NBER WORKING PAPER SERIES

INVENTING THE ENDLESS FRONTIER:
THE EFFECTS OF THE WORLD WAR II RESEARCH EFFORT ON POST-WAR INNOVATION

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

We thank Ashish Arora, Pierre Azoulay, Wes Cohen, Jon Gruber, Adam Jaffe, Simon Johnson, Tom Nicholas, Scott Stern, and audiences at the HBS Faculty Research Symposium and the Urban Economics Association meetings (discussant Alex Whalley) for helpful comments. We also thank Hayley Pallan, Greg Saldutte, and Innessa Colaiacovo for outstanding research assistance, and the Harvard Business School Division of Faculty and Research Development and NBER Innovation Policy grant (2016) for financial support. 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.

Inventing the Endless Frontier: The Effects of the World War II Research Effort on Post-war Innovation
Daniel P. Gross and Bhaven N. Sampat
NBER Working Paper No. 27375
June 2020
JEL No. H56,N42,N72,O31,O32,O33,O38,R11

ABSTRACT

During World War II, the U.S. government launched an unprecedented effort to mobilize science for war: the newly-established Office of Scientific Research and Development (OSRD) entered thousands of R&D contracts with industrial and academic contractors, spending one to two orders of magnitude more than what the government was previously investing in science. In this paper, we study the long-run effects of the OSRD-supported research effort on U.S. invention. Using data on all OSRD contracts, we show that these investments had large effects on the direction and location of U.S. invention and high-tech industrial employment, setting in motion agglomeration forces which shaped the technology clusters of the postwar era. Our results demonstrate the effects of a large, mission-driven government R&D program on the growth of domestic technology clusters and long-run technological progress.

Daniel P. Gross
Harvard Business School
Soldiers Field
Boston, MA 02163
and NBER
dgross@hbs.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

World War II was a momentous time for U.S. science and technology, perhaps unrivaled by any other. In preparation for war, President Roosevelt authorized the creation of the National Defense Research Committee (NDRC) in 1940, which later expanded into the Office of Scientific Research and Development (OSRD). Led by Vannevar Bush, the OSRD was responsible for managing research and development (R&D) in science and technology which was considered important to winning the war, including weapons development, communications and radar, electrical engineering, jet propulsion, optics, chemistry, medicine, and atomic fission. The OSRD entered into over 2,200 R&D contracts with industrial and academic contractors during the war, with a cumulative value of nearly $8 billion current (2020) dollars – one to two orders of magnitude larger than what the U.S. government was previously investing in scientific research.

In this paper, we study how the large shock to federal research spending in World War II affected the postwar U.S. innovation economy. Using newly-collected archival data on the universe of OSRD research contracts, we study the impact of this research effort on the direction of U.S. invention, the development of domestic technology clusters, and real economic outcomes in the regions where the research took place. We show that OSRD-supported research had large, long-lasting effects on the direction and the location of U.S. invention, powering the take-off of technology clusters around the country. The results appear to be driven by clusters which were already patenting more heavily before 1940, which despite not being on a differential growth path before the war, experienced explosive growth in the postwar period. In other words, the effect of the war effort was to set in motion agglomeration forces which widened existing differences in inventive productivity across regions -– a phenomenon which has contributed to increasing polarization of economic performance across different regions of the country ever since. We then show that these forces in turn shaped the geography of high-tech industries. We are currently evaluating a number of mechanisms that might explain these remarkably persistent effects of the OSRD shock, with the goal of demonstrating and explaining the long-run effects of big-push investments in innovation and place-based innovation policies (e.g., Gruber and Johnson 2019) on local economies and ultimately U.S. national competitiveness in the latter half of the 20th century.

To study the effects of the OSRD-supported research effort, we compile a new dataset of 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 data on the complete U.S. patent record, international patenting, universities and federally-funded research centers, military R&D contracts, and industrial employment, which allow us to not only study the effects of the OSRD shock, but also understand the channels through which it propagated into the postwar period. Our workhorse empirical design will compare preand post-war patenting in technology areas shocked by the war effort (i) in the U.S. versus foreign countries, and (ii) domestically, across U.S. counties. We then drill into the electronics industry, which blossomed after the war, and compare postwar industry employment in counties with higher or lower OSRD-supported patenting in related technology areas.

We begin with the aggregate view. Comparing patenting over time in more- versus less-intensively treated technology classes at the USPTO versus at other Allied countries' patent authorities (e.g., Great Britain and France), we see a clear divergence take place after the war, without pre-trends, persisting to at least 1970 (when our estimation window ends). The magnitudes indicate that in the top quartile of treated classes, U.S. patenting was >50% higher by 1970 than in Great Britain or France, relative to their respective pre-war levels. These differences attenuate in technology classes with lower treatment intensity. When we examine just USPTO filings by inventors in the U.S. versus in foreign countries, we likewise find similar effects.

We then turn our attention inwards, to the war's effects on the geography of domestic invention. Comparing patenting in more- versus less-intensively treated county-classes, we show that patenting in the most intensively-treated county-classes (i) was not growing statistically differently prior to the war, (ii) surged during the war, (iii) contracted back to pre-war levels when the war ended, and (iv) subsequently experienced a long-lived take-off, persistently growing through the end of our analysis window. These patterns show up in other views of the data, including when comparisons are conditioned to within county-years or category-years. When we break the effects out across six aggregate technology fields, we find that the results are primarily driven by invention in the "Electrical & Electronics" and "Mechanical" categories, both of which were more intensively-treated during the war – suggesting potential nonlinearities in the effects of research investments. Moreover, the results appear to be driven by counties which were already patenting more heavily in these classes before 1940 – suggesting that although the scientific war effort did not create these technology clusters per se, it set in motion forces which led to growing agglomeration and widening disparities in the inventive output of different parts of the country.

The question is then not just *what* happened, but why it led to decades of sustained technological progress in the post-war era. Although a complete answer to this question is beyond the scope of just one paper, in ongoing research (not yet reported) we are studying intermediate mechanisms that may have contributed to the persistence of these effects. We first establish that the post-war take-off in patenting is not driven by direct follow-on invention (as measured in the citation record), nor by patents from firms and inventors involved in the war effort itself. Rather, it appears entire local research ecosystems sprung up in the locations and technology areas where OSRD activity was concentrated – including universities, federally-funded research centers, and private invention. The long-lived effects of the OSRD may in turn be reflected in higher-order patent citation linkages, and supported by university spillovers to local invention, federally-funded research centers, Cold War defense R&D spending, or other local research activities – all of which we are exploring in ongoing research. As it stands, the evidence suggests the OSRD built up local scientific and technological capabilities that allowed local innovation to thrive after the war.

Our final set of results explores how the OSRD affected not just inventive output, but entire local economies. Using U.S. County Business Patterns (CBP) data, we take our analysis to county-level manufacturing employment in industries closely tied to the wartime research effort. We show that in counties with more OSRD-supported patents in electronics technology classes, employment in the Communications and Electronics industries is substantially higher after the war – even conditional on total related patenting. Empirically, a doubling of OSRD patents in the 1940s is associated with 60-65% higher employment in these industries in the 1970s. Our interpretation is that the OSRD not only deepened local technology clusters and increased local invention over the long-run, but also created jobs in associated manufacturing industries.

There is widespread recognition that World War II was a sea-change event in government-science relations and in science and technology policy. 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, published 75 years ago this July) 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.

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 the OSRD itself. Moreover, whereas most existing evidence on the effects of federal research funding is drawn from studies of incremental changes in funding for basic science at institutions at the National Institutes of Health (NIH) and National Science Foundation (NSF), the OSRD is distinctive in its large scale, broad scope, and emphasis on technology development. Our setting thus presents an unique opportunity to study the impacts of a major shock to federal research spending – the largest ever at the time, and in relative terms, perhaps the largest to this day – on the U.S. innovation system.

A historical lens such as the one the OSRD affords is especially important for understanding the long-run impact of such investments on U.S. innovation. The long view can help us to understand the persistent effects of R&D funding on not only aggregate technological progress, but also on regional disparities in the current innovation system, a key question in policy debates today that was 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. This paper also provides evidence tying research funding to employment growth in related industries, deepening our understanding of how these investments may spill over to the rest of the economy.

Most importantly, the historical experience may also present an opportunity to draw lessons for the present and future. This paper highlights the persistent effects that the World War II research effort had on U.S. innovation. As we continue our research on the causes of this persistence, we aim to provide suggestive evidence of what it would take for similar effects to be reproduced today. As important as the funding itself is, success may also hinge on the implementation – and the devil is often in the details. Reflecting this proverb, in a concurrent essay (Gross and Sampat 2020), we describe how the OSRD mobilized U.S. science for war and use the historical episode to identify key questions that organizers of crash innovation programs must answer.

We proceed as follows. In Section 1 we review the history of the World War II mobilization of U.S. science, describe the OSRD's work and legacy, and discuss related literature. Section 2 introduces our data and presents the distribution of OSRD research across regions, firms and institutions, and technology areas. Section 3 estimates the effects of the shock on the direction of U.S. invention, comparing U.S. and foreign patenting before and after the war. Section 4 examines the effects on the location of domestic innovation and growth of U.S. technology clusters. Across all of these analyses, we find persistent effects of the OSRD shock. In Section 5, we explore whether the OSRD shock extended beyond invention and into the local industrial economy. Section 6 concludes and shares lessons for open and long-running policy debates today.

1 OSRD Background and Related Literature

1.1 Historical Background

Before World War II, there was very little federal funding of research outside of agriculture. Most academic research was funded by foundations (Rockefeller and Carnegie, in particular) and industry. There was if anything an aversion among academics to public funding, reflecting concerns that this 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 research and development effort was insufficient, and that coordination would be required to mobilize the scientific and technological capabilities that had developed during the interwar era.[1]

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[2] 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, NDRC member, 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.[3] 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. This helped bring university and industry leaders on board in support of the effort. Impressed by its early successes, the NDRC organization was expanded in a 1941 Executive Order to include more development work (beyond just research), to solidify 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 annual Congressional budget appropriations. The *New York Times* wrote that this effectively made Vannevar Bush "the czar of research."[4]

The OSRD itself comprised three constituent bodies: the 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; an advisory council, which coordinated research activities across the OSRD, military, and other civilian agencies; and an engineering and transition office, which advised on the production and use of devices developed under OSRD research. The NDRC remained the chief organ of the OSRD, with 19 major divisions (e.g., Division 1, "Ballistics Re-

[1] 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.

[2] Including 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.

[3] Scholars have since described the arrangement as having "portended the beginning of a new relationship between the federal government and the nation's universities" (Geiger 1993, page 3).

[4] In addition to Bush, Compton, Conant, and Jewett, the civilian leaders of OSRD included Richard Tolman (President of the California Institute of Technology) and Conway Coe (USPTO Commissioner).

search"; 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. The OSRD entered into >2,200 R&D contracts with industrial and academic contractors over the course of the war with a cumulative value of $7.4 billion current dollars, which, as indicated, was orders of magnitude larger than the government had previously supported.

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 the 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 developed commercial applications after the war. The research activity was heavily concentrated in a network of research labs hastily organized at the start of the war, in part based on Vannevar Bush's personal network and that of his NDRC deputies. Each had a distinct technological focus and worked in tandem with corporate R&D labs which helped develop samples and prototypes, whereas basic research was typically contracted to academic researchers around the country.[5] These early "national labs" attracted scientists and engineers from around the country,[6] many of whom dispersed at the end of the war – though some also stayed.[7]

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 the OSRD, Senator Harley Kilgore (D-W. Va.). Kilgore, a New Deal Democrat, was concerned about the concentration of OSRD funding in

[5]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 (led by Arthur Compton, Enrico Fermi), UC Berkeley (Ernest Lawrence, Robert J. Oppenheimer), Columbia University (Harold Urey) 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 persist to this day, many with the same name as during the war.

[6]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).

[7]For example, Rad Lab record show that a large number of its junior staff became MIT PhD students after the war. Prominent labs like the Rad Lab are now known for having attracted and incubated some of the country's top engineering talent: for example, David Gale's (of the Gale-Shapley deferred acceptance algorithm) first job out of college was as a Rad Lab technician. The war effort also incubated a generation of science administrators: Lee DuBridge, director of the Rad Lab, would later go on to be the president of CalTech. Fred Terman, director of the Harvard RRL, would later be Stanford's provost and colloquially, the father of Silicon Valley.

big business and a handful of universities (Kevles 1977).[8,9] Kilgore had other concerns about the 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 1977, 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 any 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 a rejoinder to this approach to organizing science.[10] Like Kilgore's proposal, 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 also anticipated the "market failure" theory of federal funding[11] (e.g., Arrow 1962) and the "linear" model of science and innovation[12] (Mowery 1997, Nelson 1997).

Although the Bush Report had a strong ideological impact on U.S. policy, its specific proposals met a cool reception, including from Kilgore and other liberals, who preferred a more egalitarian peacetime approach to supporting science and technology, 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 the remaining wartime research contracts had

[8]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 wartime mobilization of scientists and business was not taking full advantage of the nation's resources beyond elite institutions and large big businesses (Maddox 1981).

[9]As we will show below, 64 percent of OSRD funding went to just 10 institutions, and nearly 50 percent to MIT, Harvard, and CalTech alone – all universities represented by OSRD leadership.

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

[11]"[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."

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

been transferred to mission agencies (the Atomic Energy Commission, the National Institutes of Health, the Office of Naval Research), precluding the single-agency approach Bush (and Kilgore) had envisioned. Though the NSF was in large part "a triumph for Bush" (Kevles 1977) – primarily focused on basic research, administered by scientists – its budget was small.[13]

While each of the other major postwar R&D funding agencies had their own rules and procedures (regarding patent rights, allocation mechanisms, indirect cost recovery, etc.), a striking feature of federal research funding in the decades that followed was its continued geographic and institutional concentration. By and large, most agencies used merit review throughout the 1950s and 1960s (Graham et al. 1997), which may have contributed to continued concentration of federal resources. Beginning in the 1960s, a variety of legislative initiatives and programs (e.g., the 1979 EPSCOR program) attempted to widen the geographic distribution of funding, channeling Kilgore's criticism of the OSRD and concerns about extending the OSRD model in peacetime. Opponents of these programs typically argue that funding should be directed to the best and brightest researchers, as determined by the scientific community, echoing Vannevar Bush.

1.2 Previous Research

Our work is informed by a large body of existing scholarship on the OSRD and postwar U.S. science. Much of this work focuses on specific research projects supported by the OSRD and its role in the war effort (e.g., Rhodes 2012, Swann 1988, Bud 2007, Rasmussen 2002a, Hoyt 2006), its effects on U.S. universities (Geiger 1993), its organizational features (e.g., patent clauses, indirect cost rates, distribution of funding, who should govern science), and the politics of the OSRD and links to later debates in postwar science policy (Hart 1998, Kevles 1977, Kleinman 1995). Some of this historical work provides quantitative evidence on the distribution of contracts across time, geography, and field (Owens 1994), which we build on in our analyses below. While many view the war as a sharp break from the previous status quo, there is also an influential body of scholarship arguing that there were important continuities between the war effort and prewar developments in science and technology (Lowen 1991, Geiger 1993, Owens 1994, Rasmussen 2002b).

Our work is also influenced by decades of empirical research in economics assessing the impact of public sector investments on private sector research and innovation. In making the case for an expanded federal role in R&D funding, the Bush Report asserted that advances in science drive technological innovation. Economists studying science and technology have since attempted to test

[13] Geiger (1993) notes that though both Bush and Kilgore recommended a single federal research funding agency, and nearly all witnesses on legislation on postwar science policy agreed, at the end of the day "instead of dominating the federal science matrix" the NSF "would inherit the remaining unoccupied spaces."

this hypothesis and quantify the magnitudes. While details vary, this research typically relates variation (e.g., over time, space, fields) in research funding to outcomes. A perennial challenge is endogeneity: public sector investments may themselves reflect changes in scientific and technological opportunities or demand, making it difficult to causally attribute changes in outputs and outcomes to changes in levels or direction of public sector investment (Jaffe 2008). Many recent papers use quasi-experimental variation to attempt to overcome this problem (e.g., Jacob and Lefgren (2011), Azoulay et al. (2019), Moretti et al. (2019)), but the problem remains challenging. While absent a true random experiment we cannot fully circumvent this problem, one advantage to studying World War II is that it was a large, mission-driven technology shock, and specific military demands (together with considerations of scientific feasibility) dictated funding decisions rather than pure scientific value.[14]

In studying the effects of the OSRD on the geography of innovation, our work connects back to the Bush-Kilgore debates, scholarship on postwar concentration of research activity (Graham et al. 1997), and modern discussion of place-based policies for innovation (Gruber and Johnson 2019, Glaeser and Hausman 2020). Our work is also connected to the empirical literature on knowledge spillovers and the extent to which economic benefits of publicly funded research are geographically localized (e.g., Jaffe et al. 1993). Unlike the majority of previous research in economics on this topic, our historical setting allows us to observe agglomeration and localized spillovers over a long timeline, and to explore potential mechanisms behind our results.

Finally, this paper links to scholarship on war, innovation, and economic growth. Ruttan (2006) argues that military procurement has been a major driver of innovation in a range of industries, and particularly important for development of general purpose technologies. Other economic historians have taken the opposite view, arguing that war has a modest or even negative impact on productivity and innovation (Mowery 2010, Field 2008). The two World Wars have also served as laboratories for specific questions about intellectual property rights and innovation, including the effects of compulsory invention secrecy (Gross 2019), the effects of compulsory licensing of foreign-owned invention (Moser and Voena 2012, Baten et al. 2017), and the effects of reduced access to foreign science on research output (Iaria et al. 2018).

[14] For example, Stewart (1948) noted of the OSRD Committee on Medical Research that "the shift in emphasis and even in direction was enormous," as many subjects "of minor importance in peacetime [e.g., tropical medicine or human physiology at high altitudes] became of controlling importance in war."

2 Data

While a complete empirical assessment of the effects of OSRD would have previously been difficult, we have created a new dataset of the universe of OSRD contracts, which were preserved at the National Archives (NARA) in College Park, MD. We have collected, transcribed, and harmonized a complete record of all 2254 OSRD contracts (to 461 distinct contractors), and the 7910 inventions, 2763 patents, and 2470 scientific publications which they produced. For each contract, we observe many features: the contractor (with locations), OSRD division which wrote the contract (indicating subject matter), total value and value obligated, security classification, patent clause (short or long form, giving the government versus the contractor the right to patent inventions developed under contract) and termination date. The OSRD records further identify 7879 OSRD-supported inventions reported by contractors, on which 3382 patents were filed and 2659 granted. We supplement these records with an automated, text-based search for continuations and divisions of these patent applications, which identifies another 104 OSRD-supported patents.

We link the data on OSRD contractors and patenting to the U.S. patent record. To do so, we separately collected data on all U.S. patents granted between 1920 and 1979, merging a USPTO master file of patents with patent number, patent class, and issue date (Marco et al. 2015) with data on (i) serial numbers and filing dates (FPO); (ii) the full network of front-page (historically, final-page) citations (FPO); (iii) harmonized assignee names (Derwent Innovation, supplemented by our own effort) and assignee types; and (iv) inventor locations (HistPat). We also identify all pre- and post-war patents of individuals listed on the rosters of two of the major OSRD-supported research labs (the Radiation Lab at MIT, and the Radio Research Lab at Harvard), to explore the role of individual inventors. In order to compare trends in U.S. and foreign patenting we also collected 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 the base layer in the USPTO data: patent number, patent class (IPC), and grant date. See Appendix A for complete details of the patent data preparation.

In Section 5, we measure county-level employment using the U.S. Census Bureau County Business Patterns (CBP), focusing on 2- and 3-digit industries related to technology areas which were a focus of OSRD research, especially Communications equipment (SIC 366), Electronic Components and Accessories (SIC 367), and Professional, Scientific, and Controlling Instruments (SIC 38), and also collecting manufacturing sector and aggregate employment. We use the 1947, 1959, 1970, and 1980 editions of the CBP to measure these outcomes at roughly 10-year intervals. The CBP data present a few unique challenges, including data suppression in small cells and (modest) changes to the SIC classification over time, which we address with imputation and reclassification, respectively. We discuss these issues and their resolution in depth in the appendix.

Distribution of OSRD funding across space and subject matter

The OSRD contracted for research in a wide range of subject areas, and with a broad array of contractors, including major firms and research laboratories like Bell Labs, General Electric, RCA, and DuPont and universities like Harvard, MIT, and CalTech, which hosted programmatic research labs to work on OSRD projects. Table 1 list the top 10 OSRD divisions, and Table 2 lists top OSRD industrial and university contractors, by the total value of their contracts.

[Tables 1 and 2 about here]

Although particular research efforts were concentrated in specific locations – like radar research at the MIT Radiation Lab and Harvard Radio Research Lab, or proximity fuze research at the Johns Hopkins Applied Physics Lab – OSRD-sponsored research projects spanned the country. Figure 1 maps locations in the continental U.S. with OSRD-supported patents filed between 1941 and 1948, demonstrating the geographic scope of the research effort.

[Figure 1 about here]

Table 3 reports the top 10 2-digit NBER categories with OSRD patents, sorted by the fraction of all patents filed between 1941 and 1948 in those classes which were OSRD patents. At the top of this list is nuclear energy. The next eight categories are broadly related to electrical, electronic, and communications equipment, including technologies such as radar, cryptography, and early computing, highlighting how much progress was being made on electronics research during the war, with applications far beyond warfighting – an observation which motivates our focus in Section 5 on the growth of industrial clusters in the electronics industry.

[Table 3 about here]

3 Direction of U.S. Invention after WWII

Given the scale of the World War II shock, our first question is whether the wartime research effort shifted the overall direction of U.S. invention in the aggregate. We take two approaches to answering this question: we first compare patent filing at the USPTO versus at foreign patent authorities in technology classes which were more- or less-heavily supported by the OSRD during the war. To do so, we aggregate to the 2-digit IPC level (131 classes in total) and compare patenting at the USPTO to that in the British and French patent offices. Using our data on the location of inventors on U.S. patents, we also aggregate USPTO filings to 2-digit patent categories (39 in total) and compare filings by domestic versus foreign inventors.

In this as well as later sections, our OSRD treatment measures is the fraction of U.S. patent applications in the given technology class filed from 1941 to 1948 which were OSRD-supported (henceforth, the "OSRD rate"). Under this definition, a given technology area is "shocked" if large fraction of its patents from this period resulted from OSRD research. We can measure the OSRD rate in levels, in logs (to account for skew), or in terms of its distribution (quantiles). Results throughout the paper are similar with any of these measures.

Here we take the latter approach: our first specification is effectively a triple-difference, comparing U.S. and foreign log patent filings, before and after war, in classes in different quartiles of the OSRD rate (calculated for classes with ≥1 OSRD patent) and an omitted category for classes with no OSRD patents. This specification is estimated over a sample of country-IPC-years from 1930 to 1970, omitting the years 1940 to 1945 (inclusive), as follows:

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

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.[15]

Figure 2 plots the estimates (with 95% confidence intervals), first for a comparison against Great Britain (left panel), and then against France (right panel). U.S. filings in the most heavily-treated classes are nearly 40 percent higher after the war than in Britain, and 60 percent higher than in France, both highly significant. The effects for the third quartile of treated classes are attenuated (roughly 25 percent and 40 percent, respectively), but still large and statistically significant, and in lower quartiles near zero and only marginally significant.

[Figure 2 about here]

[15] 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.

We also estimate a variant of Equation 1 with annual coefficients, which allows us to explore both pre- and post-treatment trends. Figure 3 presents the estimates for differences in patenting in the top quartile of treated classes over time, following the same format as the previous figure. The figure reveals no particularly notable differences between the U.S. and Great Britain prior to the war, a spike of a difference during the war, a post-war dip, and then a take-off. The pattern is similar against France, with differences even more pronounced. We will see this same empirical pattern in several other contexts throughout the rest of the paper.

[Figure 3 about here]

In Appendix Figure B.1, we repeat this estimation for lower treatment quartiles, with similar but attenuated effects. We also re-estimate the regressions in Figures 2 and 3 excluding weapons-related classes such as firearms, ordnance, and explosives (Appendix Figures B.2 and B.3, respectively), to confirm that the results aren't mechanically driven by the World War II and Cold War military buildups – in both cases finding quantitatively similar results.

We then re-estimate these specifications on a distinct sample, using U.S. patents filed by domestic and foreign inventors in 2-digit NBER patent categories with a higher or lower OSRD rate in the 1940s. Appendix Figures B.4 and B.5 provide counterpart results to the figures above (binning two untreated patent categories into the lowest treatment quartile). We find qualitatively similar effects: pre-war patenting by U.S. inventors in the most-heavily treated classes shows no particular differential trends relative to patenting by foreign inventors, but it jumps after the war – with no such differences in untreated classes. These results are similar when estimated for log patents (as specified), citation-weighted patents, or inverse hyperbolic sine patents, and remain similar at the level of USPC patent classes (>400) rather than NBER patent categories.

4 Growth of Domestic Technology Clusters

The shift in the overall direction of U.S. inventive activity is just the tip of the iceberg in documenting the effects that the war had on American science and technology. As Section 1 recounted, the OSRD experience not only deepened the country's military-scientific capabilities, but perhaps even more importantly, it laid a foundation for post-war science policy and scientific institutions. Our attention now shifts to understanding whether, and through what channels, the wartime research effort altered the country's technological landscape – quite literally, what was being invented, where, and by whom – and how broadly and quickly it changed.

Massachusetts turns out to be an instructive microcosm for this question. Figure 4 shows the time series of patents filed across the 12 largest Massachusetts counties from 1935 to 1965. Middlesex County – home to Harvard and MIT – was more inventive in levels but not on a noticeably different time trend relative to other counties prior to WW2. It then experienced a spike in patenting around the war, much of which was resulted from OSRD-funded radar research. After the war, patenting returned to near pre-war levels, and then subsequently took off.

[Figure 4 about here]

We set out to perform these comparisons in a more systematic way across the country. To do so, we compare patenting over time in counties and technology areas with higher versus lower levels of OSRD investment, on the grounds that technology clusters are often technologically specialized, and it was not entire local economies per se which were shocked, but specific research areas in these particular places, which we measure as counties crossed by 2-digit NBER patent categories. Our baseline estimating equation throughout this section is as follows:

$$Ln(Patents)_{ict} = \sum_{t=1931}^{1970} \beta_t \cdot Ln(OSRD\ Rate)_{ic} \cdot Year_t + \alpha_{ic} + \delta_t + \varepsilon_{ict} \tag{2}$$

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. Our preferred treatment measure is the county- and category-level OSRD rate (the fraction of patents filed in a county and patent category between 1941 and 1948 which were OSRD patents). In our primary specification we use a continuous measure of the logged OSRD rate, which mechanically restricts the sample to county-categories with at least one OSRD patent. In robustness checks we also re-examine our results with a specification with treatment quartiles, similar to that in Section 3, which allows us to compare segments of the distribution in a flexible way:[16]

$$Ln(Patents)_{ict} = \sum_{q=1}^{4} \sum_{t=1931}^{1970} \beta_{qt} \cdot \mathbb{1}(\text{Treatment quartile } q) \cdot Year_t + \alpha_{ic} + \delta_t + \varepsilon_{ict} \tag{3}$$

An inevitable concern is the endogeneity of the location and subject matter of OSRD research, since contractors and contracts were not randomly chosen. Indeed, in the aggregate we see that patent classes with the highest rate of OSRD patenting were growing quickly before the war and continued

[16] Note that although we have experimented with treatment intensity measures derived from OSRD contracts and obligations, there is no straightforward way to allocate these to patent classes, especially since many contracts yielded either no patents or patents in multiple classes.

growing quickly after the war. However, as we will see below, when the analysis is disaggregated to the level of county-classes, these pre-trends disappear: much like in Figure 4, it does not appear that county-classes with high levels of OSRD treatment were on different trends before the war, relative to those with low OSRD treatment or none at all.

4.1 Baseline effects

Figure 5 plots coefficients from Equation 2, shown with 95% confidence intervals. As the figure indicates, patenting in the most heavily-treated county-categories: (i) was not growing statistically differently than the untreated group through 1941, (ii) experienced a relative surge during the war, (iii) contracted from its mid-war peak when the war ended, and then (iv) took off on a path of decades of rapid, sustained growth (indeed, the underlying growth pattern is exponential). The magnitudes indicate that county-categories with double the wartime OSRD rate (e.g., from 50% to 100%) were producing 30% more patents in that technology area by 1970.

[Figure 5 about here]

Appendix C provides several supporting results. Appendix Figure C.1 estimates effects by treatment quartile (similar to Equation 3), where we see similar patterns for the top quartile of treated county-categories, attenuated effects for the third quartile, and no such effects at lower quartiles. Appendix Figure C.2 reproduces Figure 5 excluding, in sequence, California, Massachusetts, New Jersey, and New York – states which either had a concentration of OSRD patents in the 1940s or are now the locations of the largest U.S. technology hubs – to demonstrate that our baseline results are general, and not being driven by any particular region of the country.

These results raise the question of whether the rise in patenting reflects population movements (agglomeration) or underlying growth in inventive productivity (per capita). To answer this question, we collect decadal county population from the U.S. population census, interpolate between decades, and measure annual per-capita patenting. Appendix Figure C.3 estimates Equation 2 for county-categories' per-capita patenting. The pattern in Figure 5 persists, though the magnitudes are attenuated, with a doubling of the OSRD rate being associated with around a 10% increase in per-capita patenting by 1970. The OSRD shock thus appears to have increased local invention via a combination of *both* agglomeration and productivity effects.

We also confirm that these results are similar whether estimated for log patents (as specified), citation-weighted patents, or inverse hyperbolic sine patents (which can be calculated for county-categories with zero patents, which then remain in the sample). When estimated for exclusively non-OSRD patenting, the spike in the mid-1940s is smoothed out (this spike represents the OSRD patents themselves), but the subsequent take-off in patenting remains.

4.2 Heterogeneity

These patterns show up in numerous other views of the data, including when comparisons are made within county-years or category-years (versus the baseline comparison, within county-categories over time). But what these average effects invariably mask is the potential for substantial heterogeneity across space, across time, and even across technology areas.

The most striking implication of the results thus far is that World War II was the singular event setting in motion increasing agglomeration of inventive activity around the country, and ostensibly the take-off of technology clusters that persist to this day. At first view, this evidence would seem to contradict historians' continuity hypothesis (Geiger 1993, Owens 1990, Lowen 1991), which posits continuity between pre-war, wartime, and post-war U.S. science and technology (Section 1). In the context of the local or regional agglomeration of invention, the question is whether the war was an equalizing force, or merely deepened existing geographic differences.

To more explore this question, we partition county-categories into the top 5% versus bottom 95% of 1930s patenting intensity. When Equation 2 is estimated for each group, it becomes apparent that the effects are entirely driven by clusters which were already among the most inventive before World War II (Figure 6). 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 take-off. The evidence thus supports an interpretation of both continuity and change, like that seen in our Massachusetts example (in Figure 4): pre-war differences persisted, but the war caused a trend shift. In simpler terms, the OSRD's effect was to catalyze massive long-run growth in existing geographic centers of invention.

A second question is whether the OSRD effect was general across all technologies whose development it funded, or stronger for some research fields over others. We evaluate this question by partitioning the sample by 1-digit NBER categories, which represent six broader technology areas: Chemicals, Computers & Communications, Drugs & Medical, Electrical & Electronics, Mechanical, and Other. We re-estimate Equation 2 for each of these subsamples, with results in Figure 7. The OSRD's prioritization of radar, computing, and electronic controls made Electrical & Electronics one of the most active areas of OSRD patenting during the war. This was also true of weapons systems, which fall in the Mechanical category. Perhaps as a result, Figure 7 suggests that the OSRD effect is driven by precisely these two categories. The fact that the research effort played a large role in the history of the U.S. electronics industry, and an empirically important role in the take-off of U.S. electronics invention, motivates our focus on the electronics industry in the next section, when we study the effects of the OSRD on employment growth and the real economy.

In additional tests, we compare the effects across different categories of assignees, such as patenting by firms vs. individuals, and for then for patenting by incumbents vs. entrants, which we define simply as patents whose assignee does versus does not have a prior USPTO filing (back to 1920). We find large effects for all groups, indicating that the OSRD effect cuts across a wide swath of filers in the treated clusters. The question, then, is why.

Mechanisms: An ongoing investigation

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 consequences reverberated for decades. Why did a temporary shock end up having such long-lived effects? What is it about this funding shock that led to it transforming the U.S. inventive sector in the latter half of the century? And what does it teach us about how these effects can be reproduced today?

A complete answer to this question could be a golden ticket to decades of technological progress, and is beyond the scope of any single paper alone. In ongoing work we are beginning to explore the forces which may have turned this temporary shock into a long-lived one, especially by understanding how the wartime research effort shaped, fed, or interacted with other elements of the U.S. innovation system. The evidence from this effort is more likely to be hypothesis-generating than definitive, but it can bring into focus some of the postwar developments in U.S. science and technology that might have cemented the durability of the OSRD's efforts.

Candidate explanations might include direct effects of OSRD-funded research such as follow-on invention, or less direct ones such as the growth of local innovation ecosystems. We can already rule out the direct channels: in our ongoing work, we have found no evidence that the long-run effects were powered by postwar patenting by the inventors or firms involved in OSRD research (neither OSRD contractors, nor assignees on OSRD patents), or even postwar invention directly citing OSRD patents. This leads us to a tentative conclusion that the effect of the OSRD was to cultivate a local ecosystem supporting innovation in treated technology areas. One channel that may have supported or accelerated this effect was the local university infrastructure, both as a locus of research activity and as training grounds for scientific talent. Another candidate is rapid growth in the postwar network of federally-funded R&D centers (FFRDCs), which were borne out of the major OSRD-supported research labs and often had formal links to local firms and universities;

yet another is defense R&D spending during the Cold War, which engaged contractors around the country in technology development. The propagation of the OSRD shock may not be attributable to any single one of these mechanisms – it may be the whole system, rather than any one of its parts, that explains the remarkable persistence of the OSRD effect. In other words, the answer to "what changed?" might not be any one thing, but rather many things.

Several data sources are available in our effort to understand this persistence. As an initial step, we plan to expand our efforts to measure intellectual linkages to OSRD-supported research with second or third generation patent citations, and university spillovers with patent-paper citation linkages, including in-text citations (Bryan et al. 2020). Typically these types of effects are hard to observe for policy experiments, but our rich data and long observation window may allow us to do so. We also collect data on the U.S. university system from a National Academy of Science report (NAS 1963), which provides PhD production by university, year and field. We likewise measure FFRDCs from the NSF's annual Federal Funds for Research & Development report from 1958 to 1977, which lists FFRDCs, their funding agency, their administrating organization, and their location, and we supplement these data with manual research on each lab's primary subject area.

It is also possible that the long-term effects we are observing reflect OSRD building up capacity (e.g., scientific and technological capabilities, managerial knowledge, grantsmanship) in certain areas, which in turn helped institutions compete for Cold War grants and contracts, and there are localized spillovers from the latter – in other words, that the OSRD initiated a virtuous cycle of federal funding in the treated areas. Since most Cold War funding was defense related, we will start this inquiry by looking at local Department of Defense R&D contracting, which we have collected back to the 1960s. We also are collating data on total federal support for R&D by institution from the earliest NSF publications reporting these data (published in the 1960s).

One yet-unaddressed mechanism is migration, especially if certain regions became attractive for scientists and inventors working in a particular field. The geography of innovation literature (Feldman and Kogler 2010) points to cross-fertilization of ideas, thick markets for intermediate inputs, and the rise of supporting institutions (e.g. venture capital) as potential reasons why scientific and technical talent may be attracted to a given regions. Although difficult to measure, we may be able to use the postwar National Roster of Scientific and Technical Personnel (NRSTP), a biennial list of U.S. scientists, to study migration patterns. An early version of this list was introduced during World War II, reflecting Kilgorian concerns that not all relevant researchers were being mobilized through the OSRD model, but the available data span 1954 to 1970. The NRSTP includes scientists' location, field, degrees, and other information which we might be able to use to trace the mobility of scientists and inventors over the postwar era.

5 Effects on Local Employment in High-Tech Industries

The collective evidence thus far has established that World War II was a sea-change event for the U.S. inventive economy, and especially for inventive geography, but a remaining question is whether these effects spilled over into the broader population. What are the real effects of a big-push investment in science and technology like the one undertaken by the U.S. government in World War II? What happens to population growth, employment, wages, land prices, and other economic outcomes? These questions sit at the heart of contemporary debates around place-based policies to jump-start local innovation and increase both the growth rates and resilience of local economies (e.g., Gruber and Johnson 2019, Glaeser and Hausman 2020).

To study this question, we need to a data series on local outcomes that is available over a long horizon and will vary at the level of the OSRD treatment itself, namely counties crossed by subject areas. For this exercise, we use CBP data from 1959, 1970, and 1980 to measure local employment in high-tech industries. We relate the OSRD's wartime R&D investments to future employment, in a test of whether its effects extended beyond R&D and patenting by shaping population centers and the economic geography of high-tech manufacturing.[17]

Our focus here is on the postwar U.S. electronics industry. Not only was electronics one of the technology categories that appears to have embodied the OSRD's effect on inventive agglomeration (Figure 7), but given (i) the concentration of OSRD research in predecessor fields like microwave radiation and electrical communication, (ii) our ability to draw relatively clean links between these industries and related patent classes, and ultimately (iii) the rapid growth of these industries and their importance to U.S. technological progress and global technological supremacy in the second half of the 20th century, these industries are most ripe for study.

We focus on SIC 366 and 367, *Communications Equipment* and *Electronic Components and Accessories* manufacturing. Appendix A lists examples of 4-digit industries and products included in these SICs. We combine these SICs and crosswalk patent classes to them using a USPC-SIC concordance from the USPTO, which we in turn use to measure (i) total wartime patents and (ii) OSRD patents for each county and industry – in other words, an industry-level measure of the local OSRD shock. From the CBP data we measure employment in SIC 366-367, as well as total manufacturing sector employment and total employment.[18]

[17]Pre-1970 CBP data are unavailable in electronic format and require digitization from Census publications. The 1959 CBP is the earliest edition available which provides county-level employment at the 3-digit SIC level, which is sufficiently disaggregated to serve our purposes, and which we can crosswalk to patent classes using the USPTO's USPC-SIC crosswalk. As Section 2 explains, because digitizing the entirety of the 1959 CBP is prohibitively expensive, we selected SICs of interest to be digitized for all U.S. counties.

[18]Because a few counties' employment for these industries is suppressed, we have also used county-industry data on

We regress county-level log employment in SIC 366-367 in 1959, 1970, and 1980 on the log number of technologically-related OSRD patents during the 1941-1948 period. In all specifications, we control for the log number of total patents (OSRD and non-OSRD) during this period to account for overall inventive productivity. In successive specifications we also add controls for a county's log manufacturing employment (across all manufacturing SICs) and log total employment (across all SICs). Formally, we estimate the following regression:

$$
\begin{aligned}
Ln(Employment)_{it} = & \sum_t \beta_t \cdot Ln(OSRD\ Patents)_i \cdot Year_t \\
& + \sum_t \gamma_t \cdot Ln(All\ Patents)_i \cdot Year_t + \delta_t + X_{it}\phi + \varepsilon_{it}
\end{aligned}
\tag{4}
$$

where i indexes counties and t indexes years, δ_t represent year fixed effects, X_{it} is a vector of controls, and standard errors are clustered at the county level. Note that as specified, the regression sample will be restricted to county-years with nonzero employment in SIC 366-367 and at least one associated OSRD patent, such that their logged values are defined.

Table 4 presents the results. Column (1) shows our baseline specification without controls, where we see a qualitative, increasing relationship between local OSRD patenting and SIC 366-367 employment over time, with effects marginally significant by 1980. Once we control for total manufacturing employment in Columns (2) and (3), a pattern more clearly emerges: counties with more OSRD patents in the 1940s in electronics-related classes had higher electronics industry employment in the postwar period, with the magnitudes suggesting a doubling of OSRD patents being associated with roughly 60-65% higher employment in 1970 and 1980. In Column (4), we replace the dependent variable with the fraction of county employment in SIC 366-367, where we see qualitatively similar effects: a doubling of related OSRD patents in the 1940s is associated with a 0.25 p.p. increase in local SIC 366-367 employment share by 1970 and 0.4 p.p. by 1980.

[Table 4 about here]

Given that these effects are largest in columns controlling for total employment, our interpretation of this evidence is that counties where the wartime research was performed were subsequently more likely to be loci of electronics and communications manufacturing, and in turn more specialized

the firm size distribution to impute employment levels, regressing observed employment on the firm size distribution for county-industries reporting employment, and using the resulting parameter estimates to generated "predicted" employment levels, which we can use in robustness checks to expand the sample. Due to a preference for precise measurement and unknown risks of bias in the imputation method, we currently are presenting results using only the recorded data, although results based on imputed employment levels are similar.

in these high-tech industries versus others. The upshot, however, is that the seeds planted by the wartime research effort appears to have also sprouted local industrial activity, spilling over from the inventive sector to the local economy more broadly.

6 Implications and Conclusion

As the war drew to its close, a flurry of media coverage revealed the scientific research which had taken place behind the scenes, and the significant role it played in the Allied victory. Yet as the OSRD's Deputy Director Irvin Stewart wrote in 1948, anticipating its transformative effects, the "full impact" of this effort may not yet have been realized (Stewart 1948). Despite a large historical and science policy literature on the long-run effects of OSRD on the institutions of postwar science policy, we believe this paper to be first quantitative empirical assessment of the long-run effects of this "unique experiment" on innovation and employment outcomes.

Using a new dataset on all OSRD contracts, we show evidence of the long run effects of this temporary funding shock on the regions and technology areas where OSRD funding was concentrated. If interpreted as causal, these results support Vannevar Bush's argument that federally-funded research can lead to persistent growth in innovation and economic performance. Here it is important to recognize that although Bush argued for funding "basic" research in *Science: The Endless Frontier*, asserting that the wartime effort built on prior advances in fundamental science, the OSRD's funding was primarily for applied research and development. Our results thus speak most directly to the effects of large-scale federal investments in applied research.[19]

The results add to several long-running policy and academic debates. On the concentration of federal R&D spending, Kilgore was right: much of the OSRD support was directed to researchers at elite institutions and research labs (Table 2). While it is hard to know whether the elite funding model of OSRD was the most efficient one for wartime (i.e., were there, literally, any lost Einsteins?), it is also hard to argue with the results. Nonetheless, 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. 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.

In addition to efficiency-equity tradeoffs, another theme in postwar science and technology policy relates to accountability of federally funded researchers. The Bush Report is often read as a plea for

[19] Similarly, it is interesting that the Bush report emphasized scientific self-governance, but OSRD was decidedly driven not by scientific curiosity but instead by battlefield demands.

limited direct accountability of scientists to policymakers, as part of the postwar "social contract" for science. Calls for greater accountability and performance evaluation, often in times with tight budgets, ruffle scientists' feathers (Cozzens 1999). Our results suggest the need for long evaluation time horizons. Were the effects of OSRD research on non-OSRD patenting evaluated in 1950 as opposed to 1970, we would miss a large part of its economic impact.

Our paper also connects to the literature on the economics of World War II, and on the economic effects of war in general. There is a debate in the literature about whether war is good or bad for economic performance, and whether defense R&D generates economic spillovers (e.g., Moretti et al. 2019). Our results suggest that there is a long-run effect of defense R&D on patenting and economic performance, even for a research program that was explicitly not focused on creating "dual use" technologies. It is difficult to perform a rate of return calculation, however, not only because valuing patent outcomes is difficult, but also because without better understanding of mechanisms generating the long-run effects (including potential additional funding from other agencies and actors) it is unclear we are measuring all the relevant inputs.

We previously referenced a large historical literature describing the wartime effort as a sharp break from the past, and a countervailing school arguing that wartime research was in fact a continuation of preceding technological developments. Our results lend support to both hypotheses: the main effects of OSRD funding are seen in county-classes which were already more heavily patenting before the war – but they were not on differential trends. The wartime effort appears to have magnified these differences, setting in motion persistent agglomeration effects.

Our work is preceded by a large literature in the economics of science which estimates the impact of marginal changes to research funding on patenting, generally finding positive effects. Much of this evidence is based on grants for scientific research, such as at the NIH, which – reflecting Vannevar Bush's vision in *Science: The Endless Frontier* – currently comprise the majority of the federal government's (civilian) extramural research support. Our work complements this body of evidence, suggesting that the major applied research effort undertaken in World War II had large, long-run effects on the U.S. national innovation system.

References

Akcigit, Ufuk, John Grigsby, and Tom Nicholas. 2017. *The Rise of American Ingenuity: Innovation and Inventors of the Golden Age*. NBER Working Paper No. 23047.

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

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 University Press, pp. 609–626.

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

Baten, Joerg, Nicola Bianchi, and Petra Moser. 2017. "Compulsory licensing and innovation: Historical evidence from German patents after WWI," *Journal of Development Economics*, Vol. 126, pp. 231–242.

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

Bryan, Kevin A, Yasin Ozcan, and Bhaven Sampat. 2020. "In-text patent citations: A user's guide," *Research Policy*, Vol. 49, No. 4, p. 103946.

Bud, Robert. 2007. *Penicillin: Triumph and tragedy*. Oxford: Oxford University Press.

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

Cozzens, Susan E. 1999. "Are new accountability rules bad for science?" *Issues in Science and Technology*, Vol. 15, No. 4, pp. 59–66.

Feldman, Maryann P and Dieter F Kogler. 2010. "Stylized facts in the geography of innovation," in *Handbook of the Economics of Innovation*, Vol. 1: Elsevier, pp. 381–410.

Field, Alexander J. 2008. "The impact of the Second World War on US productivity growth," *The Economic History Review*, Vol. 61, No. 3, pp. 672–694.

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.

Graham, Hugh Davis, Nancy Diamond et al. 1997. *The rise of American research universities: Elites and challengers in the postwar era.*: JHU Press.

Gross, Daniel P. 2019. *The consequences of invention secrecy: Evidence from the USPTO Patent Secrecy Program in World War II*. Working paper.

Gross, Daniel P. and Bhaven N. Sampat. 2020. *Mission-oriented research in a National Emergency: Lessons from the Committee on Medical Research in World War II*. Working paper.

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

Hart, David M. 1998. *Forged consensus: Science, technology, and economic policy in the United States, 1921–1953*. Princeton: Princeton University Press.

Hoyt, Kendall. 2006. "Vaccine innovation: Lessons from World War II," *Journal of Public Health Policy*, Vol. 27, No. 1, pp. 38–57.

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.

Jacob, Brian A. and Lars Lefgren. 2011. "The impact of research grant funding on scientific productivity," *Journal of Public Economics*, Vol. 95, No. 9–10, pp. 1168–1177.

Jaffe, Adam B. 2008. "The 'Science of Science Policy': Reflections on the important questions and the challenges they present," *The Journal of Technology Transfer*, Vol. 33, No. 2, pp. 131–139.

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.

Kevles, Daniel J. 1977. "The National Science Foundation and the debate over postwar research policy, 1942–1945: A political interpretation of Science–The Endless Frontier," *Isis*, Vol. 68, No. 1, pp. 5–26.

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

Lowen, Rebecca S. 1991. "Transforming the university: Administrators, physicists, and industrial and federal patronage at Stanford, 1935–49," *History of Education Quarterly*, Vol. 31, No. 3, pp. 365–388.

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

Moretti, Enrico, Claudia Steinwender, and John Van Reenen. 2019. *The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers*. NBER Working Paper No. 26483.

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.

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

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.

Owens, Larry. 1990. "MIT and the Federal 'Angel': Academic R & D and Federal-private cooperation before World War II," *Isis*, Vol. 81, No. 2, pp. 188–213.

——— . 1994. "The counterproductive management of science in the Second World War: Vannevar Bush and the Office of Scientific Research and Development," *Business History Review*, Vol. 68, No. 4, pp. 515–576.

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.

Rasmussen, Nicolas. 2002a. "Of 'Small Men', Big Science and Bigger Business: The Second World War and biomedical research in the United States," *Minerva*, Vol. 40, No. 2, pp. 115–146.

——— . 2002b. "Steroids in arms: Science, government, industry, and the hormones of the adrenal cortex in the United States, 1930–1950," *Medical History*, Vol. 46, No. 3, pp. 299–324.

Rhodes, Richard. 2012. *The making of the atomic bomb*. New York: Simon and Schuster.

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

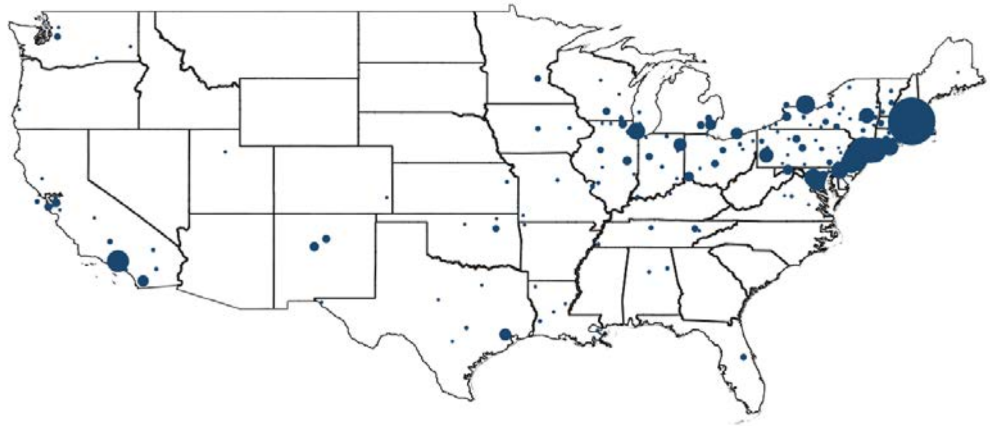
Ruttan, Vernon W. 2006. *Is war necessary for economic growth? Military procurement and technology development*. Oxford: Oxford University Press.

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.

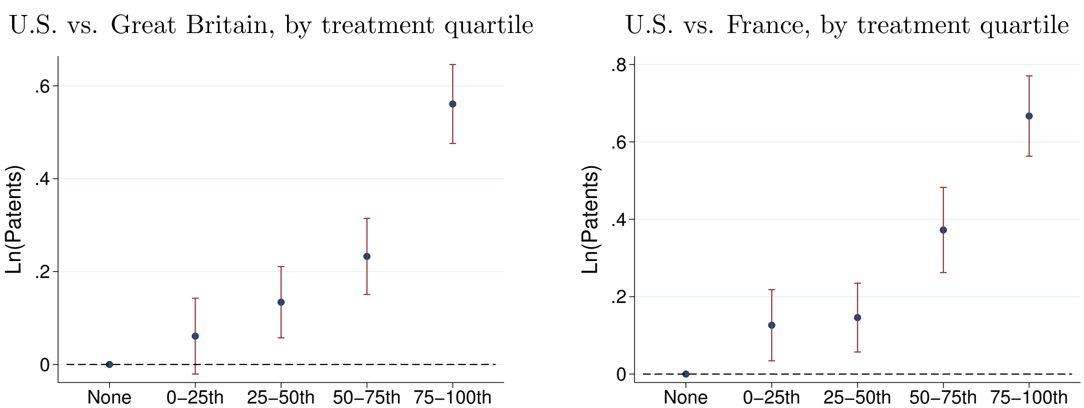
Swann, John Patrick. 1988. *Academic scientists and the pharmaceutical industry: Cooperative research in twentieth-century America*. Baltimore: Johns Hopkins University Press.

Figure 1: Number of OSRD Patents, by County



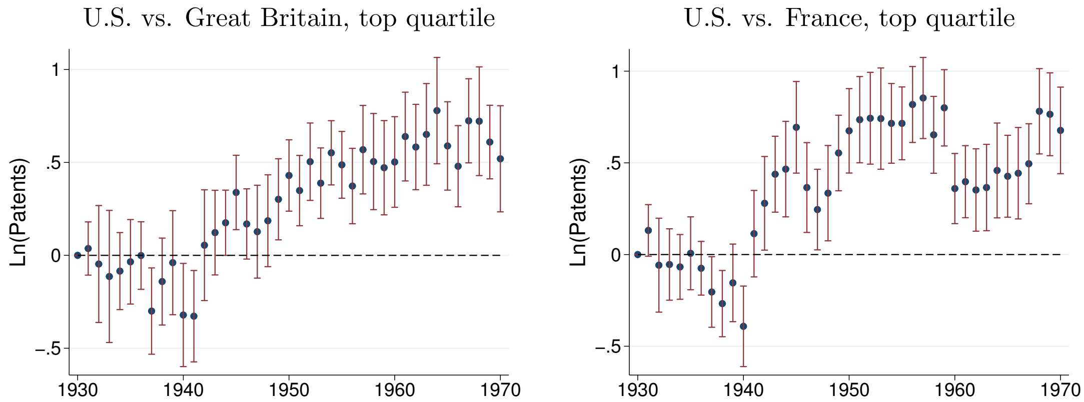
Notes: Figure maps counties with OSRD-supported patents. Bubble sizes proportional to each county's total number of OSRD patents.

Figure 2: Patenting at the USPTO versus select foreign patent authorities in each quartile of treated patent classes, difference-in-differences, pre-1940 vs. post-1945



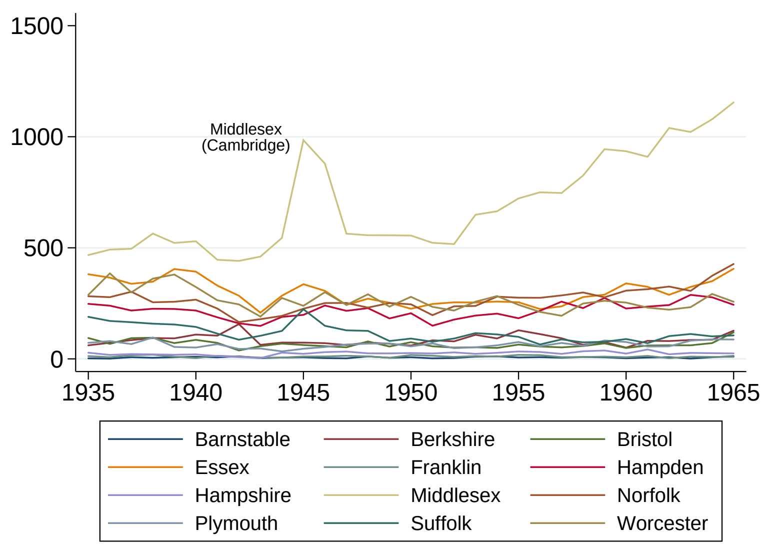
Notes: Figure shows difference-in-difference estimates of the effects of the OSRD shock on U.S. versus foreign patenting, in technology classes (2-digit IPCs) at different quartiles of OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-supported. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-class level.

Figure 3: Patenting at the USPTO versus select foreign patent authorities in the top quartile of treated patent classes (relative to untreated classes), annual estimates, 1930-1960



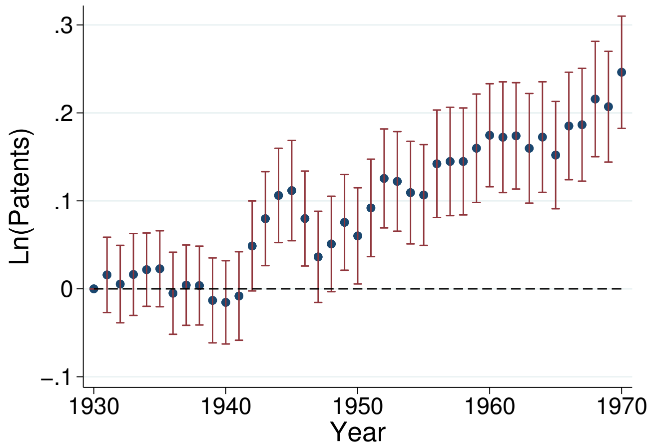
Notes: Figure shows annual difference-in-differences estimates of the effects of the OSRD shock on U.S. versus foreign patenting, in technology classes (2-digit IPCs) in the top quartile OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-supported, versus those without OSRD treatment. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-class level.

Figure 4: Annual patenting in top 12 Massachusetts counties, 1935 to 1965



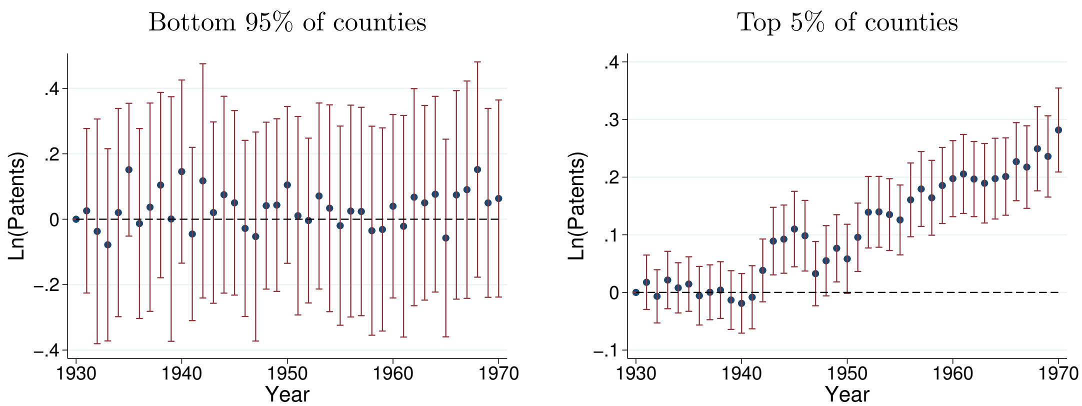
Notes: Figure shows total annual patents filed in the top 12 Massachusetts counties. The figure illustrates, for Middlesex County (location of Cambridge, home to Harvard and MIT): (i) relatively constant, pre-1940 level differences in patenting; (ii) a mid-1940s spike (doubling) of patenting, driven 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 5: Effect of the OSRD shock on patenting in county-categories, 1930-1970



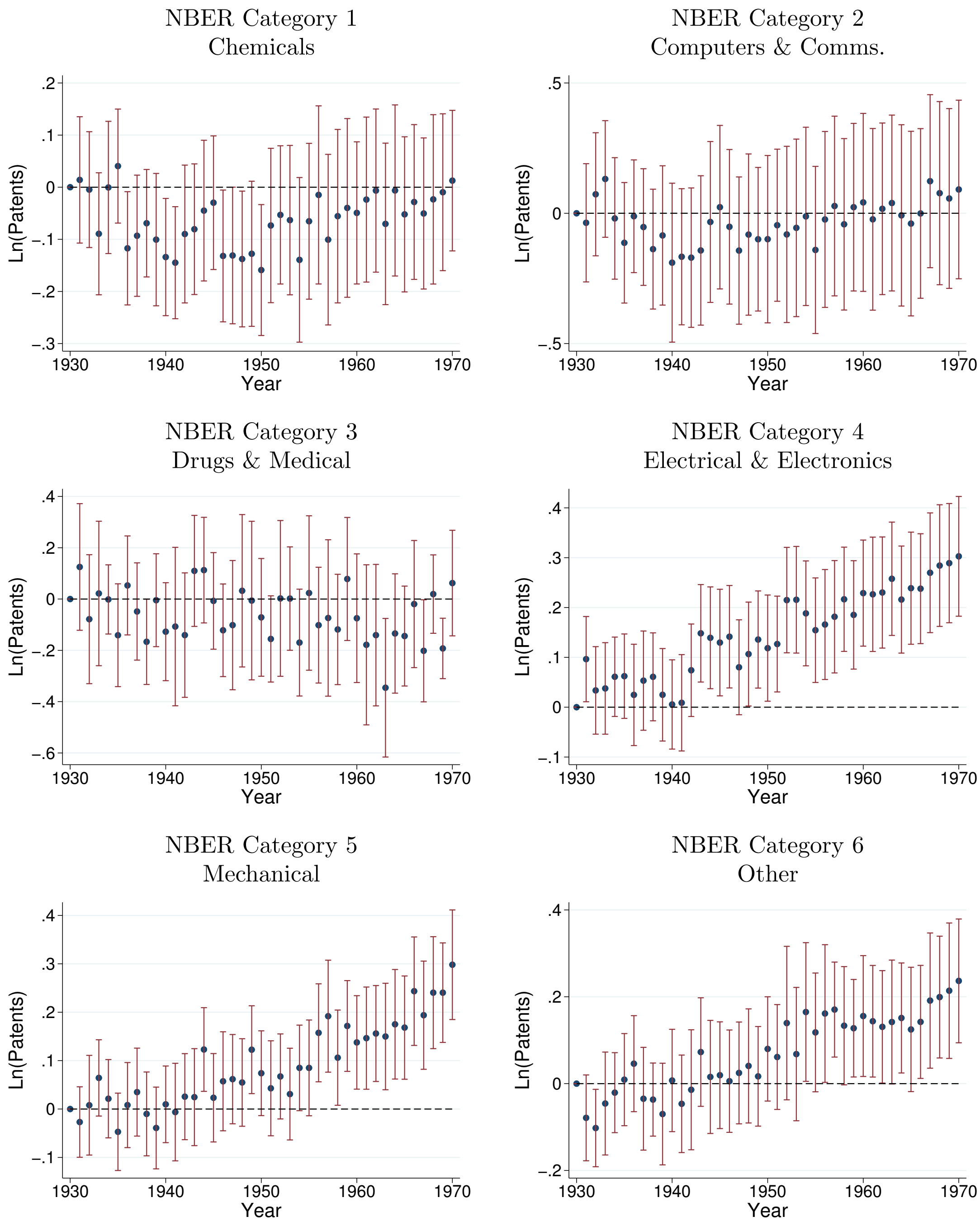
Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting. The independent variable measures the log fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-supported. Error bars represent 95% confidence intervals, computed from SEs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

Figure 6: Effects of the OSRD shock on local patenting, for county-categories in the bottom 95% versus top 5% of 1930s patenting (existing clusters), 1930-1970



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 7: Effects on local patenting by NBER top-level category, 1930-1970



Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting, by major technology category. Error bars represent 95% confidence intervals, computed from SEs clustered at the county-category level.

Table 1: Top 10 OSRD divisions, by contract obligations

Division	Description	Contracts	Total oblg. (mil.)	Pct. of oblg.
14	Radar	137	$129.6	28.0%
03	Rocket Ordnance	14	$82.9	17.9%
06	Subsurface Warfare	50	$47.2	10.2%
15	Radio Coordination	46	$27.3	5.9%
T	Proximity Fuze Program	63	$25.3	5.5%
CMR	Committee on Medical Research	572	$21.4	4.6%
S-1	Atomic Energy Program	100	$14.4	3.1%
04	Ordnance Accessories	43	$13.9	3.0%
11	Chemical Engineering	151	$11.7	2.5%
05	New Missiles	50	$11.7	2.5%

Notes: Table lists the top 10 OSRD divisions by total obligations. Percentages measure each division's percent of total OSRD research spending.

Table 2: Top OSRD contractors, by contract obligations

Top 10 Firms			Top 10 Universities		
Contractor	Total oblg.	Percent	Contractor	Total oblg.	Percent
Western Electric Co.	$15.2 mil.	3.3%	Massachusetts Inst. of Tech.	$106.8 mil.	23.1%
General Electric Co.	$7.6	1.6%	California Inst. of Tech.	$76.6	16.6%
Radio Corp. of America	$6.0	1.3%	Harvard University	$29.1	6.3%
E. I. Dupont De Nemours & Co.	$5.4	1.2%	Columbia University	$27.1	5.9%
Monsanto Chemical Co.	$4.5	1.0%	University of California	$14.6	3.2%
Eastman Kodak Co.	$4.3	0.9%	Johns Hopkins University	$10.8	2.3%
Zenith Radio Corp.	$4.2	0.9%	George Washington University	$6.9	1.5%
Westinghouse Elect. & Mfg. Co.	$3.9	0.8%	University of Chicago	$5.7	1.2%
Remington Rand, Inc.	$3.7	0.8%	Princeton University	$3.6	0.8%
Sylvania Electric Products, Inc.	$3.1	0.7%	University of Pennsylvania	$2.9	0.6%
Total	$57.81	12.5%	Total	$284.03	61.5%

Notes: Table lists the top 10 firms and universities with OSRD contracts by total obligations. Percentages measure each contractor's percent of total OSRD research spending.

Table 3: Top 10 2-digit NBER categories by OSRD rate, 1941-1948

Category	Description	# Patents from OSRD contracts	Pct. of patents from OSRD contracts, 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%
23	Computer peripherals	2	3.9%	12.5%
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%

Notes: Table lists the 10 2-digit NBER categories with the highest fraction of 1941-1948 patents produced under OSRD contracts, in descending order, as well as the number of patents and the maximal fraction of patents which were OSRD patents in any single year.

Table 4: Effects on post-war county employment in high-tech manufacturing industries

	Log employment in 366/367	Log employment in 366/367	Log employment in 366/367	Pct. of employment in 366/367
Ln(OSRD patents, 1941-1948) * Year==1959	0.014	0.428***	0.419**	0.269*
	(0.180)	(0.157)	(0.161)	(0.145)
Ln(OSRD patents, 1941-1948) * Year==1970	0.245	0.595**	0.578**	0.238**
	(0.227)	(0.234)	(0.235)	(0.099)
Ln(OSRD patents, 1941-1948) * Year==1980	0.353*	0.642***	0.638***	0.392**
	(0.196)	(0.184)	(0.186)	(0.161)
Ln(Patents, 1941-1948) * Year==1959	0.800***	-0.006	-0.014	0.059
	(0.254)	(0.252)	(0.257)	(0.053)
Ln(Patents, 1941-1948) * Year==1970	0.409	-0.149	-0.141	0.121**
	(0.293)	(0.286)	(0.289)	(0.050)
Ln(Patents, 1941-1948) * Year==1980	0.314*	-0.254	-0.270	-0.257
	(0.185)	(0.212)	(0.216)	(0.247)
Constant	3.447***	-1.883	-2.625	1.101
	(1.119)	(1.824)	(2.168)	(0.909)
N	122	122	122	231
R^2	0.31	0.42	0.43	0.15
Log mfg. empl.		X	X	
Log all empl.			X	

Notes: Table estimates the relationship between counties' post-war manufacturing employment in Communications Equipment & Electronic Components and Accessories (SIC 366 and 367, including electronic controls and semiconductors) and wartime OSRD patenting in classes which crosswalk to these industries (according to the USPTO's USPC-SIC crosswalk). County-level industrial employment measured from the 1959, 1970, and 1980 editions of the U.S. County Business Patterns (CBP). The outcome in the first three columns is log employment in SIC 366 and 367; in the fourth column, the fraction of total county employment in these SICs. All columns control for a given county's total wartime patenting, and successive columns add controls for log manufacturing employment and log total employment in the given CBP year. Sample restricted to counties with at least one OSRD patent in the associated classes (by construction). *, **, *** represent significance at the 0.1, 0.05, and 0.01 levels, respectively. SEs clustered by county in parentheses.

Web Appendix

A Data Appendix

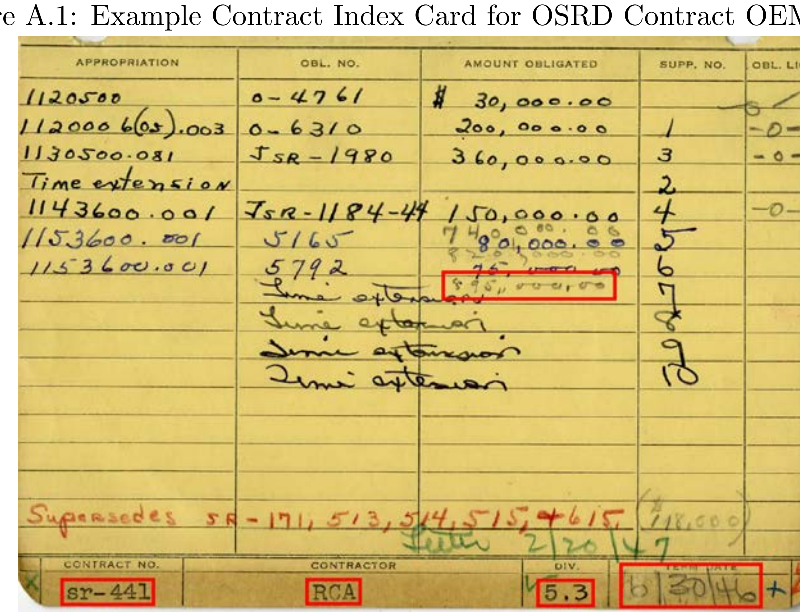
A.1 Data collected from OSRD archival records

A.1.1 OSRD contractors, contracts, and inventions

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

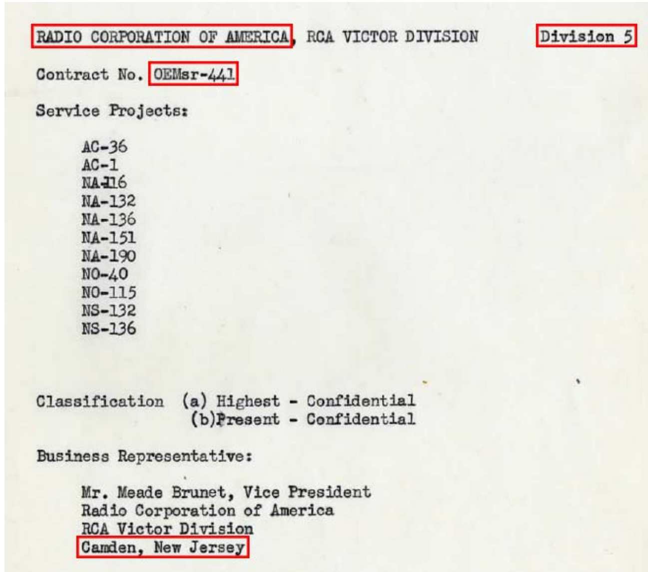


APPROPRIATION	OBL. NO.	AMOUNT OBLIGATED	SUPP. NO.	OBL. LIC
1120500	0-4761	$ 30,000.00		
1120006(05).003	0-6310	200,000.00	1	-0-
1130500.081	JSR-1980	360,000.00	3	-0-
Time extension			2	
1143600.001	JSR-1184-44	150,000.00	4	-0-
1153600.001	5165		5	
1153600.001	5792		6	

Supersedes SR-171, 513, 514, 515, & 615

CONTRACT NO. sr-441 | CONTRACTOR RCA | DIV. 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:

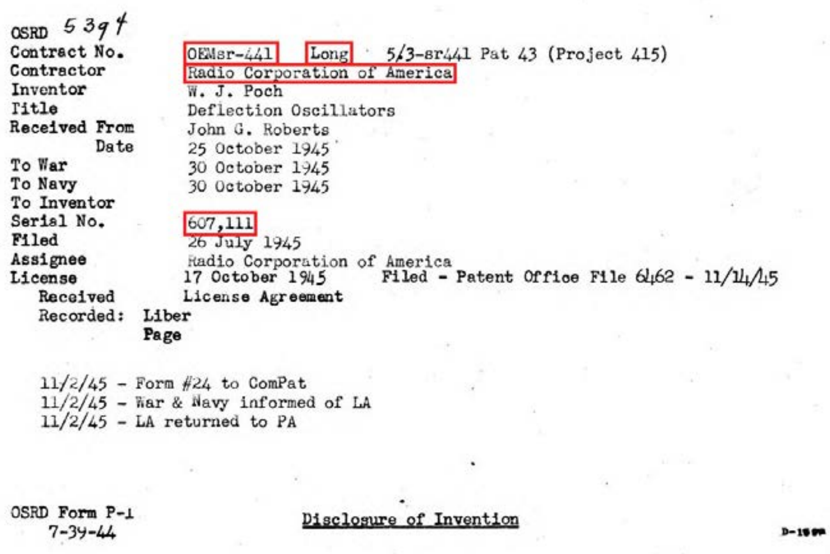
AC-36
AC-1
NA-116
NA-132
NA-136
NA-151
NA-190
NO-40
NO-115
NS-132
NS-136

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 5397

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

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.1.2 Medical research contracts and publications

For the 573 CMR contracts, we collect additional data from other contemporary publications and archival records. Although the CMR was a division of the OSRD, it itself was organized into six (sub)divisions, covering some of the following medical subjects:

1. *Division of Medicine.* Infectious diseases, venereal diseases, convalescence, psychiatry.
2. *Division of Surgery.* Wounds and burns, neurosurgery, other surgical specialties.
3. *Division of Aviation Medicine.* Oxygen deprivation, G-force stresses.
4. *Division of Physiology.* Blood, blood derivates, blood substitutes, shock, nutrition, acclimitization, other miscellaneous physiological studies.
5. *Division of Chemistry.* Treatment of gas casualties, and rodent and insect control.
6. *Miscellaneous.* Malaria, adrenocortical hormones (steroids), antibiotics.

We obtain the CMR division and medical subject of 571 of our 573 CMR contracts from Andrus (1948), which also provides one-sentence synopses for each contract. The OSRD archival records also contain ledgers with more detailed summaries of 567 these contracts: for each contract, the ledger reports the CMR division, contracting institution, principal investigator(s), project title, performance dates, contract budget, a one-page summary of results, and a list of academic publications and technical reports which were produced from this research.

In addition to these ledgers, the OSRD archival records include a collective bibliography of CMR publications, which appears to have been compiled near the end of the war. We construct a dataset of CMR publications as the union of these two data sources (which only partly intersect), with

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

the associated contracts, and link these publication records to PubMed (PM) and Web of Science (WOS) identifiers.[5] To connect publications to contracts, we use the CMR ledgers where possible (which explicitly link papers to contracts). For publications which were listed in the bibliography only (for which we have no direct measure of the associated contract), we take two approaches: (i) we manually look for other papers by the same author(s) on the same subject (based on paper titles) for which the contract is known, and assign these cases to that contract; and (ii) if the publication's contract is still unknown, we manually match authors and paper titles to contract PIs and project titles/descriptions. Through this procedure we are able to match all papers to CMR contracts. Between the CMR ledgers and the CMR bibliography, we identify 2,230 publications, which we matched to 1,715 papers in WOS and 1,146 papers in PM.

In the course of this data work, we noticed that these papers often acknowledge financial support from "the Committee on Medical Research" of the "Office of Scientific Research and Development" in a footnote. For completeness – especially to pick up papers which resulted from CMR-funded research but were published after the war (i.e., after our primary data sources were created) – we searched Google Scholar for these terms, and manually examined the results for new hits. Through this process we find an additional 210 CMR-supported publications, 202 of which we successfully link to CMR contracts. In total, we therefore have 2,440 CMR publications in our data (with 1,912 matched to Web of Science, and 1,315 matched to PubMed).

A.2 Construction of U.S. patent datasets

A.2.1 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. 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): (mostly) standardized assignee names[6]

[5] Nearly 1,300 papers were listed in both the CMR ledgers and the CMR bibliography, but sometimes with different author sequences, titles, or even just formatting. To make sure we do not double-count publications, we manually reviewed all papers to standardize the formatting, and harmonized authors and titles where we were confident that two slightly different observations represented the same paper (e.g., where the journal/volume/issue/pages were the same but the titles slightly differed). Similarly, in linking these papers to Web of Science and PubMed, we carefully review candidate matches and allow for slight deviations in authors and titles.

[6] 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

- HistPat dataset: lead inventor state and county (U.S. inventors only)[7]
- Fleming et al. (2019): indicators for government-owned or -supported patents

A small subset of patents are missing filing dates and assignees. Table A.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).

Table A.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 (of grant).

Patented, OSRD-funded inventions are identified in the OSRD archival records by the serial number of the patent application. For the purposes of this paper, it is thus critical to have accurate data on serial numbers. The application-level data (serials and filing dates) from FPO were therefore manually reviewed and validated 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.

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.

[7] See Petralia et al. (2016) for the HistPat data and documentation, as well as Berkes (2018) and Andrews (2019) for additional discussion of historical patent geography data.

Although these corrections are valuable for matching patents to secrecy orders, the low error rate for this sample also indicates that such errors are not widespread in the data.

A.2.2 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 unique list of assignees in alphabetical order, and for each assignee recording other nearby assignees up to 9 positions before and 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.

A.2.3 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 ≥ 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).

A.2.4 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, especially given the poor quality of patent text OCR. 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 (e.g., Akcigit et al. 2017, Berkes 2018, Petralia et al. 2016), which are summarized by Andrews (2019). Of these, the Petralia et al. (2016) HistPat dataset[8] is the only dataset available for public use, and is the source we use to measure inventor locations for the patents in our sample.

Measuring inventor locations from patent text for the universe of patents is too large a problem for manual transcription efforts, and researchers who have tried to do so typically rely on a mix of OCR with rule-based regex string parsing, machine learning, and manual correction. Although the earliest release of the Petralia et al. (2016) data contained errors which caused the geolocation rate (the fraction of patents with a location) to drop precipitously in certain grant years, the most recent data release (v8, January 2019) patches these errors and is the version we use. The Petralia et al.

[8]See https://dataverse.harvard.edu/dataverse/HistPat.

(2016) data provide a state, city, and county for each inventor and assignee on historical patents. The dataset also includes an accuracy measure, reporting the expected accuracy in discrete bins ("Low" to "High") based on the authors' own validation methods, which varies between 95-99%. An additional advantage of this dataset is that it provides FIPS codes for each county, with fixed (modern) county boundaries (i.e., accounting for county boundary changes over time), which we can use to merge the patent data with other county-level datasets.

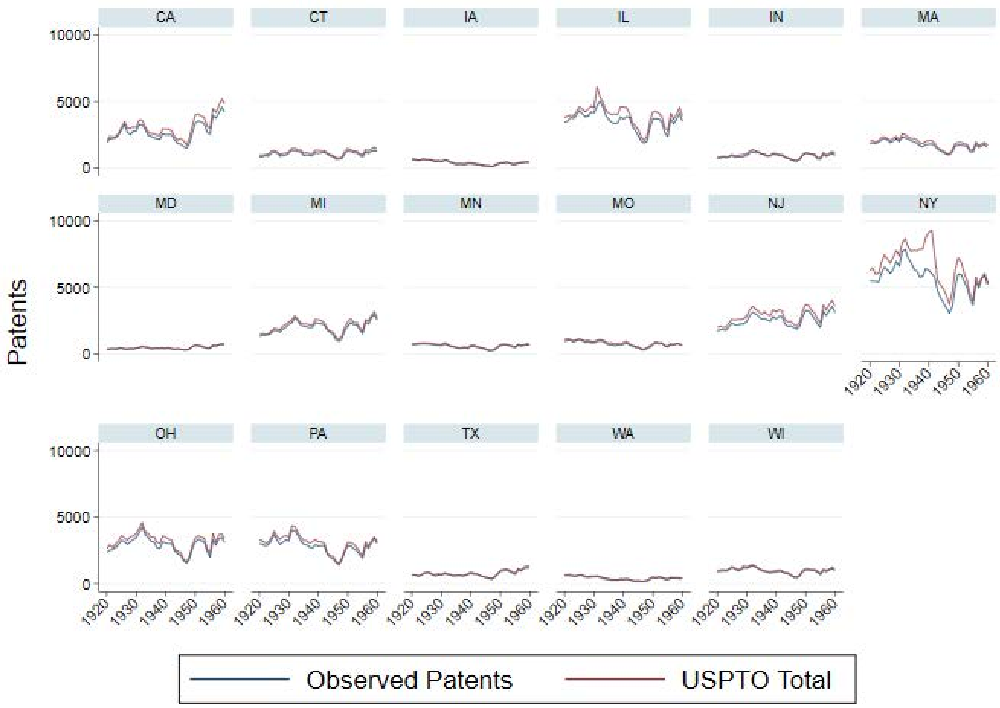
Throughout this paper, we use these data as given, identifying the first-listed inventor's location as the principal place of invention (as is convention). In parallel, we have undertaken a multi-pronged effort to validate these data ourselves. We begin by comparing counts of granted patents in our data by state and year to administrative totals from the USPTO's 1977 Technology Assessment & Forecast report.[9] We also collect two hand-entered, patent-level validation samples. For the first validation sample, we pick 200 random patents from 1930, 1940, 1950, 1960, and 1970, and have Mechanical Turk workers load the patent document and record inventor and assignee locations. For a second sample, we randomly select two pages from the USPTO's Index of Patents volumes for 1930, 1940, and 1950, and have RAs enter inventor and assignee names and locations for every patent listed on each of the sampled pages.[10] As a final check, we compare HistPat to the Berkes (2018) data, and can release results when the data are public.

Figure A.4 plots the state-year totals in our data against the USPTO administrative totals, by state, for the 1930 to 1970 period. With the exception of deviations in a handful of state-years (e.g., NY in the 1930s, RI around 1950, VT in 1940, ID in 1950), and one or two states (e.g., WV), the totals from the HistPat data closely track those in the USPTO data.

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

[10] Unfortunately, the 1960 and 1970 volumes to not list locations.

Figure A.4: State-year patent totals, HistPat vs. Administrative data, 1930-1970



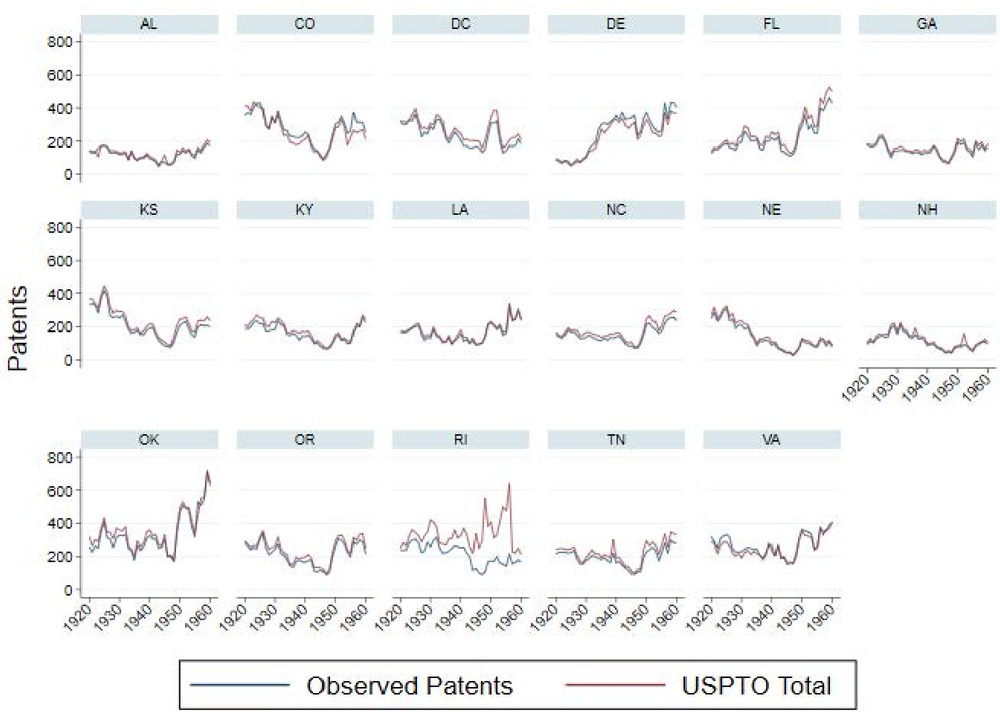
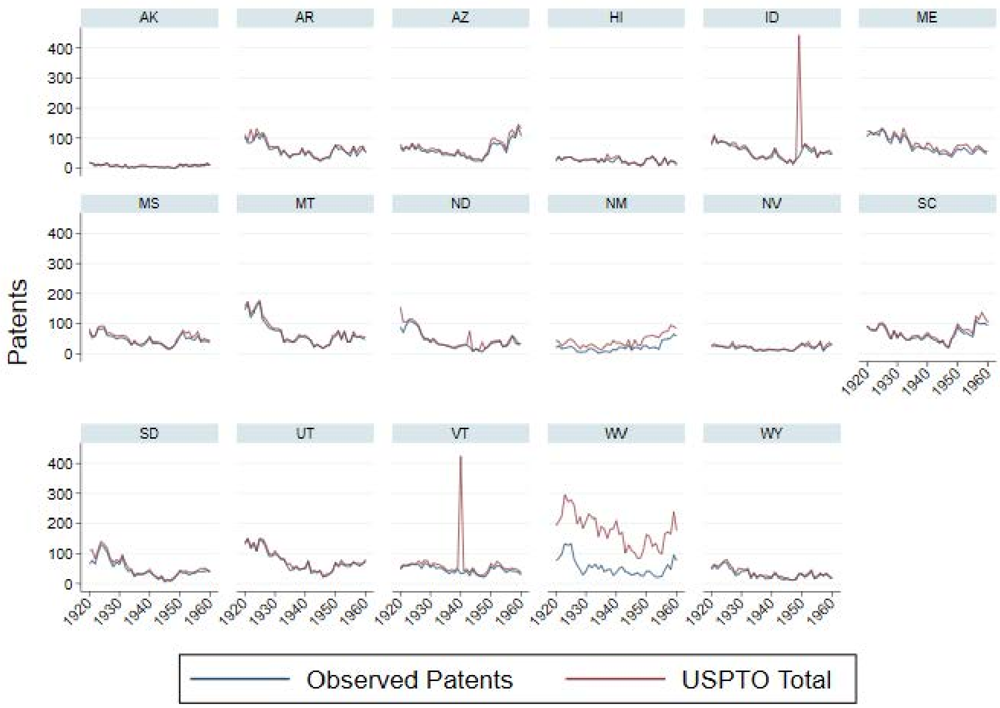


Figure A.4: State-year patent totals, HistPat vs. Administrative data, 1930-1970



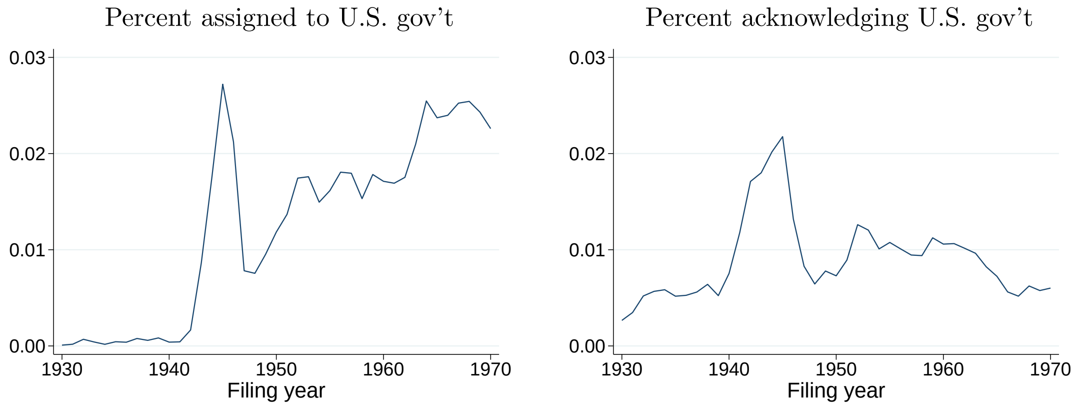
Notes: Figure plots counts of granted patents in our data by state and grant year (first inventor locations as measured in HistPat) to administrative totals from the USPTO. Sample divided into the top third of states by average annual patents over the sampled period, middle third, and bottom third for ease of viewing.

That the data so closely matches administrative totals at the state level is confidence-inspiring, but our analysis is especially dependent on the quality of the more localized city and county measurement. Using our validation samples as a ground-truth sample, we can evaluate the accuracy of the HistPat measures. Results from this exercise are forthcoming.

A.2.5 Supplementary data on government-supported patents

We also identify government-owned or -supported patents using the data of Fleming et al. (2019). We specifically borrow measures of whether a patent is owned by a U.S. government agency, or whether a patent acknowledges direct financial support from a U.S. government agency. Although these measures are imperfect – particularly the latter, due to changes in reporting requirements over the historical period we study – they can nevertheless be used to separate patents known to have been directly supported by the U.S. government. Figure A.5 below shows the time series of the fraction of annual filings that are (i) assigned to U.S. government agencies (left panel), or (ii) acknowledge U.S. government agencies (right panel).

Figure A.5: Percent of filed patents assigned to or acknowledging U.S. government, 1930-1970



Notes: Calculated using patent-level data from Fleming et al. (2019).

A.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.

A.4 Other datasets

A.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.[11]

[11] 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.

For this paper, we use 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). Due to resource constraints, we focus our county-level data collection efforts to 3-digit manufacturing industries associated with technology areas which were a focus of OSRD research, namely SIC 366 ("Communications Equipment"), SIC 367 ("Electronic Components and Accessories") – which are available from the 1959, 1970, and 1980 CBP editions – as well as all 4-digit industries within them, manufacturing sector totals, and all sector totals.[12] We also collect totals by state and industry, which were published in standalone tables and we use to validate the county-level data where possible.

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 use 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, but at these more aggregated levels, definitions are stable: all changes to 4-digit industries occur entirely within these branches. This stability is another motivation for performing our analysis at the 3-digit SIC level.

2. *Data suppression.* A more challenging problem is data suppression: the CBP historically suppressed employment and wages in county-industries with very few establishments, to avoid disclosing data on individual establishments – and this is particularly the case for later editions (the 1947 edition does not appear to have suppressed data). Because data suppression was much more common for 4-digit industries than for the 3-digit industries we study, and because the industries we study are relatively geographically concentrated, this is less of a problem for this paper than it would otherwise be – this is yet another reason to study industrial activity at these more aggregated levels. However, employment values are nevertheless suppressed for our focal industries (366 and 367) in some counties and years.

   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.

[12]In measuring industry-specific local OSRD treatment, we crosswalk patents to industries using a USPTO concordance between USPCs and (grouped-up) SICs. Given our focus on the electronics industry, we ultimately restrict attention to patents which crosswalk to the "366+" SIC group in this concordance, which represents SIC codes 366 and 367, "Electronic components and accessories and communications equipment".

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.[13] 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, 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 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 between 1947

[13] 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.

and 2000 (the year to which we fix the geographic analysis throughout the paper, to match the patent data), due to counties being combined, divided, or simply shifting borders. 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.[14] Over this period, 22 counties and independent cities (Virginia) were created, 9 were abolished, and 1 was both.

We handle these changes by combining the affected counties and independent cities into single units for analysis of CBP outcomes, since we can safely add their values together whereas we do not have solid grounds for how to divide counties into smaller pieces. For the counties and independent cities which were eliminated between 1947 and 2000, we combine them with the absorbing county in the CBP data (bringing the data forward to 2000). For the counties and independent cities which were established between 1947 and 2000, we add them back to their parent counties in the patent data (bringing these cases back to 1947).

The Historical Atlas and the Census webpage also discuss other boundary changes, most of which are boundary redefinitions or transfers of small areas between counties – often between counties and independent cities in Virginia, specifically. Because the affected population is small, these adjustments would not have a material impact on the data or results, and they are omitted. We also run our analysis of the CBP data excluding all of Virginia, and find our results unchanged, particularly as many of these cases are already omitted due to county-independent city binning in the 1959 data.

Given the focus on SIC 366 and 367, it is also useful to describe some of the industries and products therein. 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

[14] For county boundary changes from the Atlas of Historical County Boundaries, go to https://publications.newberry.org/ahcbp/index.html, select a state, and click through to the chronology of state and county boundaries. Also see the U.S. Census Bureau's list of substantial changes to county boundaries since 1970, at https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.html.

- 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

A.4.2 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.

A.4.3 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 and are now referred to as Federally-funded Research and Development Centers (FFRDCs). Our source is the NSF's annual *Federal Funds for Research,* *Development, and Other Scientific Activities* publication, which includes an appendix listing active FFRCs 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 Labs, which was funded by the DoD and managed by MIT. Having compiled a list of FFRCs by year, we then manually research each FFRC's (i) type (research lab, study and analysis center, test facility, or systems/technical direction), (ii) location, and (iii) subject matter. With these data, we can study the role that FFRCs might have played in the growth of local inventive activity after WW2.

A.4.4 Identities and locations of top PhD-granting institutions

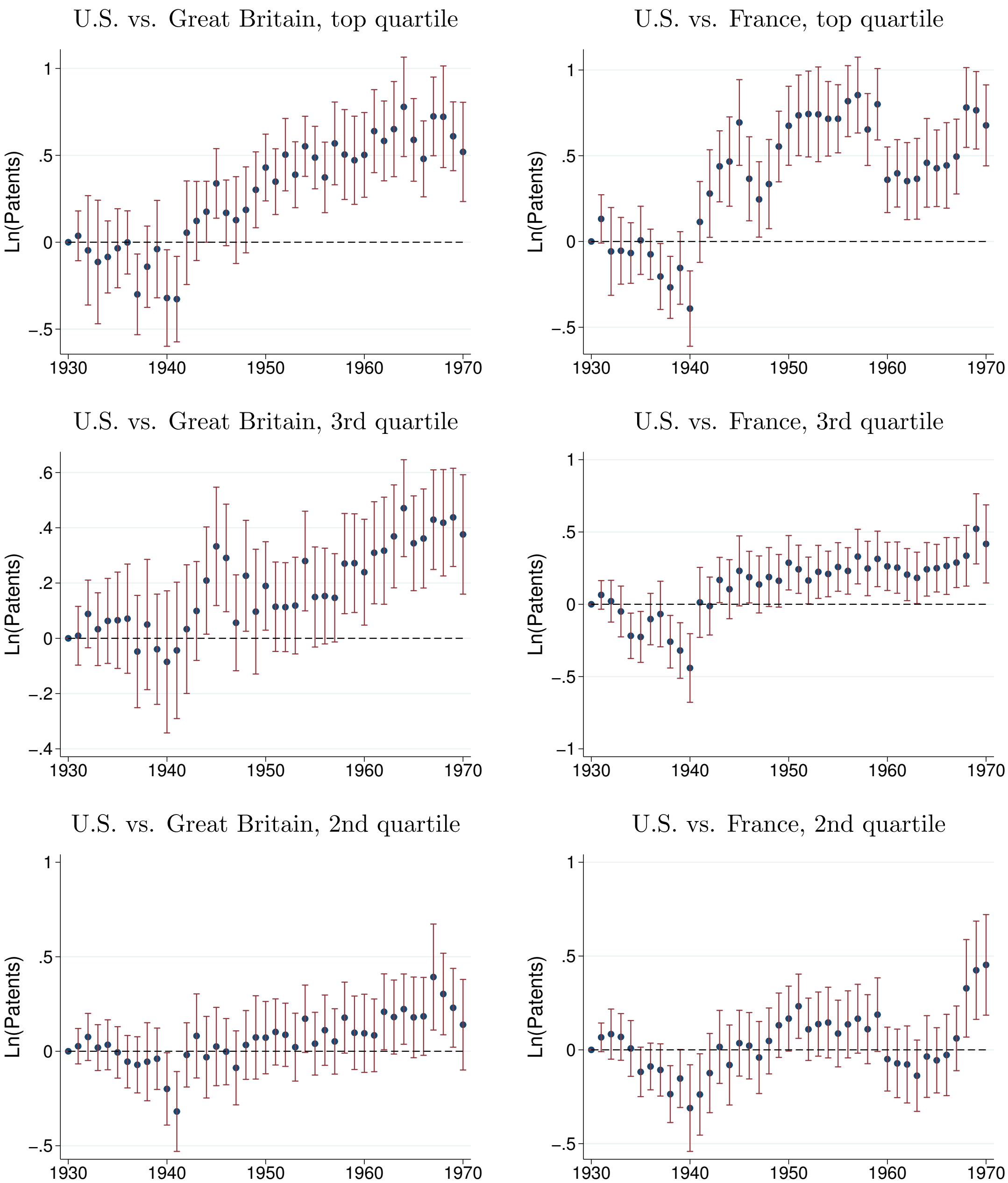
We similar collect data on historical PhD production at U.S. universities, by field and decade. Our source for these data is the 1963 report *Doctorate Production in United States Universities, 1920-1962*, published jointly by the National Academy of Sciences (NAS) and National Research Council (NRC). 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 identify the top 40 institutions of all doctoral production in the 1930s, and analogously the top 40 in the physical sciences, which includes physics, chemistry, and engineering (the other four main fields reported are biological sciences, social sciences, arts, and education, and are less germane to invention). Using these data we can test whether the long-run effects of the WW2 research effort varied across locations with versus without a top university.

A.4.5 Department of Defense R&D contracting data

We also collect data on Department of Defense (DOD) R&D contracting in the late 1960s and early 1970s, which is available in electronic format from NARA and provides contract action level data for the universe of U.S. defense contracts. These data report all the information that appears on a DOD procurement action form, including the contractor, the location, the subject matter, the value, the estimated completion date, information on whether the contractor is a small business or nonprofit institution, whether the contract was competed, and numerous other fields. We begin by cleaning the data and building a crosswalk to link DOD county codes to FIPS codes. We then use the provided Federal Supply Class (FSC) codes to identify R&D contracts, and we collapse the raw data to measure annual, county-level R&D contracts, unique contractors, and contract value to (i) all organizations in a given county, (ii) small businesses, and (iii) nonprofit institutions. We also prepare a version of this dataset disaggregated to the level of county x FSC codes over time, but with over 1500 military branch FSC codes, they vary between broad and specific in their subject matter (e.g., "Test and Evaluation" vs. "F-4/F-105 Protective System"), and are more useful for identifying specific programs of interest than for cross-program comparisons.

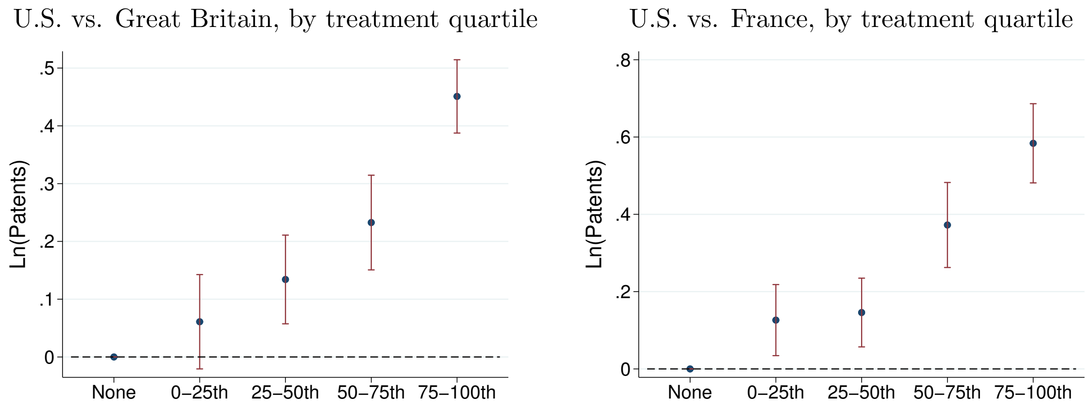
B Supplemental Results: Direction of U.S. Invention

Figure B.1: Patenting at the USPTO versus select foreign patent authorities by quartile of treated patent classes (relative to untreated classes), annual estimates, 1930-1970



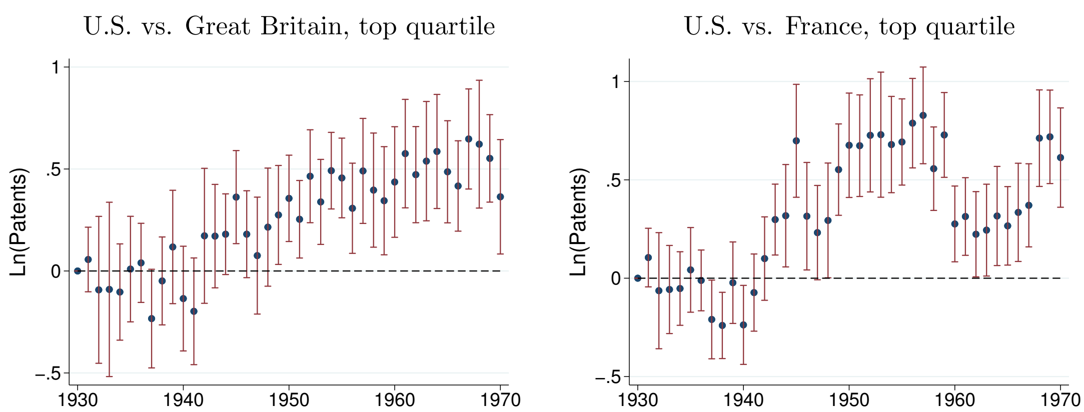
Notes: Figure shows annual difference-in-differences estimates of the effects of the OSRD shock on U.S. versus foreign patenting, in technology classes (2-digit IPCs) in the top, 3rd, and 2nd quartiles OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-supported, versus those without OSRD treatment. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-class level.

Figure B.2: Patenting at the USPTO versus select foreign patent authorities in each quartile of treated patent classes, difference-in-differences, pre-1940 vs. post-1945, excl. patents in weapons-related classes (firearms/ammunition/ordnance/explosives)



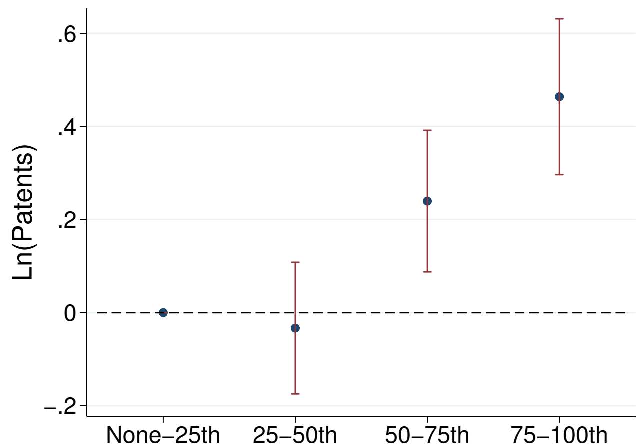
Notes: Figure shows difference-in-difference estimates of the effects of the OSRD shock on U.S. versus foreign patenting, in technology classes (2-digit IPCs) at different quartiles of OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-supported. Sample excludes technology classes with primarily military application (e.g., firearms, ammunition, ordnance, explosives). Error bars represent 95% confidence intervals, computed from SEs clustered at the country-class level.

Figure B.3: Patenting at the USPTO versus select foreign patent authorities in the top quartile of treated patent classes (relative to untreated classes), annual estimates, 1930-1970, excl. patents in weapons-related classes (firearms/ammunition/ordnance/explosives)



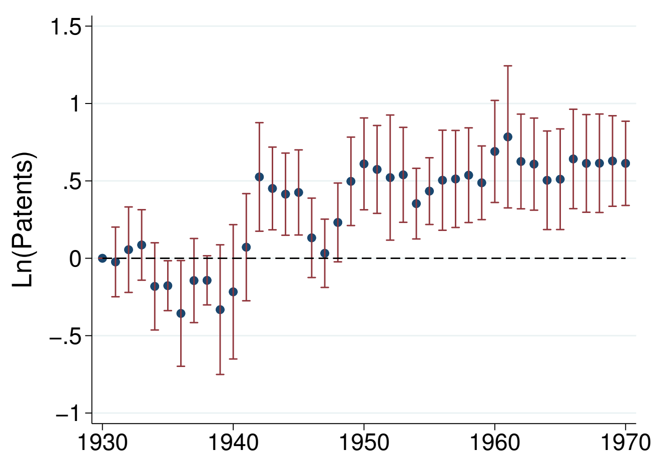
Notes: Figure shows annual difference-in-differences estimates of the effects of the OSRD shock on U.S. versus foreign patenting, in technology classes (2-digit IPCs) in the top quartile OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-supported, versus those without OSRD treatment. Sample excludes technology classes with primarily military application. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-class level.

Figure B.4: Patenting at the USPTO by domestic versus foreign inventors in each quartile of treated patent categories, difference-in-differences, pre-1940 vs. post-1945



Notes: Figure shows difference-in-difference estimates of the effects of the OSRD shock on domestic versus foreign inventor patenting *at the USPTO*, in technology areas (2-digit NBER categories) at different quartiles of OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-supported. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-category level.

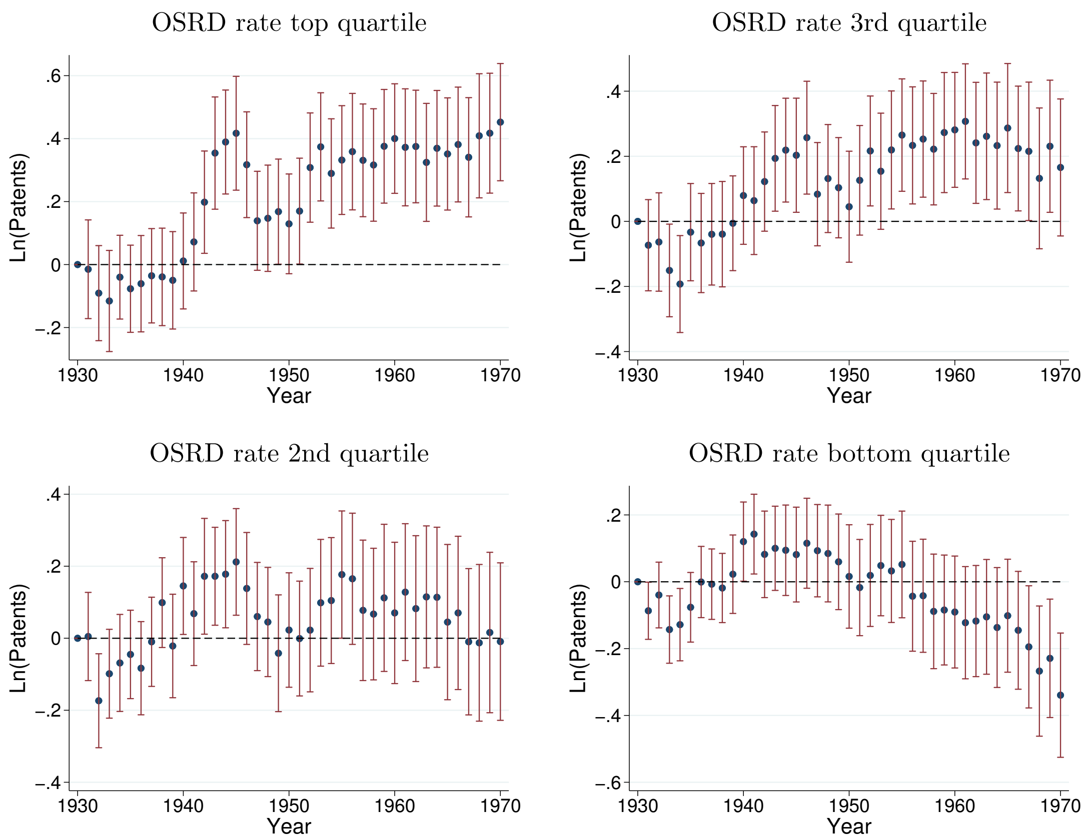
Figure B.5: Patenting at the USPTO by domestic versus foreign inventors in the top quartile of treated patent categories (relative to bottom quartile and untreated), annual estimates, 1930-1970



Notes: Figure shows annual difference-in-differences estimates of the effects of the OSRD shock on domestic versus foreign inventor patenting *at the USPTO*, in technology areas (2-digit NBER categories) in the top quartile of OSRD treatment, as measured by the fraction of U.S. patents in those classes between 1941-1948 which were OSRD-supported, versus those in the bottom quartile or without OSRD treatment. Error bars represent 95% confidence intervals, computed from SEs clustered at the country-category level.

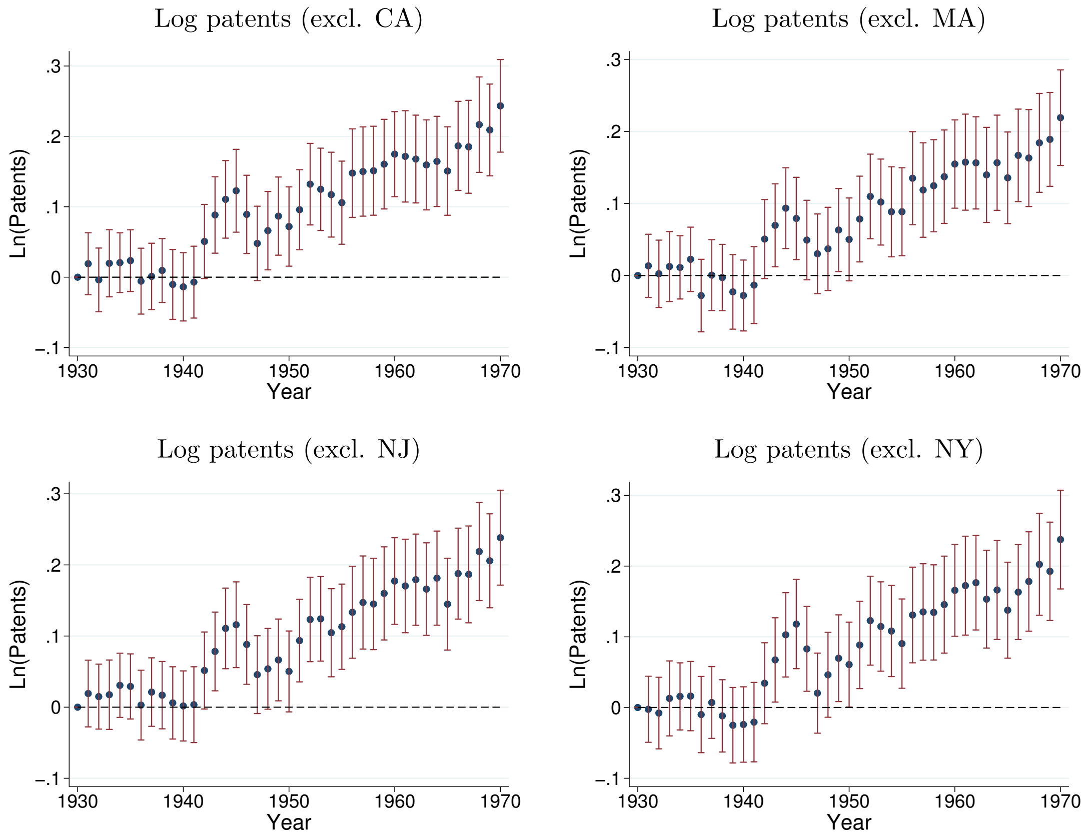
C Supplemental Results: Geography of U.S. Invention

Figure C.1: Effect of the OSRD shock on local patenting, by treatment quartile, 1930-1970



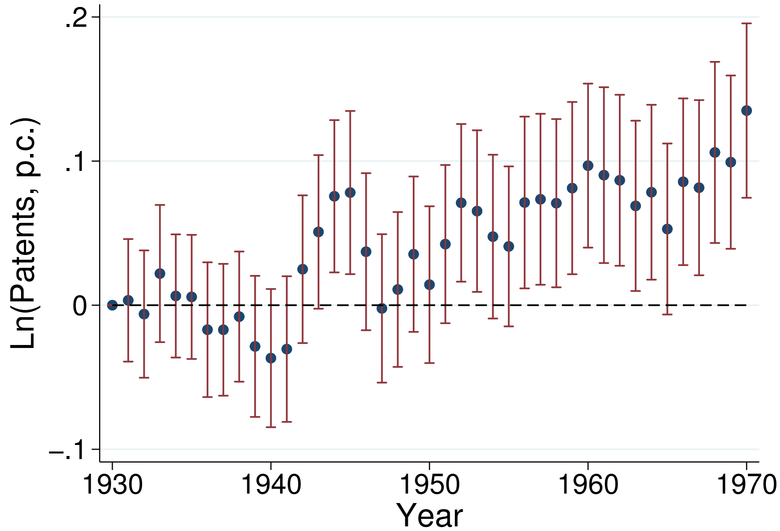
Notes: Figure shows annual estimates of the effects of the OSRD shock on county-category patenting. The independent variable measures quartiles of the fraction of U.S. patents in each county-category between 1941-1948 which were OSRD-supported, and the figure plots results by treatment quartile. Error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

Figure C.2: Effect of the OSRD shock on local patenting, excluding select states, 1930-1970



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

Figure C.3: Effect of the OSRD shock on per-capita patenting in county-categories, 1930-1970



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