
367	Acoustic wave systems/devices	95	20.4%	41.0%
102	Ammunition and explosives	174	16.6%	39.4%
343	Radio wave antennas	104	12.6%	33.9%
380	Cryptography	29	12.0%	21.4%
250	Nuclear energy	120	12.0%	23.8%
342	Directive radio wave systems/devices (radar)	200	10.9%	21.9%
333	Wave transmission lines and networks	108	10.5%	20.0%
327	Misc. electrical devices, circuits, and systems	88	10.3%	25.8%
708	Electrical computers	15	9.4%	22.2%

Notes: Table lists the top patent classes ranked by the fraction of 1941-1948 patents which were OSRD-funded, and the maximal fraction of OSRD-funded patents in any year.

C.3 Additional results

Figure C.2: Effects of OSRD on citation-weighted patents and per-capita patents


Panel (A): Citation-weighted patents

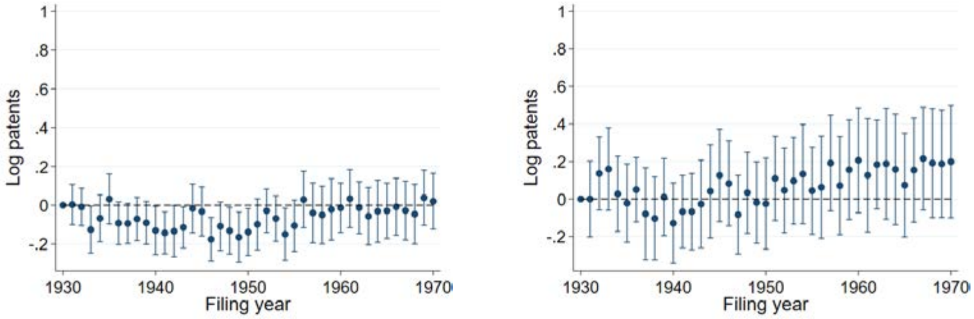
Panel (B): Per capita patents

Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting on citation-weighted patents and per-capita patents. The independent variable measures the log fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure C.3: Effects of OSRD on cluster patenting, 1930-1970, by technology area

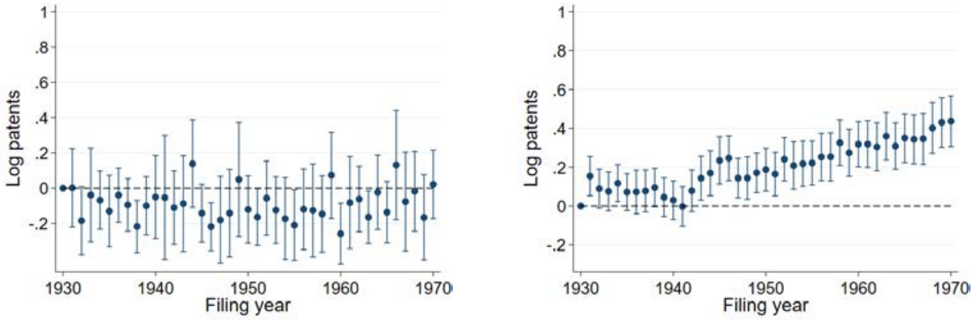
Panel (A): Chemicals (10s)

Panel (B): Communications (20s)



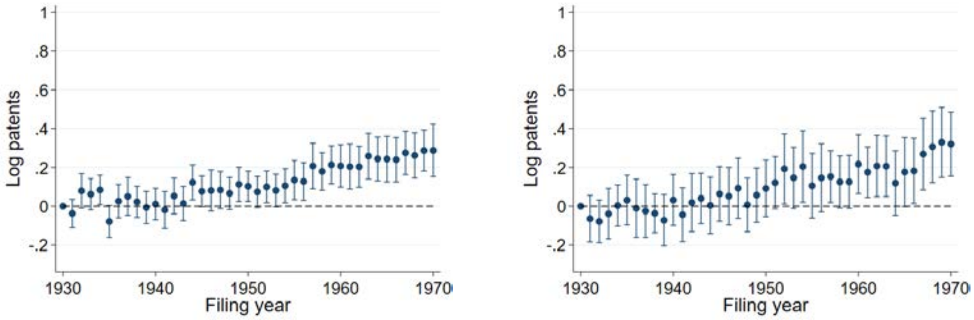
Panel (C): Drugs & Medical (30s)

Panel (D): Electrical/Electronics (40s)



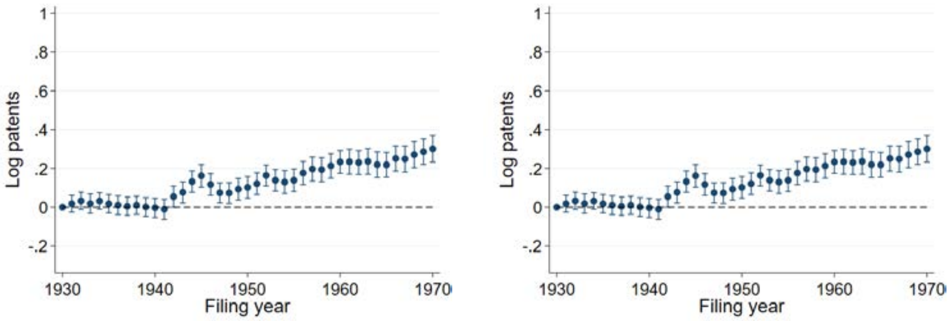
Panel (E): Mechanical (50s)

Panel (F): Other (60s)

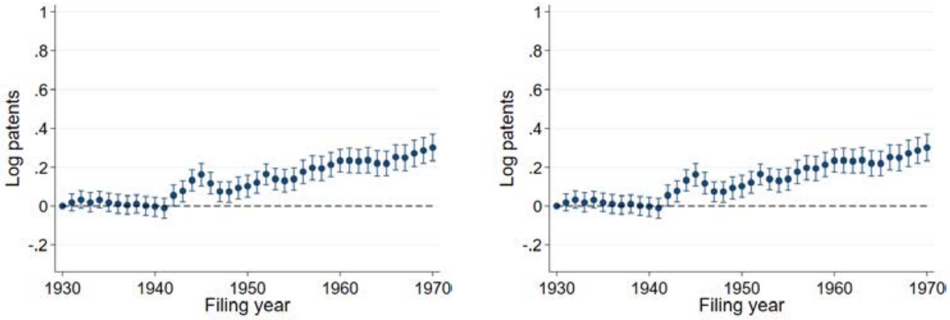


Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting, splitting the sample into high-level technology categories (Hall et al. 2001). The independent variable measures the log fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure C.4: Effects of OSRD on cluster patenting, 1930-1970, excluding select states


Panel (A): Omitting CA

Panel (B): Omitting MA


Panel (C): Omitting NJ

Panel (D): Omitting NY

Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting, omitting select states from the estimation sample (CA, MA, NJ, NY). The independent variable measures the log fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

C.4 Alternative treatment measure: OSRD intensity quartiles

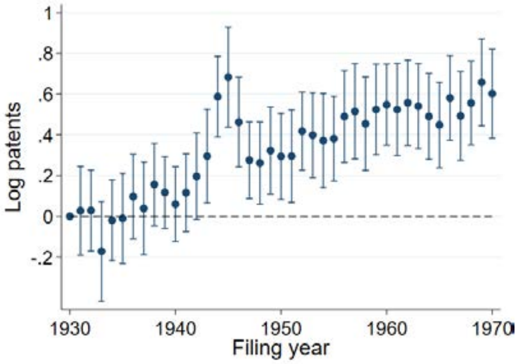
The following figures re-estimate our agglomeration results using a categorical treatment measure. We group clusters (county-categories) in four quartiles of treatment intensity (fraction of 1941-1948 patents OSRD-funded, conditional on any), and estimate changes in cluster patenting over time by treatment quartile, relative to a reference category of clusters without any OSRD-funded invention. Concretely, the specification we estimate here and elsewhere is:

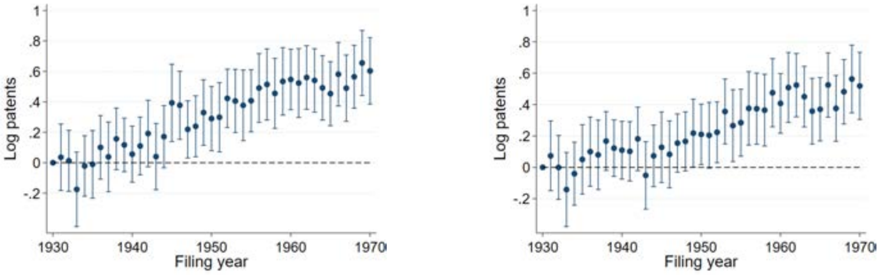
$$Ln(Patents)_{ict} = \sum_{q=1}^{4} \sum_{t=1931}^{1970} \beta_{qt} \cdot \mathbb{1}(\text{Treatment quartile } q) \cdot Year_t + \alpha_{ic} + \delta_t + \varepsilon_{ict}$$

where i indexes counties, c indexes patent categories, and t indexes years, and the sample runs from 1930 to 1970, with standard errors clustered at the county-category level. Figure C.5 presents our key result for the top quartile of treated clusters, reproducing Figure C.5 of the paper with this alternative treatment measure. Figure C.6 presents estimates for all four treatment quartiles, where the effect of the World War II shock can be seen to diminish in treatment intensity. Our preferred treatment measure for this paper is continuous, rather than categorical, because continuous measure preserve power (whereas categorical measures slice the sample into smaller cells, with wider confidence bands, as can be seen in these figures).

Figure C.5: Effects of OSRD on cluster patenting, 1930-1970, top treatment quartile

Panel (A): All patents

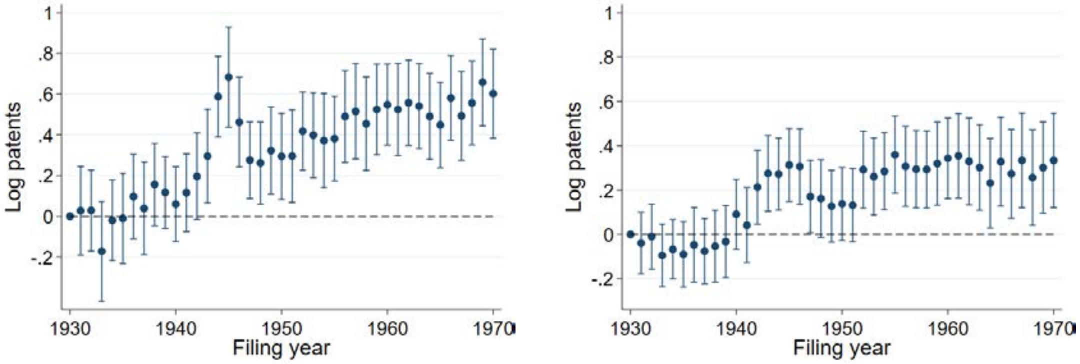



Panel (A): Non-OSRD patents

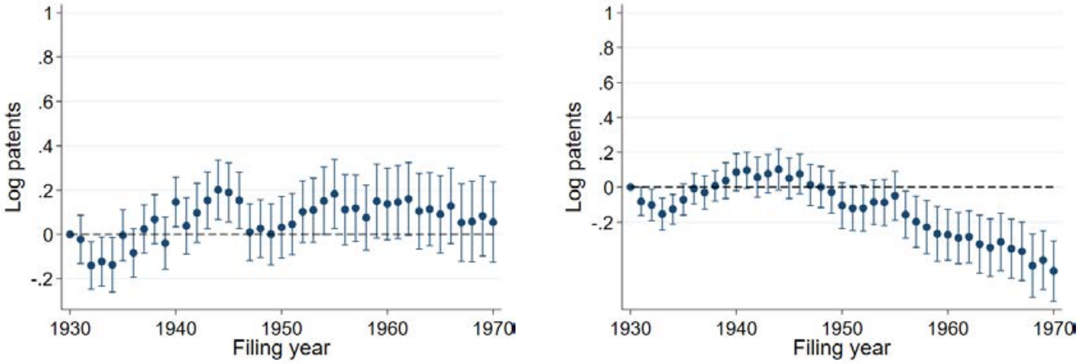
Panel (B): Non-govt. interest patents

Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting. The independent variable measures the log fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure C.6: Effects of OSRD on cluster patenting, 1930-1970, by treatment intensity quartile


Panel (A): Top quartile

Panel (B): 3rd quartile


Panel (C): 2nd quartile

Panel (D): Bottom quartile

Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting. The independent variable measures the quartile of treatment intensity, conditional on treatment (the fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded, conditional on any), with parameters across all four panels estimated jointly (in one regression) relative to a reference group of county-categories without any OSRD patents. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

C.5 Correlation of OSRD and postwar government-funded patenting

Table C.3 estimates the relationship of World War II OSRD patenting to postwar government-funded patenting across county-categories. The outcome variable is measured as the fraction of 1950-1969 patents that are government-supported, in: levels (Column 1); logs (Column 2); relative to the median, conditional on non-zero (Columns 3 and 4); and in quartiles, conditional on nonzero (Columns 5 to 8). The explanatory variable is measured as the fraction of 1941-1948 patents that were OSRD-funded, in levels, logs, and quartiles. The table illustrates the strong path dependence in the local intensity of government-funded invention.

Table C.3: Correlation of OSRD patenting and postwar govt-funded patenting across clusters

	USG rate, 1950-1969		Rel. to median		Quartiles			
	(1) Level	(2) Log	(3) Above	(4) Below	(5) Top	(6) Upper-mid	(7) Lower-mid	(8) Bot.
Wartime OSRD rate, 1941-1948	0.370*** (0.071)							
Ln(Wartime OSRD rate)		0.463*** (0.040)						
Wartime OSRD rate in top quartile			0.417*** (0.045)	0.095*** (0.036)	0.200*** (0.052)	0.218*** (0.038)	0.118*** (0.032)	-0.023 (0.019)
Wartime OSRD rate in 3rd quartile			0.348*** (0.049)	0.262*** (0.043)	0.108** (0.043)	0.240*** (0.048)	0.223*** (0.037)	0.039 (0.025)
Wartime OSRD rate in 2nd quartile			0.257*** (0.044)	0.460*** (0.042)	0.013 (0.022)	0.243*** (0.038)	0.290*** (0.038)	0.170*** (0.037)
Wartime OSRD rate in bot. quartile			0.063* (0.038)	0.672*** (0.039)	-0.022* (0.011)	0.085*** (0.031)	0.229*** (0.036)	0.443*** (0.048)
N	14517	606	14517	14517	14517	14517	14517	14517
R^2	0.02	0.30	0.04	0.06	0.02	0.03	0.03	0.03
Y Mean	0.04	-2.48	0.10	0.17	0.04	0.07	0.08	0.09

Notes: Table relates county-category postwar government-funded patents to wartime OSRD patents. All columns include county-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county-category in parentheses.

C.6 County-level institutional outcomes

The bulk of our paper focuses on the effects of World War II R&D on cluster-level outcomes, where we define clusters as the cross of counties by technology areas (e.g., semiconductors in Santa Clara, CA). Table C.4 examines more aggregated county-level treatment and outcomes, including population, PhD graduates from local universities (overall and by field), and the establishment of federally-funded research centers (FFRDCs). Population and PhD production are measured decadally, for both the pre-war and postwar period, whereas FFRDCs are only measured for the 1950s and 1960s. We do not find a statistical relationship between the county-level OSRD treatment and population or PhD production. Counties with high treatment intensity are, however, more likely to have a postwar FFRDC. Broadly, this is capturing the network of nuclear energy research laboratories which emerged out of the World War II atomic fission program, initiated by OSRD and later transferred to the Army under the Manhattan Project.[15]

Table C.4: County-level treatment and outcomes: PhD production and FFRDC installations

		PhD production, by field					FFRDCs	
	(1) Pop	(2) All	(3) Phys. Sci.	(4) Biol. Sci.	(5) Soc. Sci.	(6) Any	(7) Electronics	(8) Nuclear
Ln(OSRD rate) * 1(1940-1949)	0.010	-0.027	-0.046*	0.006	0.037			
	(0.009)	(0.041)	(0.024)	(0.014)	(0.033)			
Ln(OSRD rate) * 1(1950-1959)	0.003	-0.078	-0.035	0.023	0.013	0.056***	0.015	0.040**
	(0.019)	(0.093)	(0.068)	(0.050)	(0.078)	(0.019)	(0.011)	(0.016)
Ln(OSRD rate) * 1(1960-1970)	-0.004	-0.044	-0.023	-0.016	0.016	0.052***	0.012	0.035**
	(0.027)	(0.089)	(0.070)	(0.064)	(0.063)	(0.018)	(0.009)	(0.015)
N	6888	840	840	840	840	2184	2184	2184
R^2	0.97	0.78	0.78	0.78	0.77	0.05	0.02	0.04
Y mean	3.63	1.56	1.18	1.00	0.99	0.14	0.03	0.06
County FEs	X	X	X	X	X			
Year FEs	X	X	X	X	X	X	X	X

Notes: Table estimates the effect of the OSRD shock on county-level population, PhD production and federally-funded research centers (FFRDCs). PhD graduate counts by institution, decade and subject obtained from NAS (1963) and aggregated up to counties, with regressions estimated at the county-decade level. FFRDC data obtained from NSF (1957) and manually coded by subject. All columns include year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county in parentheses.

---

[15] These include sites in remote locations such as Oak Ridge (Anderson, TN), Hanford (Benton, WA), and Los Alamos (Los Alamos and Bernalillo, NM), where wartime research comprised the majority of their 1940s patenting, as well as site in larger, urban counties such as Berkeley/Livermore (Alameda, CA) and Argonne (Cook, IL). These were among the most intensively-treated counties, when measured at the county level.

C.7 Aggregate outcomes

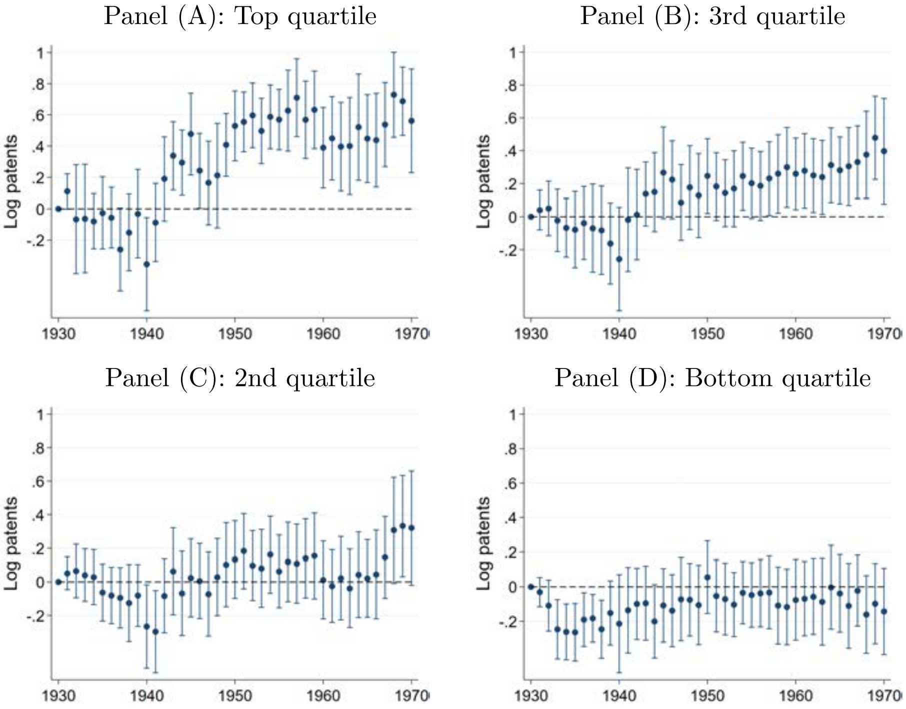
The following figures present annual estimates of the effects of World War II R&D on the direction of U.S. invention, disaggregating the simple difference-in-differences presented in Figure 9 in the body of the paper. We estimate the following specification:

$$Ln(Patents)_{ict} = \sum_{q=1}^{4} \sum_{t=1931}^{1970} \beta_{qt} \cdot \mathbb{1}(\text{Country } i = \text{US}) \cdot \mathbb{1}(\text{Class } c \in \text{quartile } q) \cdot Year_t$$
$$+ \alpha_{ic} + \delta_{it} + \gamma_{ct} + \varepsilon_{ict}$$

where i, c, and t index countries, technology classes, and years, the latter terms represent interacted fixed effects, and standard errors are clustered at the country-class level.

Figure C.7 presents the estimates for patenting at the USPTO vs. UK patent authorities. Figure C.8 does so for USPTO patents of U.S. versus foreign inventors. In both cases, the effect of the World War II shock is visible but diminishes in treatment intensity.

Figure C.7: Patenting at the USPTO versus the UK patent authority, by quartile of treated patent classes (relative to untreated classes), 1930-1970


Panel (A): Top quartile

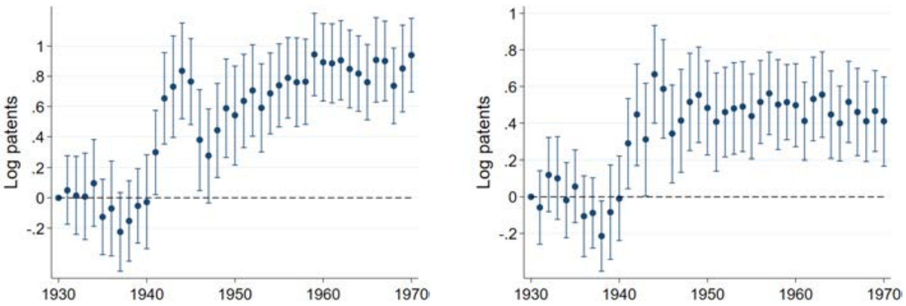
Panel (B): 3rd quartile

Panel (C): 2nd quartile

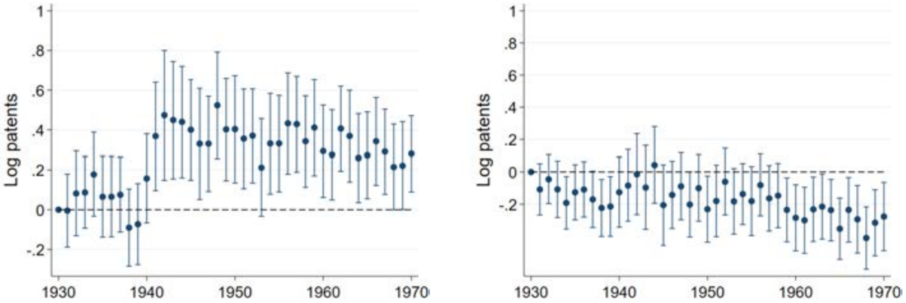
Panel (D): Bottom quartile

Notes: Figure shows annual difference-in-differences estimates of the effects of the OSRD shock on patenting at USPTO versus British patent authority, in technology classes (IPC classes) in each quartile of OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-funded, versus those without OSRD treatment. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-class level.

Figure C.8: Patenting at the USPTO by domestic versus foreign inventors, by quartile of treated patent classes (relative to untreated classes), 1930-1970


Panel (A): Top quartile

Panel (B): 3rd quartile


Panel (C): 2nd quartile

Panel (D): Bottom quartile

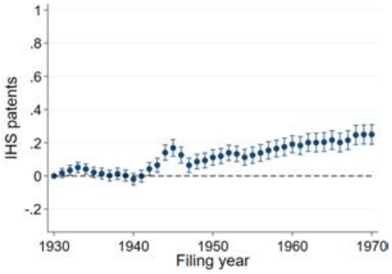
Notes: Figure shows annual difference-in-differences estimates of the effects of the OSRD shock on USPTO patents with U.S. versus foreign inventors, in technology classes (USPCs) in each quartile of OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-funded, versus those without OSRD treatment. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-class level.

D Alternative outcome measures

This appendix reproduces our main results on the postwar agglomeration of U.S. invention, replacing logged outcome measures with the inverse hyperbolic sine transformation. Figures D.1 to D.5 reproduce Figures 4 to 8. Tables D.1 to D.2 reproduce Tables 3 to 4.

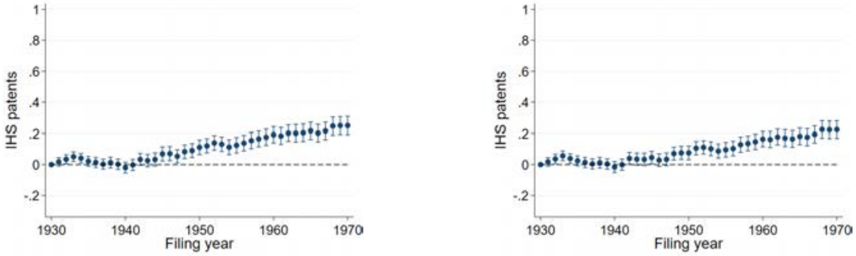
Figure D.1: Effects of OSRD on cluster patenting, 1930-1970

Panel (A): All patents



Panel (B): Non-OSRD patents

Panel (C): Non-govt. interest patents

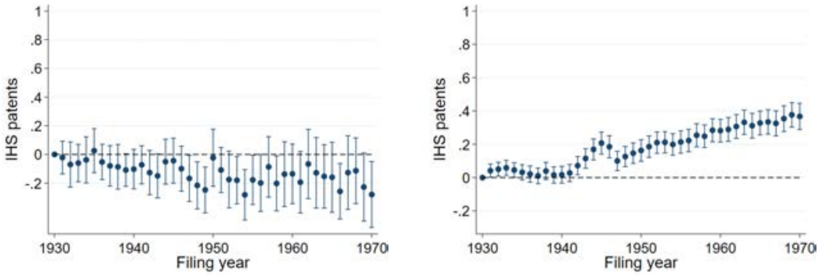


Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting. The independent variable measures the IHS fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure D.2: Effects of OSRD on cluster patenting, for clusters in counties in the bottom 95% versus top 5% of 1930s patenting

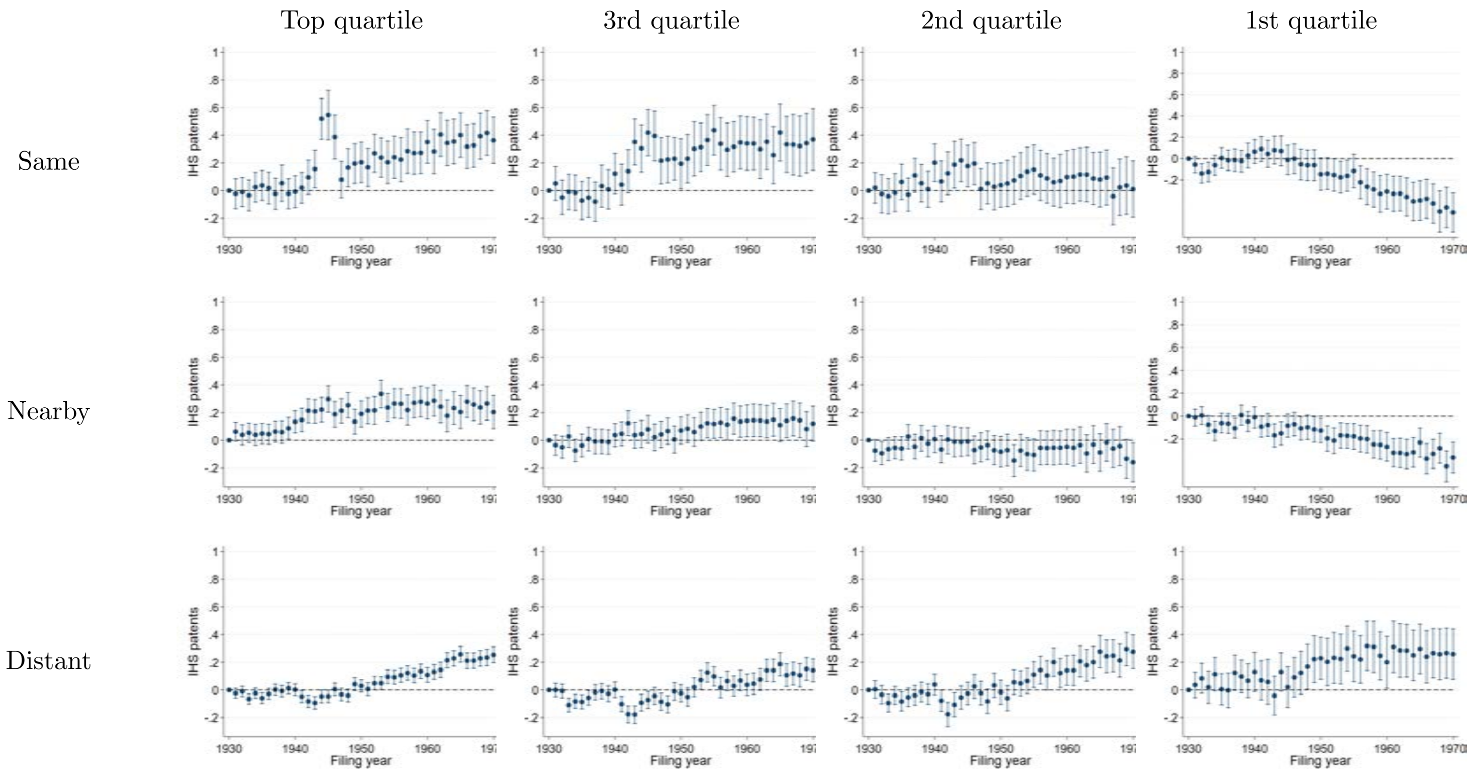
Panel (A): Bottom 95%

Panel (B): Top 5%



Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting, for counties in the bottom 95% and top 5% of 1930s patenting (i.e., existing technology clusters). Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure D.3: Effects of OSRD on cluster patenting, 1930-1970, cross-technology area spillovers
horserace regression of treatment in (i) same technology area, (ii) nearby technology areas, (iii) more distant technology areas

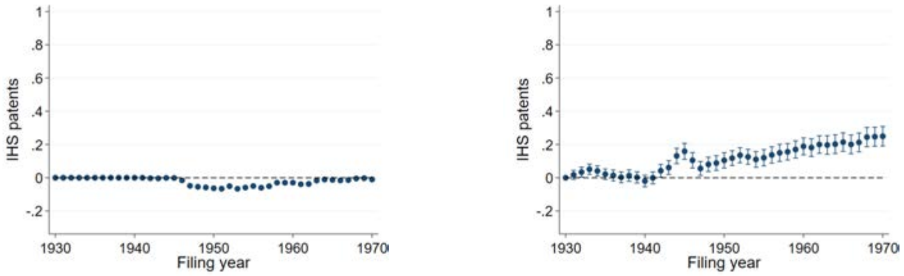


Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting. The independent variable measures the quartile of treatment intensity, conditional on treatment (the fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-funded, conditional on any), of three types: (i) in the given county-category (top row); (ii) in other, similar county-categories (same 1-digit NBER category, per Hall et al. (2001); middle row); and (iii) in other, more distant county-categories (other 1-digit NBER categories; bottom row). Parameters across all panels are estimated jointly (in one regression) relative to a reference group of county-categories without any OSRD patents. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure D.4: Effects not explained by OSRD's direct impacts on local invention

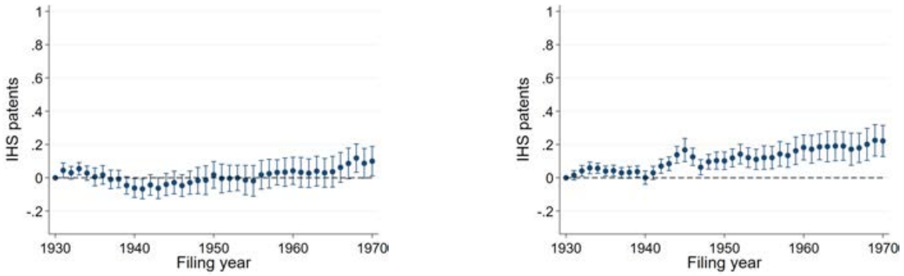
Panel (A1):
Patents citing OSRD patents

Panel (A2):
Patents not citing OSRD patents



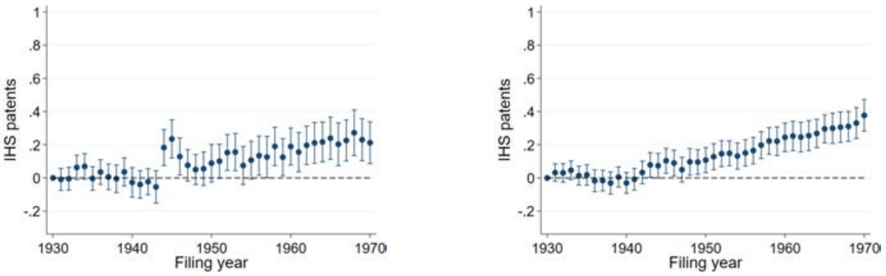
Panel (B1):
Patents assigned to OSRD contractors

Panel (B2):
Patents assigned to others



Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category (i) patents citing versus not citing OSRD patents, and (ii) patents assigned to OSRD contractors versus others, as an exploration of the direct impacts of OSRD on postwar invention in the treated clusters. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Figure D.5: Effects not explained by sustained government investment in local invention


Panel (A):
High postwar govt. patent clusters

Panel (B):
Low postwar govt. patent clusters

Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patents in county-categories with above and below median postwar (1950-1969) government-funded patent rates, as an exploration of the role of sustained public R&D investment as an explanation for persistence. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Table D.1: Firm patents: All, incumbents, and entrants
(incumbents by geographic and technological proximity)

		Same county		Diff. county		
	(1) All	(2) Same field	(3) Diff. field	(4) Same field	(5) Diff. field	(6) New
Ln(OSRD rate) * 1(1935-1939)	-0.019*	-0.040***	0.026***	0.029***	0.011**	0.021**
	(0.010)	(0.013)	(0.008)	(0.010)	(0.005)	(0.009)
Ln(OSRD rate) * 1(1940-1944)	-0.016	-0.056***	0.038***	0.042***	0.016***	0.051***
	(0.018)	(0.021)	(0.010)	(0.010)	(0.006)	(0.009)
Ln(OSRD rate) * 1(1945-1949)	0.018	-0.013	0.053***	0.058***	0.022***	0.035***
	(0.024)	(0.028)	(0.011)	(0.012)	(0.007)	(0.012)
Ln(OSRD rate) * 1(1950-1954)	0.047	0.009	0.064***	0.064***	0.017**	0.038**
	(0.032)	(0.037)	(0.019)	(0.016)	(0.008)	(0.017)
Ln(OSRD rate) * 1(1955-1959)	0.063	0.024	0.031	0.025	0.014	0.025
	(0.040)	(0.046)	(0.024)	(0.022)	(0.011)	(0.023)
Ln(OSRD rate) * 1(1960-1964)	0.090**	0.043	0.051*	0.045*	0.007	0.020
	(0.045)	(0.050)	(0.029)	(0.023)	(0.011)	(0.026)
Ln(OSRD rate) * 1(1965-1970)	0.117**	0.077	0.052*	0.027	-0.001	0.006
	(0.049)	(0.056)	(0.030)	(0.028)	(0.012)	(0.029)
N	28413	28413	28413	28413	28413	28413
R^2	0.81	0.79	0.55	0.48	0.23	0.61
Y mean	1.88	1.57	0.56	0.52	0.12	0.47
County-cat FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

Notes: Table estimates the effect of the OSRD shock on the growth of firm patenting. Observations are at the county-category-year level, and outcome variables are measured in IHS. Column (1) measures all firm patents; Column (2), patents invented by firms with prior inventions in the same county and technology area; Columns (3) to (5), firms with prior patents only in a different county and/or technology area; and Column (6), firms with no prior patents. All columns include county-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county-category in parentheses.

Table D.2: All patents vs. government-funded patents (total and by agency)

	(1) All	(2) Non govt.	(3) Govt.	(4) DOD	(5) DOE	(6) NASA
Ln(OSRD rate) * 1(1935-1939)	-0.017*	-0.018**	-0.015			-0.001
	(0.010)	(0.009)	(0.010)			(0.001)
Ln(OSRD rate) * 1(1940-1944)	-0.092***	-0.011	-0.085***	-0.005*		-0.003
	(0.013)	(0.014)	(0.014)	(0.003)		(0.004)
Ln(OSRD rate) * 1(1945-1949)	-0.069***	0.021	-0.065**	-0.004		-0.007
	(0.025)	(0.021)	(0.025)	(0.009)		(0.008)
Ln(OSRD rate) * 1(1950-1954)	-0.077**	0.066**	-0.079**	-0.004		-0.004
	(0.030)	(0.026)	(0.031)	(0.007)		(0.006)
Ln(OSRD rate) * 1(1955-1959)	-0.059*	0.092***	-0.062*	0.001		-0.004
	(0.032)	(0.034)	(0.033)	(0.011)		(0.005)
Ln(OSRD rate) * 1(1960-1964)	-0.016	0.137***	-0.025	0.009	-0.012*	-0.002
	(0.035)	(0.039)	(0.032)	(0.011)	(0.007)	(0.004)
Ln(OSRD rate) * 1(1965-1970)	0.035	0.174***	0.032	0.010	-0.010	-0.002
	(0.031)	(0.043)	(0.027)	(0.010)	(0.011)	(0.003)
N	28413	28413	28413	28413	28413	28413
R^2	0.49	0.83	0.48	0.33	0.16	0.36
Y mean	0.51	2.15	0.42	0.06	0.01	0.01
County-cat FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

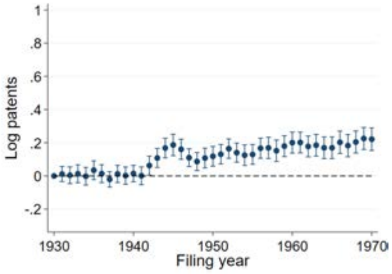
Notes: Table estimates the effect of the OSRD shock on the growth of government-funded patenting. Observations are at the county-category-year level, and outcome variables are measured in logs. Column (1) measures all government-funded patents; Column (2), non-government funded patents; and Columns (3) to (6), patents by agency: Department of Defense (DOD), Department of Energy (DOE), National Aeronautics and Space Administration (NASA), and Department of Agriculture (USDA). Columns are labeled with modern agencies, but the DOD and DOE categories include predecessor agencies before they were established in 1947 and 1977, respectively. All columns include county-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county-category in parentheses.

E CBSA-level analysis

This appendix reproduces our main results on the postwar agglomeration of U.S. invention, replacing logged outcome measures with the inverse hyperbolic sine transformation. Figures E.1 to E.5 reproduce Figures 4 to 8. Tables E.1 to E.2 reproduce Tables 3 to 4.

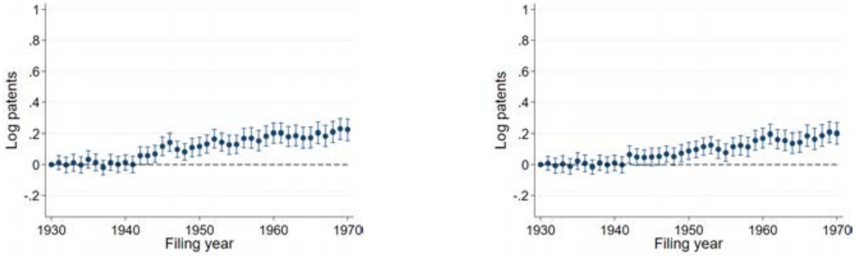
Figure E.1: Effects of OSRD on cluster patenting, 1930-1970

Panel (A): All patents



Panel (B): Non-OSRD patents

Panel (C): Non-govt. interest patents

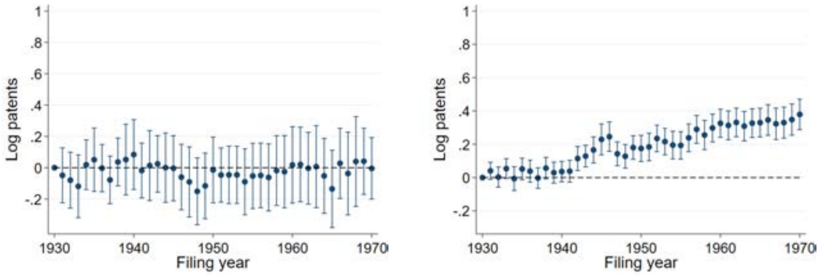


Notes: Figure shows annual estimates of the effects of the OSRD shock on CBSA-category patenting. The independent variable measures the log fraction of U.S. patents in each CBSA-category between 1941-1948 which were OSRD-funded. Error bars represent 95% confidence intervals, computed from SEs clustered at the CBSA-category level.

Figure E.2: Effects of OSRD on cluster patenting, for clusters in CBSAs in the bottom 95% versus top 5% of 1930s patenting

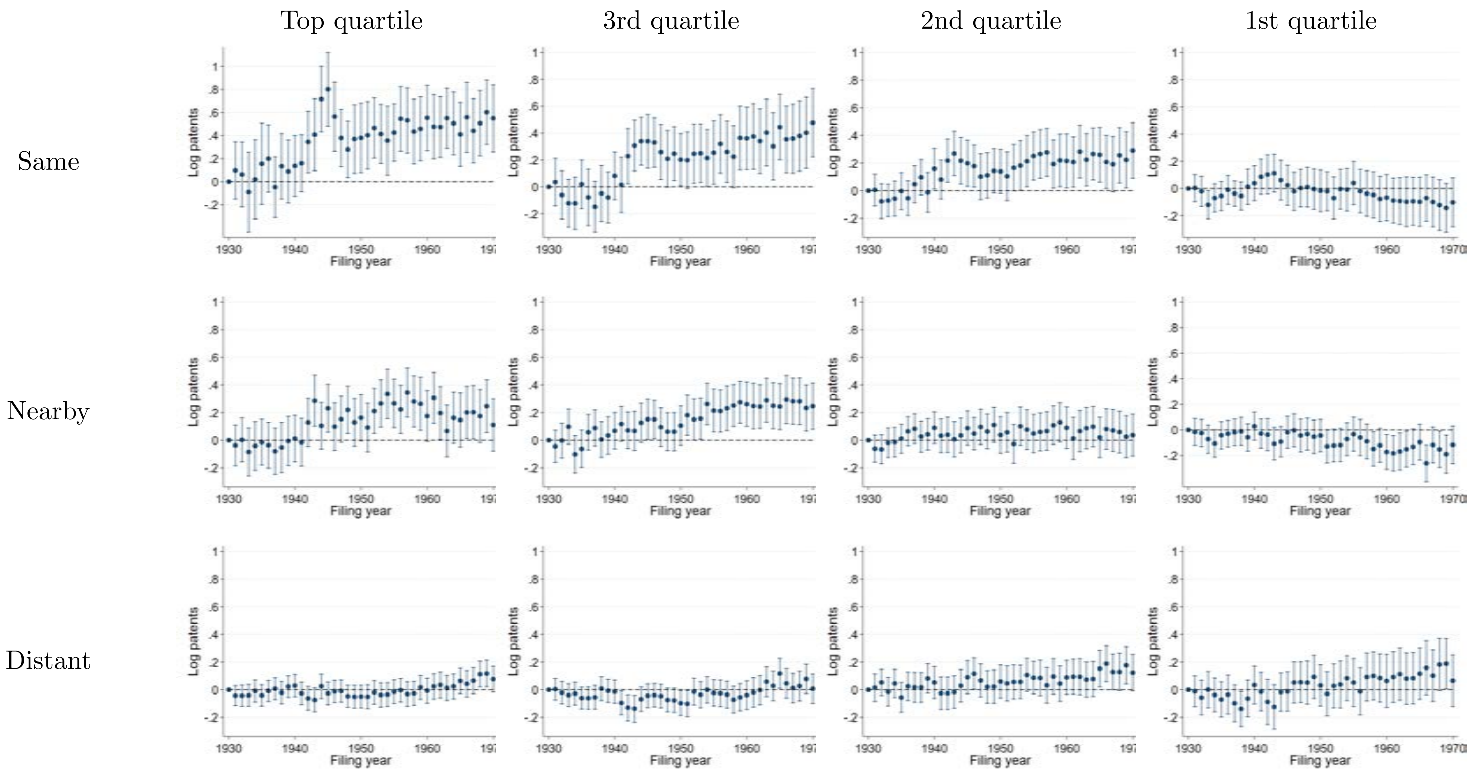
Panel (A): Bottom 95%

Panel (B): Top 5%



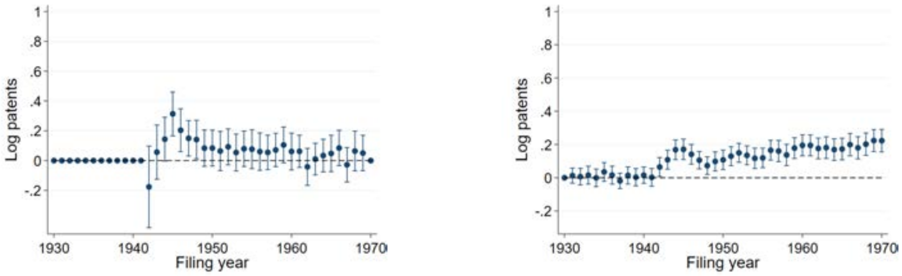
Notes: Figure shows annual estimates of the effects of the OSRD shock on CBSA-category patenting, for CBSAs in the bottom 95% and top 5% of 1930s patenting (i.e., existing technology clusters). Error bars represent 95% confidence intervals, computed from SEs clustered at the CBSA-category level.

Figure E.3: Effects of OSRD on cluster patenting, 1930-1970, cross-technology area spillovers
horserace regression of treatment in (i) same technology area, (ii) nearby technology areas, (iii) more distant technology areas



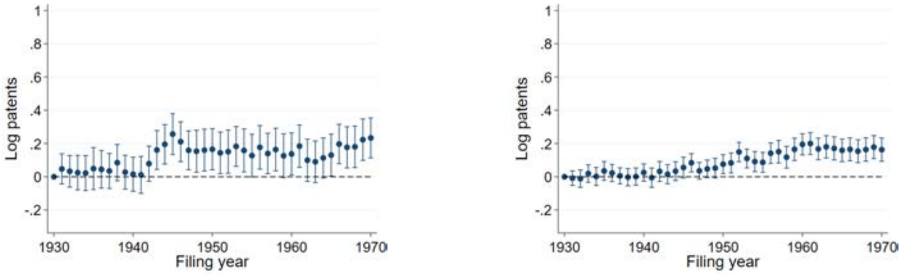
Notes: Figure shows annual estimates of the effects of the OSRD shock on CBSA-category patenting. The independent variable measures the quartile of treatment intensity, conditional on treatment (the fraction of U.S. patents in each CBSA-category between 1941-1948 which were OSRD-funded, conditional on any), of three types: (i) in the given CBSA-category (top row); (ii) in other, similar CBSA-categories (same 1-digit NBER category, per Hall et al. (2001); middle row); and (iii) in other, more distant CBSA-categories (other 1-digit NBER categories; bottom row). Parameters across all panels are estimated jointly (in one regression) relative to a reference group of CBSA-categories without any OSRD patents. Error bars represent 95% confidence intervals, computed from SEs clustered at the CBSA-category level.

Figure E.4: Effects not explained by OSRD's direct impacts on local invention



Panel (A1):
Patents citing OSRD patents

Panel (A2):
Patents not citing OSRD patents

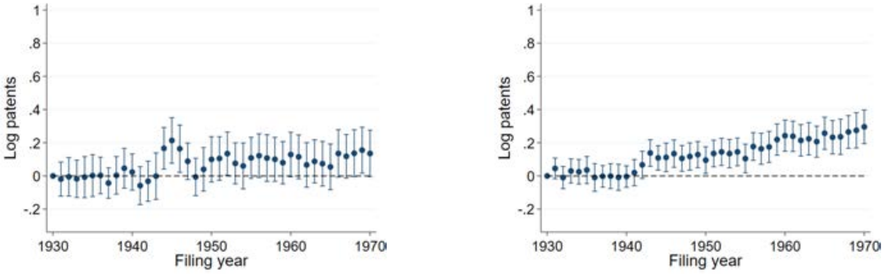


Panel (B1):
Patents assigned to OSRD contractors

Panel (B2):
Patents assigned to others

Notes: Figure shows annual estimates of the effects of the OSRD shock on CBSA-category (i) patents citing versus not citing OSRD patents, and (ii) patents assigned to OSRD contractors versus others, as an exploration of the direct impacts of OSRD on postwar invention in the treated clusters. Error bars represent 95% confidence intervals, computed from SEs clustered at the CBSA-category level.

Figure E.5: Effects not explained by sustained government investment in local invention



Panel (A):
High postwar govt. patent clusters

Panel (B):
Low postwar govt. patent clusters

Notes: Figure shows annual estimates of the effects of the OSRD shock on CBSA-category patents in CBSA-categories with above and below median postwar (1950-1969) government-funded patent rates, as an exploration of the role of sustained public R&D investment as an explanation for persistence. Error bars represent 95% confidence intervals, computed from SEs clustered at the CBSA-category level.

Table E.1: Firm patents: All, incumbents, and entrants (incumbents by geographic and technological proximity)

		Same county		Diff. county		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Same field	Diff. field	Same field	Diff. field	New
Ln(OSRD rate) * 1(1935-1939)	0.001	0.035*	0.014	0.051**	0.006	0.022
	(0.022)	(0.020)	(0.016)	(0.022)	(0.026)	(0.024)
Ln(OSRD rate) * 1(1940-1944)	0.054*	0.077**	0.066***	0.082***	0.022	0.067**
	(0.032)	(0.033)	(0.022)	(0.022)	(0.027)	(0.033)
Ln(OSRD rate) * 1(1945-1949)	0.106**	0.138***	0.073***	0.094***	0.017	0.067**
	(0.043)	(0.052)	(0.020)	(0.016)	(0.029)	(0.030)
Ln(OSRD rate) * 1(1950-1954)	0.114**	0.141***	0.108***	0.089***	0.057***	0.096***
	(0.045)	(0.050)	(0.024)	(0.019)	(0.017)	(0.029)
Ln(OSRD rate) * 1(1955-1959)	0.120***	0.163***	0.086***	0.052**	-0.010	0.097***
	(0.044)	(0.044)	(0.028)	(0.022)	(0.029)	(0.030)
Ln(OSRD rate) * 1(1960-1964)	0.145***	0.171***	0.115***	0.100***	-0.002	0.077**
	(0.046)	(0.048)	(0.032)	(0.022)	(0.018)	(0.032)
Ln(OSRD rate) * 1(1965-1970)	0.168***	0.213***	0.104***	0.058**	0.001	0.070*
	(0.050)	(0.051)	(0.035)	(0.025)	(0.019)	(0.039)
N	13140	11805	8175	7007	2349	7180
R^2	0.85	0.82	0.66	0.48	0.24	0.73
Y mean	2.43	2.36	1.05	0.73	0.28	1.03
County-cat FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

Notes: Table estimates the effect of the OSRD shock on the growth of firm patenting. Observations are at the CBSA-category-year level, and outcome variables are measured in logs. Column (1) measures all firm patents; Column (2), patents invented by firms with prior inventions in the same CBSA and technology area; Columns (3) to (5), firms with prior patents only in a different CBSA and/or technology area; and Column (6), firms with no prior patents. All columns include CBSA-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by CBSA-category in parentheses.

Table E.2: All patents vs. government-funded patents (total and by agency)

	(1)	(2)	(3)	(4)	(5)	(6)
	All	Non govt.	Govt.	DOD	DOE	NASA
Ln(OSRD rate) * 1(1935-1939)	-0.015	0.006	-0.040			0.201**
	(0.027)	(0.014)	(0.036)			(0.078)
Ln(OSRD rate) * 1(1940-1944)	0.125**	0.034*	0.115**	-0.986***		0.102
	(0.056)	(0.020)	(0.051)	(0.053)		(0.093)
Ln(OSRD rate) * 1(1945-1949)	0.212***	0.060**	0.232***	-0.854***		0.038
	(0.074)	(0.029)	(0.076)	(0.055)		(0.072)
Ln(OSRD rate) * 1(1950-1954)	0.156**	0.107***	0.188**	-0.898***		0.070
	(0.077)	(0.030)	(0.080)	(0.036)		(0.099)
Ln(OSRD rate) * 1(1955-1959)	0.194**	0.118***	0.229**	-0.869***		-0.019
	(0.080)	(0.031)	(0.087)	(0.047)		(0.093)
Ln(OSRD rate) * 1(1960-1964)	0.217***	0.164***	0.249***	-0.839***	0.026	0.113
	(0.074)	(0.034)	(0.072)	(0.040)	(0.077)	(0.111)
Ln(OSRD rate) * 1(1965-1970)	0.252***	0.183***	0.271***	-0.844***	0.108	-0.075
	(0.069)	(0.039)	(0.062)	(0.044)	(0.083)	(0.056)
N	7624	14348	6470	1526	359	389
R^2	0.61	0.89	0.61	0.50	0.39	0.47
Y mean	1.06	2.54	1.00	0.53	0.44	0.59
County-cat FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

Notes: Table estimates the effect of the OSRD shock on the growth of government-funded patenting. Observations are at the CBSA-category-year level, and outcome variables are measured in logs. Column (1) measures all government-funded patents; Column (2), non-government funded patents; and Columns (3) to (6), patents by agency: Department of Defense (DOD), Department of Energy (DOE), National Aeronautics and Space Administration (NASA), and Department of Agriculture (USDA). Columns are labeled with modern agencies, but the DOD and DOE categories include predecessor agencies before they were established in 1947 and 1977, respectively. All columns include CBSA-category and year fixed effects. *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by CBSA-category in parentheses.

Appendix references

Andrews, Michael. 2019. *Comparing Historical Patent Datasets*. Working paper.

Bergeaud, Antonin and Verluise Cyril. 2022. "PatentCity: a dataset to study the location of patents since the 19th century," URL: https://doi.org/10.7910/DVN/PG6THV.

Bergeaud, Antonin, Cyril Verluise et al. 2022. *A new Dataset to Study a Century of Innovation in Europe and in the US*. Working paper.

Berkes, Enrico. 2018. *Comprehensive Universe of U.S. Patents (CUSP): Data and Facts*. Working paper.

Duranton, Gilles, Peter M Morrow, and Matthew A Turner. 2014. "Roads and Trade: Evidence from the US," *Review of Economic Studies*, Vol. 81, No. 2, pp. 681–724.

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.

Gross, Daniel P., Maria P. Roche, and Bhaven N. Sampat. 2022. *Coordinated R&D Programs and the Creation of New Industries*. Working paper.

Gross, Daniel P. and Bhaven N. Sampat. 2022. *The World War II crisis innovation model: What was it, and where does it apply?*. NBER Working Paper No. 27909.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2001. *The NBER patent citation data file: Lessons, insights and methodological tools*. NBER Working Paper No. 8498.

Marco, Alan C., Michael Carley, Steven Jackson, and Amanda Myers. 2015a. *The USPTO Historical Patent Data Files: Two Centuries of Innovation*. SSRN working paper, available at https://ssrn.com/abstract=2616724.

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

NAS. 1963. *Doctorate production in United States universities: 1920-1962*. National Academy of Sciences–National Research Council Publication No. 1142.

NSF. 1957. *Federal Funds for Science*. 6th to 25th editions, 1957 to 1976.

Petralia, Sergio, Pierre-Alexandre Balland, and David Rigby. 2016a. "HistPat Dataset," URL: https://doi.org/10.7910/DVN/BPC15W, DOI: 10.7910/DVN/BPC15W.

Petralia, Sergio, Pierre-Alexandre Balland, and David L. Rigby. 2016b. "Unveiling the geography of historical patents in the United States from 1836 to 1975," *Scientific data*, Vol. 3, No. 1, pp. 1–14.

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

Seamans, Robert, Eunhee Sohn, and Daniel Sands. 2018. *Technological opportunity and the locus of innovation: Airmail, aircraft, and local capabilities*. Working Paper.

Stewart, Irvin. 1948. *Organizing scientific research for war: The administrative history of the Office of Scientific Research and Development*. Boston: Little, Brown, and Company.

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.