Federally-Funded Science as Engines of Regional Innovation and Prosperity*

Susan Helper, Resem Makan, and Daniel Shoag†

Case Western Reserve University

March 21, 2025

Abstract

Can federal science funding jumpstart local clusters of innovation and prosperity? After World War 2 the federal government built sixteen national laboratories to conduct directed research in basic and applied science. The labs were largely built outside of pre-existing innovation hubs. We show that the locations where these laboratories were built, relative to both identified runner-up locations and synthetic controls, saw a sustained increase in innovation as measured by patenting both at the lab and by nearby unrelated innovators We find evidence of local spillovers in citations, and in a shift in research areas toward those of the labs. Further, using newly discovered and digitized county-level national data on income and retail sales from 1937-1971 (which will soon be available to download), we quantify the aggregate economic impact relative to runner-up and statistical controls. The effects are large – well above what could be explained by traditional fiscal spending multipliers – though heterogeneous across locations. Using matched person-level records from the 1940-1950 Censuses, we show that these economic gains benefited both prior residents ("stayers") and new residents. These results may have implications for the design of future funding for science and for innovation eco-systems.

Keywords: Innovation Ecosystem, Innovation Clusters, National Laboratories, R&D, Federal Investments, Placed-Based Policies, Knowledge Spillovers.

*This is a preliminary draft.

†We thank participants at the Case Western Reserve University Economics Brownbag as well as the NBER Place-based Policies and Entrepreneurship conference for helpful comments. We also gratefully acknowledge the research assistance provided by Dawson Fishman, Megan Abel, Joanna Chiu, Alivia Michalski, Giuseppe Llovet-Nava, and John McCormick. And lastly, we thank the National Science Foundation (NSF) for funding this project.

1 Introduction

Every urban economist tells their class that Silicon Valley isn't rich because it has a lot of silicon. Yet, we still don't understand how innovation clusters originate, even though they drive some of the most prosperous regional economies in the country. Can government funding actively create regional innovation hubs, even in places far from the existing technological frontier?

Much of the existing literature looks backward, studying the origins of clusters that are already known to be successful. In contrast, this paper examines the establishment of multiple national laboratories without conditioning on whether they ultimately became successful innovation hubs. Sites for national labs were often chosen for reasons largely exogenous to their innovation potential; these choices were instead based on geopolitical, military, or logistical concerns. Using newly discovered data, we show that the creation of government-funded research labs successfully triggered local innovation takeoff and broader economic growth, even in regions that previously lacked innovation infrastructure. We find strong evidence that such interventions can indeed catalyze sustained innovation and economic spillovers, benefiting both pre-existing residents and those who newly moved in.

The goal of catalyzing innovation-based regional ecosystems that deliver broad economic and societal benefits has been the subject of significant academic and policy attention. For instance, Gruber and Johnson (2019) proposed to "jump-start America" by establishing a competition in which municipalities vie for federal grants to become new technology hubs. Partially in response to this work, the U.S. Congress in the last 5 years authorized $80 billion for regions to "invest in clusters of technology, innovation, and competitiveness."[1] These initiatives include Tech Hubs (heavily influenced by the Gruber and Johnson proposal), RECOMPETE, and the Build Back Better Regional Challenge, with additional billions authorized under the CHIPS and Science Act. Similarly, the NSF Engines program "aims to fund regional coalitions of partnering organizations to catalyze technology and science-based regional innovation ecosystems."[2] A crucial question that arises, however, is whether these efforts to stimulate innovation ultimately translate into tangible and sustained economic gains for regions.

This paper studies the impacts of one of the largest targeted federal investments in United States history: the establishment of the National Laboratory System during the Cold War (Jaffe and Lerner, 2001). Built primarily in rural areas during the 1940s and 1950s, some of these labs—such as Los Alamos and Argonne—have remained prominent into modern times

[1] Breaking down an $80 billion surge in place-based industrial policy; Sizing and seizing Washington's $40 billion down payment on place-based industrial policy.

[2] Dear Colleague Letter.

(see Figure 1 for the geographic distribution of the labs). Though most were constructed with the primary goal of securing American supremacy in nuclear technology, over the decades, they have also stressed their goals of fostering local economic development. A key question this paper attempts to answer is whether these national labs generated self-sustaining local innovation hubs and/or broad-based prosperity in the regions where they were established. We focus on the period starting with the establishment of most of the labs in the 1940s, and follow them into the 1970s (when the labs' missions changed, and their budgets doubled due to the energy crisis).

While a number of studies examine the effects of federal funding on local innovation and economic growth, they largely focus on government contracts awarded to existing frontier firms or universities (Gross and Sampat, 2023; Kantor and Whalley, 2023). What distinguishes the national labs is that they involve the direct government creation of large-scale R&D infrastructure built from scratch in largely rural counties, necessitating the building of new infrastructure and creation of new organizations of researchers and technical staff. As such, they can be viewed as investments at the extensive margin.

In this paper, we investigate several types of impacts of the labs on local eco-systems:

- Innovation spillovers in the form of patents from non-lab entities in the same county; citations; changes in fields of inquiry; and informal interactions.
- Fiscal multipliers from the influx of federal dollars;
- Impacts on incomes of both movers and stayers. In addition, we look at the distribution of gains, for example as measured by changes in incomes to those who resided in the county both before and after a lab was established.[3]

For each of our channels of impact, we compare counties that received lab investments with control counties determined by a) counties that were considered for a lab's location, but did not receive a lab and b) synthetic controls (the construction of which is discussed further below).

Innovation spillovers: Using both identified "runner-up" locations and a synthetic control approach, we find that the counties where the labs were established experienced a surge in patenting activity, an effect that has persisted for at least several decades after their founding. Moreover, through archival research, we are able to distinguish between patents assigned to the labs themselves and those granted to other entities in the same counties. Notably, we find a substantial increase in the latter category following the establishment of each lab.

[3]In future work, we plan to look at human capital spillovers, as measured by changes in education in treated counties, and business formation in the historical county business patterns (Eckert et al., 2022).

These non-lab patents also co-move strongly with lab-assigned patents, which we interpret as evidence of knowledge spillovers from the labs to the broader region. Further supporting this interpretation, we find that non-lab patents shift their fields of innovation—measured by international patent classification codes—toward those of the nearby lab, suggesting technological diffusion beyond the immediate research institutions. When analyzing the geographic nature of lab patent citations, we find suggestive evidence that distance to the labs predict the likelihood of citing the lab patents. The closer a county is to a lab the more likely it is to have patents citing the labs.

Economic impact: To examine broader economic impacts, we introduce a novel dataset that allows us to analyze county-level economic outcomes such as retail sales and family income dating back to the 1930s. Using these newly digitized data, we find that the establishment of the labs led to significant increases in aggregate economic activity. Importantly, these economic effects persisted into the 1970s, the latest period for which we have data, indicating long-lasting regional benefits rather than temporary economic booms. For many of the labs, we observe multiplier effects much larger than those resulting from other forms of federal spending (eg, on physical infrastructure), again supporting the view that the lab investments played a significant role in generating a vibrant local innovation ecosystem.

For both patents and retail sales, we also observe significant heterogeneity in outcomes. To explore the sources of this variation, we use historical documents to obtain data on the operating costs of the labs. We find positive and statistically significant elasticities between lab operating costs and both patenting activity and retail sales. While acknowledging that the size of federal funding for the labs is likely endogenous to their perceived success, the positive relationship between investment levels and observed economic outcomes suggests a meaningful impact of federal investment on regional development.

To further strengthen our analysis, we incorporate counterfactual locations for select labs, identified through historical documentation as serious contenders during the site selection process. Using these runner-up counties as control units, we observe trends that align with our synthetic control estimates — counties that hosted the labs experienced greater long-term growth in both patents and retail sales, with divergence emerging particularly after the labs were founded.

Distributional impact. Additionally, leveraging US Census data linking individuals between 1940 and 1950, we investigate how prior residents of lab-hosting counties fared relative to those in runner-up locations. We find that their wages increased by 5% in the treated counties compared to the counterfactual locations, suggesting that the economic benefits of the labs extended beyond those directly employed by them, positively impacting the broader workforce.

1.1 Contribution to the Literature

Our first contribution to the literature is in highlighting the role of federal investment in fostering innovation clusters. A substantial body of qualitative research has demonstrated the importance of linkages between scientists and industry in creating environments conducive to innovation (e.g., Porter and Stern, 2001), while others have emphasized the role of large cities in facilitating agglomeration economies and attracting scientific talent (Glaeser, 2010; Kerr, 2010; Chatterji et al., 2014; Feldman and Francis, 2002). We build on this literature by providing causal evidence that policy interventions can actively stimulate such agglomeration effects, even in regions with no prior history of innovation activity.

Recently, several papers have explored the relationship between government-sponsored R&D and private-sector innovation. Using defense-related R&D, Moretti et al. (2025) found significant complementarity between private and public R&D spending. Similarly, Myers and Lanahan (2022) found that government R&D grants yielded both patents by the recipient and spillover patenting. Azoulay et al. (2019) showed that NIH spending spurred private patenting, and again, Gross and Sampat (2023) found increased patenting due to World War II R&D. Our paper differs in that it looks at direct government provision of R&D, but like these papers, finds spillovers to the private sector. In particular, "little research has been done on the impact of national labs" (Howell, 2024), something we help to rectify in this paper.

While much of the literature has focused on their formation at the intensive margin (Buera and Kaboski, 2012; Moretti, 2021), relatively little is known about the long-term economic activity driven by federal funding to create such economic clusters. Our results reveal economic gains that significantly exceed traditional Keynesian effects, suggesting that the government innovation spending unleashes self-sustaining growth trajectories in treated areas, rather than providing merely short-term fiscal stimulus.

Thirdly, we contribute to the literature on place-based policies (Greenstone et al., 2010; Diamond, 2016). The most widely accepted rationale for such policies is to benefit the existing local residents (Kline and Moretti, 2014). Canonical spatial models predict the impacts of policies to improve local incomes are dissipated over time by in-migration. In contrast, our findings suggest the possibility of generating persistent economic growth even in largely rural areas. While we document evidence for such a phenomenon, we also observe heterogeneity based on educational and occupational characteristics. This finding regarding the national labs may be in contrast to other interventions. For example Glaeser and Gottlieb (2009) make a case that many policies targeted towards less productive areas are ineffective. While we do not investigate the general equilibrium effects in this paper or test whether there could have been more productive alternative uses to the federal funds (the efficiency argument),

we nonetheless find evidence consistent with improving outcomes in poorer regions.

2 Background on National Labs

Most studies of the impact of research institutions focus on universities. In this paper, we focus on the economic impact of U.S. national laboratories. This focus is useful for several reasons. The formation of a university involves a complicated mix of treatments over decades and even centuries. The national laboratory program, which is focused solely on research, is perhaps a more accurate approximation of government programs that focus on building innovation eco-systems. While the laboratories do host post-docs, they do not house tens of thousands of students.[4] Unlike universities, they are much more tied to federal funding and a federal agency (the Department of Energy). Understanding where these more targeted institutions can, nevertheless, have large spillovers is important for understanding how to craft ongoing interventions like the NSF Regional Engines program.

The historical records suggest that, while the labs were to engage in advanced research, they were to be located in relatively secluded areas because of the sensitive nature of their work (nuclear research). For example, when deciding where to set up Argonne National Lab, the siting committee felt it best to locate it in the Argonne Forest in DuPage IL, away from the bustle of the Greater Chicago area.[5] Likewise, Los Alamos and Oak Ridge labs were built as top-secret missions and consequently needed to be as secluded as possible (Meade, 2014). Nonetheless there were also a few that were built in existing university campuses. For example Ernest Lawrence, a physics professor at UC Berkeley, helped build the Lawrence Berkeley National Lab in the same campus he worked in.[6]

All but one of the 17 national laboratories are Federally Funded Research and Development Centers (FFRDCs).[7] FFRDC's are " government-owned contractor-operated facilities (GOCOs)"; the rationale for this design is to give the labs insulation from both market and political pressures, and to enable attraction of talented personnel because they did not have to conform to civil service rules (Jaffe and Lerner, 2001; Mitre, 2015).[8] Today, these are

[4] According to the NSF, these labs as of Fall 2021 collectively hosted 3,357 postdocs (link).

[5] For its 75th anniversary, the lab released on their website a series of writings of the lab's early days titled "Those days as we remember them." According to journalist reports of Christopher Kramer, Head of Media Relations at Argonne, "There was nothing out here but some farmland and forests back then [when the lab was founded]."

[6] Berkeley lab's website tells the story of Ernest Lawrence securing a building to build a radiation lab. Over time it would expand to areas beyond the campus (https://history.lbl.gov/.)

[7] The exception is the National Energy Technology Laboratory (NETL) which is formally part of DOE. NTEL is not included in our sample.

[8] https://www.mitre.org/news-insights/publication/primer-federally-funded-research-and-development-centers-21st-century

Figure 1: National labs are geographically dispersed



Notes: Lab name followed by founding year- (1) Pacific Northwest 1949, (2) Lawrence Berkeley 1933, (3) SLAC 1952, (4) Lawrence Livermore 1952, (5) Idaho Lab 1949, (6) National Renewable Lab 1977, (7) Los Alamos 1943, (8) Sandia 1952, (9) Ames 1947, (10) Argonne 1946, (11) Fermi 1967, (12) Oak Ridge 1943, (13) Savannah River 1951, (14) National Energy Tech 1911, (15) Thomas Jefferson 1984, (16) Princeton Plasma 1951, (17) Brookhaven 1947.

overseen by the U.S. Department of Energy (which was not founded until 1977).

The National Laboratory system emerged from the extraordinary scientific efforts of World War II, particularly the Manhattan Project, which was tasked with developing the first nuclear weapons. Following the war, the U.S. government maintained these research institutions in an effort to continue advancing scientific knowledge and technology. During the Cold War, national labs played a pivotal role in nuclear research and development, for both nuclear weapons and nuclear energy. It appears that 14 of the 17 national labs were originally created to develop nuclear weapons (or other kinds of advanced weapons).[9] Most came directly out of the Manhattan Project, with different labs working on different aspects of that work to produce an atomic bomb – Oak Ridge and Ames producing high concentration

[9]See the history pages for each of the labs. The three labs that didn't follow this course are:

- National Energy Technology Laboratory (NETL), which grew out of research on coal mines and petroleum production dating back to 1910.
- National Renewable Energy Laboratory (NREL), which was established to focus on wind, solar, etc, and traces its origins to 1977.
- Lawrence Berkeley Laboratory, which was founded in 1931 to study radiation using a cyclotron (developed by Ernest Lawrence) and then did some of the theoretical work to develop the atomic bomb as well as developing medical radiation.

uranium, Hanford (now Pacific Northwest National Lab) producing plutonium, Los Alamos developing the actual bomb, and others (Argonne) working on the chemistry and metallurgy required. The labs also studied radiation and how radiation affected humans, instruments for measuring radiation, explosion physics, etc.

After World War II, several more labs were created to develop hydrogen fusion bombs (Livermore, Sandia, Savannah River, etc.), nuclear reactors for submarines (Idaho) and then for navy ships, and eventually for commercial use. After nuclear bomb testing was outlawed in 1963, it became even more important to be able to study nuclear reactions in the laboratory (Princeton Plasma Lab), through accelerators that smashed elemental particles into each other (Brookhaven, SLAC, Thomas Jefferson Lab, Fermilab), and using supercomputers that could simulate the reactions (many of the labs).

As Cold War tensions eased and funding went down, the labs (with their thousands of scientists and engineers) looked for other projects (Tarter, 2018; Johnson and Schaffer, 1994). Some labs focused on one or another of these scientific aspects of weapons production and some developed new topics to research (medical radiation, cryptography, etc.) and began seeking contracts from other agencies of the US government and from commercial businesses. In 1964 Congress placed a spending cap on the funds that the Department of Defense could obligate to FFRDCS (Dale and Moy, 2000).

However, by 1964, the focus of these labs broadened beyond nuclear research to encompass a wide range of scientific fields, including renewable energy, environmental science, and advanced computing.(Westwick, 2003) Today, these labs collaborate with universities, industries, and government agencies to drive technological advancements and stimulate economic growth. The labs developed the technology for CDs and DVDs, rechargeable lithium batteries, the cathodes most used in electric vehicle batteries, and many, many other technologies (U.S. Department of Energy, 2023). The research conducted at Argonne and other national labs has had far-reaching implications, impacting not only national security but also contributing significantly to economic growth and public health. These labs serve as hubs of innovation, fostering the development of advanced technologies and creating high-quality jobs in the surrounding regions. Through partnerships with various sectors, the labs help stimulate local economies by attracting skilled workers and catalyzing business activity.

These labs vary in size (both by location and over time). Typically, however, they represent significant infusions of resources into their local communities, especially when they were first established. Because of the danger and classified nature of their early work, the government often sought to place them in isolated locations. As their missions broadened, most of the labs have become more public-facing and are important parts of their local innovation ecosystems (though some argue that the labs are excessively "cloistered" (Tartari

and Stern, 2021; see also Dale and Moy, 2000 and Jaffe and Lerner, 2001).

One example is Argonne National Laboratory, located in Lemont, Illinois. Established in 1946 as the first national lab in the U.S., Argonne evolved from the University of Chicago's Metallurgical Laboratory, which was instrumental in the Manhattan Project. Initially, its mission centered on developing nuclear reactors for both peaceful and defense purposes. Argonne's rural location was specifically chosen to facilitate safe testing of nuclear technologies (previously the testing was done underneath the University of Chicago's football stadium, which was located in a populated area). Argonne had 2,727 employees in 1957, 2.5% of the total workforce of DuPage County (see figure A.6 for employee trends of Sandia Lab in New Mexico). In 2022, it had 3,723 full-time employees and a budget of $1.1 billion.[10]

Over the years, Argonne has expanded its research scope to include a broad array of scientific disciplines, such as physics, chemistry, biology, and engineering. Its collaborations with universities, industries, and government agencies have been instrumental in solving complex scientific challenges and driving innovations in various fields.

3 Data and Methodology

We utilize a rich collection of data from multiple sources including archival documents. For credible identification we employ various strategies to get causal estimates of the labs' impacts. In this first subsection we give more details on both publicly available data as well as novel ones we collected, followed by a methodology subsection.

3.1 Data Sources

3.1.1 New Historical County-level Retail Sales and Income Data (1936-1970)

We are not aware of publicly available annual county-level data on economic activity for this period prior to this project. As part of the work for this paper, we identified novel data collected by the *Sales Management's Survey of Buying Power* (Sales and Management Magazine, 1972) that are preserved in a handful of libraries across the United States. Starting in 1929 the series was published annually with the aim of "improving the accuracy and completeness and facility of arrangement of regional market data."[11] We obtained access to them at the Cleveland Public Library for the years 1936 to 1970. This series provides yearly county-level data on a number of economic variables including total retail sales and buying income per family. The construction of retail sales is based on the magazine's own

[10] "Argonne National Laboratory Annual Report: 1958" link p. 99; "Argonne: By the Numbers" link.

[11] See Figure A.2 in the Appendix for an image of a scan we took.

aggregation technique—using the Census of Business totals, the Department of Commerce's estimates of retailers, bank debits, Federal Reserve reports of department store sales, and the Census of identical stores—and its accuracy is further bolstered by corroboration with actual transaction tax records when available. "Buying power income" is defined as aggregate income (wages, salaries, dividends, interest, rents, entrepreneurial income, government payments) minus total taxes (both at the county level).[12] To construct this variable, they use information from the Department of Commerce and the Bureau of Internal Revenue on income tax receipts plus other state information for non-income tax levying ones. The magazine then reports buying income per-family by dividing this measure of total buying income by an estimate of total families at the county level. The Magazine states that more details about their calculations can be obtained from the Superintendent of Documents in Washington. Unfortunately we are not able to track it this late in time. Nonetheless we take their estimates to be the best option we have, in the sense that they happen to provide harmonized county-level economic data for our study period.

3.1.2 Exogenous Siting: Identifying Runner-up, Counterfactual Locations

In general, labs were sited based on rationales orthogonal to the quality of the local innovation eco-system. For example, 'ample water and power supplies' were a key criterion for siting the aptly-named Savannah River National Lab along the Savannah River (Reed et al., 2013, pg 27). A site in rural DuPage County Illinois (in the Argonne Forest) was chosen for Argonne National Lab, because continuing to perform nuclear reactions under the University of Chicago football stadium was not considered safe for such a densely populated area. Seventeen sites were initially considered for the location of Brookhaven National Laboratory, but the list was eventually narrowed down to two runner-up locations (Monmouth, NJ, and Westchester, NY) and the final site in Suffolk County, which was deemed "equally inaccessible to everybody" (Crease, 1999). As the scale of work increased, administrators sought even more isolated locations, such as Los Alamos in New Mexico,[13] Oak Ridge in Tennessee,[14] and Idaho National Lab.[15] In these locations initially there were hardly any residents (eg, 200 in Los Alamos (Hunner, 2007)). In Los Alamos and Oak Ridge, the Atomic Energy Commission built the towns from scratch, constructing houses and streets as well as labs.

For six labs we identify counterfactual locations using a variety of archival sources (see Figure A.1). These sources, which are detailed in section 2 (as well as in Appendix Table

[12]For more details, see page 30 of the 1940 issue of the Sales and Management Survey of Buying Power.

[13]Los Alamos: Hunner (2007). This book describes a number of accidents and near accidents that make clear the benefits of a remote site.

[14]Oak Ridge Lab: Jones (1985) pg 70.

[15]Idaho National Lab: Holl (1994) pg 16.

A1), give historical accounts of other locations that site scouts considered seriously before finalizing on the eventual ones. The ultimate decisions often came down to arbitrary factors; for example press accounts attribute the Idaho Falls site's selection over the town's competitors to a cocktail party thrown for the official in charge. "Guest lists were carefully crafted to include the young wives in town who were 'as winsome as possible,'" recounted the Idaho Statesman.

3.1.3 Patent Data

The historical patent data is sourced from the HistPat database (Petralia et al., 2016), which is notable for its extensive coverage of patent information dating back to 1836. This dataset not only provides detailed insights into the innovation landscape over nearly two centuries but also records the county of residence of inventors, allowing us to examine geographical trends and the spatial diffusion of technological advancements. In addition to these datasets, we extract unstructured data from the *Official Gazette of the United States Patent and Trademark Office (USPTO)*. Published on a monthly basis, the Gazette includes the patent assignee name for every patent granted by the USPTO. This allows us to disaggregate patents in treated counties into lab versus non-lab categories, a critical step for measuring knowledge spillovers and assessing the broader impact of lab-based innovations. Moreover, for our analysis of budgetary factors, we rely on the annually published Financial Reports of the Atomic Energy Commission which provide detailed accounts of operating costs during the labs' earlier years. Complementary to these, we also utilize the Patstat data (De Rassenfosse et al., 2014), which provides detailed information on patent classifications and citations, offering further insights into the technological and innovation dimensions at a granular level.

3.1.4 Individual Level Panel Data from Linked Censuses

Lastly, to capture income dynamics at the individual level, we incorporate the 1940 and 1950 individual-linked Census data (Ruggles et al., 2024). This dataset is especially valuable because it tracks the same people across census years, thereby allowing us to control for individual-specific unobservable factors. The linking matches about 54 million individuals across the two census years.

3.2 Identification Strategies

Synthetic Control Method: Given the relatively small number of treated counties and the aggregated nature of our data sources, we primarily employ the synthetic control method in our analysis. As outlined in Abadie et al. (2010), this method is particularly well-suited

for evaluating "interventions that are implemented at an aggregate level affecting a small number of large units (such as a cities, regions, or countries), on some aggregate outcome of interest." For each treated county hosting a lab, we construct a synthetic control using a donor pool comprised of other counties within the same state. This strategy enables us to effectively control for state-specific unobservables such as local policies, political dynamics, and broader economic trends that might confound our estimates.[16]

For our predictor variables we use lagged outcomes for 15 years prior to treatment, or when data doesn't permit, to the earliest possible year (Grier and Maynard, 2016; Magness and Makovi, 2023). These lagged outcomes ensure that the pre-treatment trajectories of the treated and synthetic control units are closely aligned. Our results, however, remain robust to including county-level baseline wages and education levels as well as additional economic variables as predictors. We supplement this with additional inference-based methods to confirm the robustness of our findings. Instead of pooling all labs we report results for each individual lab. The small number of treated units affords us this approach while also allowing lab-specific visualizations. Our results are robust to using a consistent set of predictor variables across measures (see Appendix Figure A.8).[17]

Runner-up counties: As detailed in the previous subsection, we also leverage our identification of runner-up counties to use them as control locations. These were places that came close to being selected but narrowly lost out, and hence might serve as a cleaner set of controls. The site selection committees used a whole set of criteria to identify potential sites and hence these counterfactuals will likely alleviate the problem of unobservable confounders in our analysis. A few influential studies follow this approach for causal identification. For example, Greenstone et al. (2010) use runner-up plants as controls in their study of agglomeration spillovers. Other recent studies using similar runner-up strategies to study the effects of educational infrastructure include Andrews (2023) and Howard et al. (2024).

Difference-in-differences: In instances where individual-level or more disaggregated data is available, we complement the synthetic control approach with the difference-in-differences (DiD) method. For example we take advantage of the linked US Census between 1940 and 1950 so as to track the same individuals and analyze their wage gains from exposure to the labs. By integrating both aggregate and individual-level analyses, our approach provides a robust framework to assess the impacts of technological and economic interventions across

[16]Los Alamos county was carved out of Sandoval county (and to a lesser extent from Santa Fe) in 1949. For our analyses, therefore, we will treat Sandoval county as the treated county and from 1949 onwards we merge Los Alamos with Sandoval.

[17]Due to concern that metro-Chicago may be atypical in ways not captured by the synthetic controls approach, for Argonne we replicate our results using only other counties in the metro area as a counterfactual. The results, which are very similar, are in Appendix figure A.4.

diverse contexts and time periods.

Due to data limitations, we restrict our sample to 12 labs only. One (NETL) was founded before our data range (1910). Two of them—NREL and Thomas Jefferson Lab—were founded in the 1970s and later, outside of our data range. Another two were founded in the same counties as an existing lab but in later years—Fermi Lab founded in DuPage county, IL of Argonne Lab and Lawrence Livermore Lab founded in Alameda county, CA of Lawrence Berkeley.

4 Results

4.1 Balance Checks

To ensure that our treatment and constructed controls are balanced at baseline, we compare the treated counties with the synthetic units (and weights) from the main results on a number of demographic and economic variables. Table A2 in the Appendix shows strong balance between the two groups on traits not used in the construction of the controls. This is very reassuring since it suggests the synthetic controls and weights do a fairly good job of creating comparable units with the treated ones, despite using our lagged outcome variables as the predictors in the synthetic control.

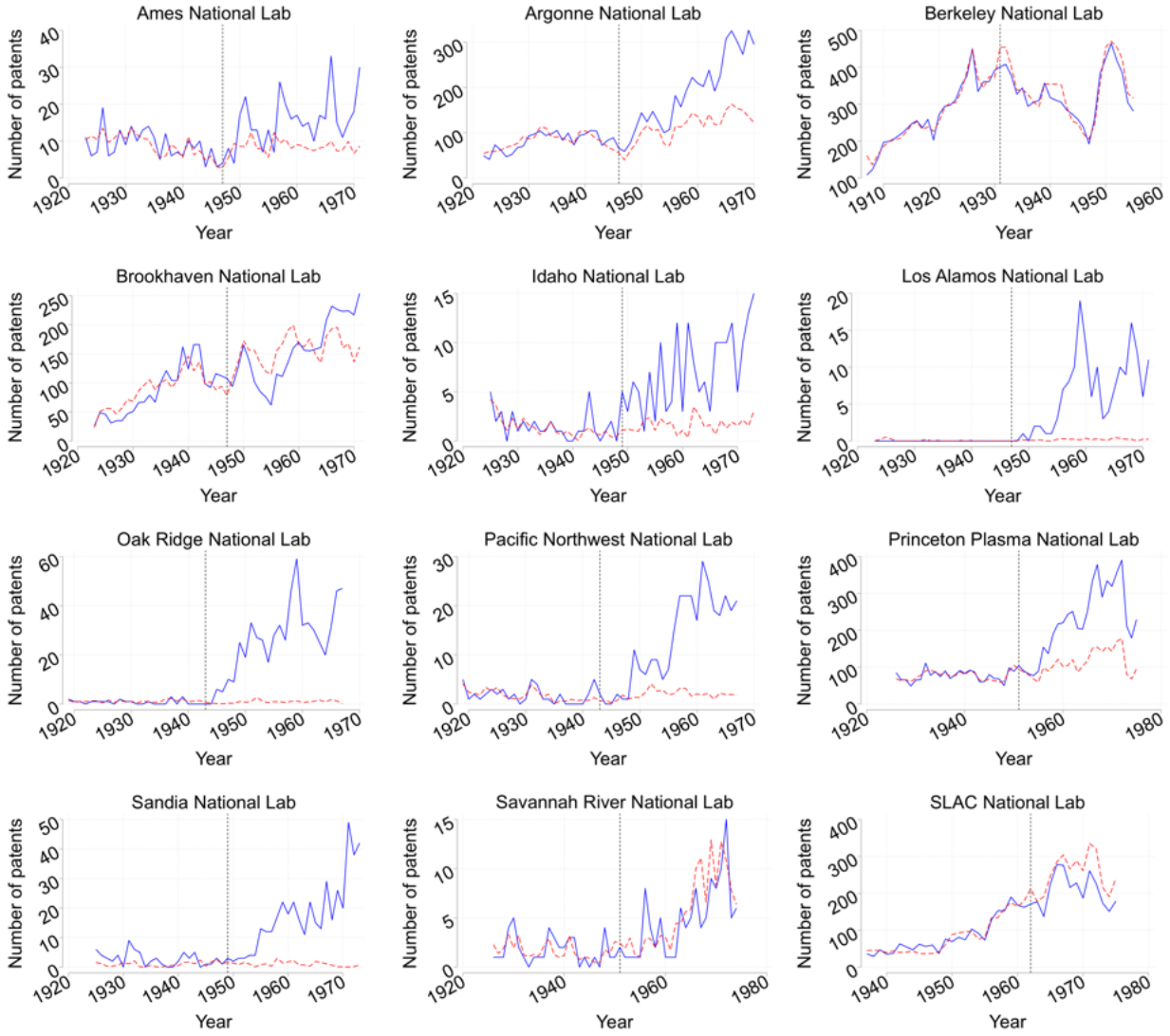
4.2 Innovation Takeoff

We first present our results focusing on patent outcomes. Our sample is restricted to inventors listed in the HistPat dataset.[18] Figure 2 illustrates the key findings using the synthetic control method. In each subfigure, the blue (solid) line represents the actual number of patents in the treated county, while the red (dashed) line shows the corresponding synthetic control series constructed from a weighted combination of counties that did not host a lab. Notably, the synthetic control series fits the treated county data quite well during the pre-treatment period, which lends credibility to our approach and sets the stage for analyzing post-treatment differences.

Examining the post-treatment period, we observe that, for two thirds of the treated counties, there is a clear uptick in the number of patents granted compared to those in the synthetic control counties. For 8 out of the 12 labs, the treated counties deviate significantly

[18] Our results barely change, quantitatively speaking, when assignees are included as well. The reason is that, almost always, for each assignee there is at least one inventor with the same patent number in the same county. As a result our sample does not change much when including the former once we drop duplicates and focus on an individual patent number in a particular county.

Figure 2: Patent changes (synthetic control method)



Notes: Blue (solid) lines plot the time series of the treated county and the red (dashed) ones plot the synthetic controls. The vertical dashed line in each subfigure indicates the founding year of the labs.

from their synthetic counterparts. Even in cases where the divergence is less pronounced, the treated counties still show an upward trend in patent counts, suggesting that the establishment of labs may have a positive impact on local innovation activity.

We obtain very similar results if using the identified runner-up counties in place of synthetic controls, as demonstrated in Figure A.3. Our results are also robust to using additional predictor variables including economic ones. Furthermore we conduct inference tests which are reported in Table A3 in the Appendix and find significance for 8 of the 12 labs tested.

Whenever we observe significant divergence of the treated county with its synthetic control, we also observe significance at conventional levels.[19]

Ten years after establishment, patents typically increase, compared to control counties, by about 57% at the median. At the mean this number is 267% though driven primarily by Oak Ridge Lab.

Lastly, for Brookhaven lab in Suffolk NY, we also have information on the number of publications during its early years. From about 195 publications to its name in 1950, this number explodes to 660 by 1960.[20]

4.3 Innovation Spillovers

If labs indeed help to create innovation ecosystems in the places they are located, we should find that innovation by labs promotes innovation by other nearby actors. We investigated several types of innovation spillovers: 1) impacts on the numbers of patents from non-lab entities in the same county; 2) citations; 3) changes in fields of inquiry; and 4) informal interactions.

4.3.1 Disaggregating Lab vs. Non-Lab Patents

To provide evidence of knowledge spillovers, we further disaggregate the patent data within treated counties into two categories: patents directly affiliated with the labs and those not linked to the labs. This disaggregation is achieved using documents from the *Official Gazette of the United States Patent Office*, which provides patent numbers and assignee names of all patents granted by the office. This information is published monthly. We identify all patents assigned to the Atomic Energy Commission—the parent organization overseeing the labs—from 1947 to 1974.[21] We then match them with the HistPat data using the unique patent numbers to identify their county locations, which we in turn interpret as a stochastic match to lab patents.[22]

Figure 3 presents this disaggregated analysis. In each subfigure, the blue (solid) line depicts the time series of lab-associated patents, while the red (dashed) line represents non-lab patents. A clear pattern emerges from these plots: the two lines tend to move in tandem

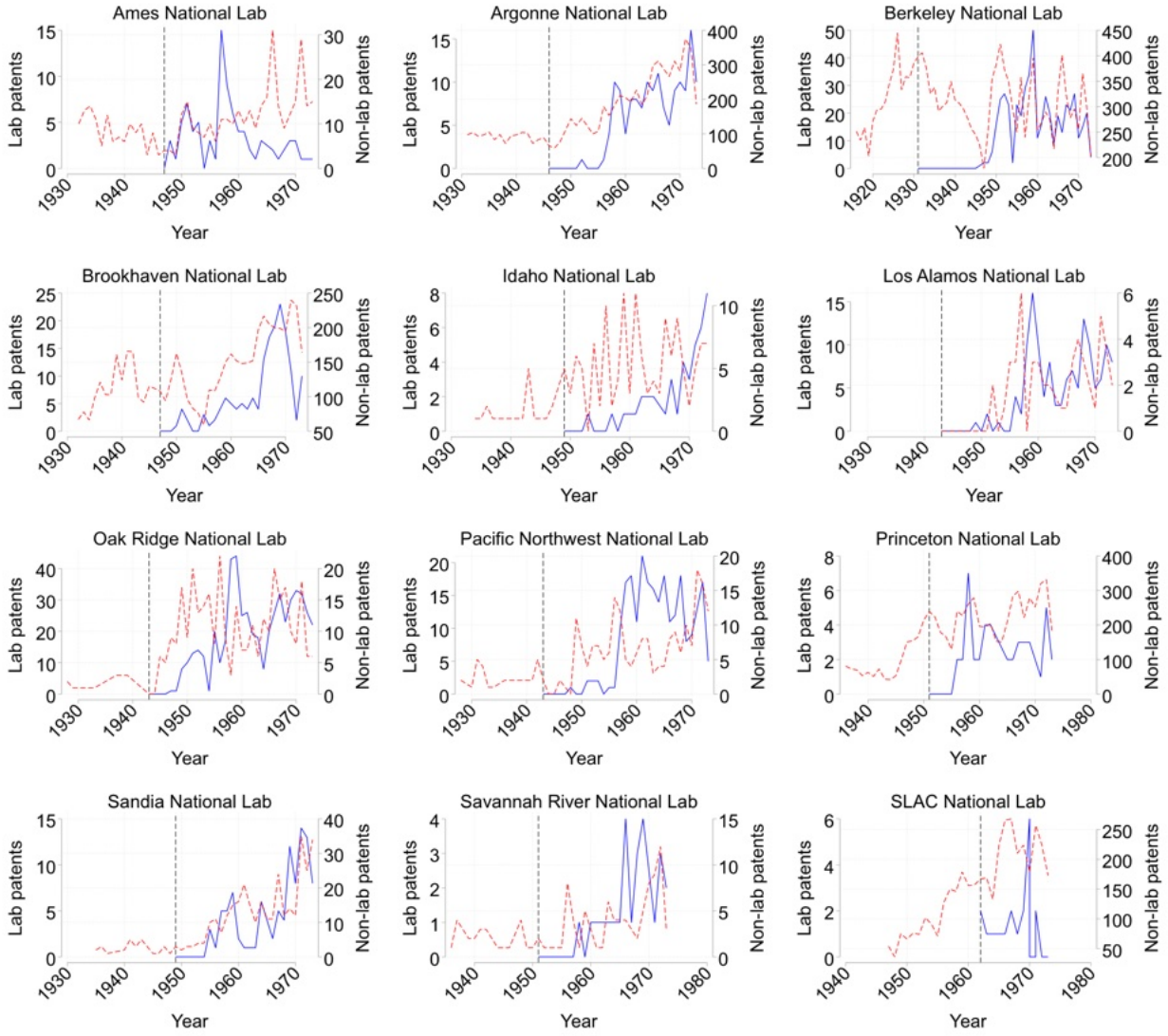
[19] In future research, we intend to investigate the causes of this heterogeneity in impact. Section 4.6 of our paper delves a little into this heterogeneity, where we assess the role of operating costs as a measure of treatment dose to further understand the differential impacts across labs.

[20] This information is taken from the lab's annual report available on Hathitrust.

[21] The Atomic Energy Commission was abolished in 1974 under the Energy Reorganization Act of 1974.

[22] To get a sense of the accuracy of this matching process we validate it with a document (Argonne National Lab Patent Portfolio compiled by B.C Huguelet) obtained from the Office of Scientific and Technical Information that provides a list of all patents granted to Argonne Lab from its founding in 1946 till 1976. The method we use matches this record with a high level of accuracy, close to a one-to-one.

Figure 3: Non-lab patents also rise after lab establishment



Notes: Time series of patents disaggregated by those granted to the lab (solid blue) vs non-lab (dashed red).

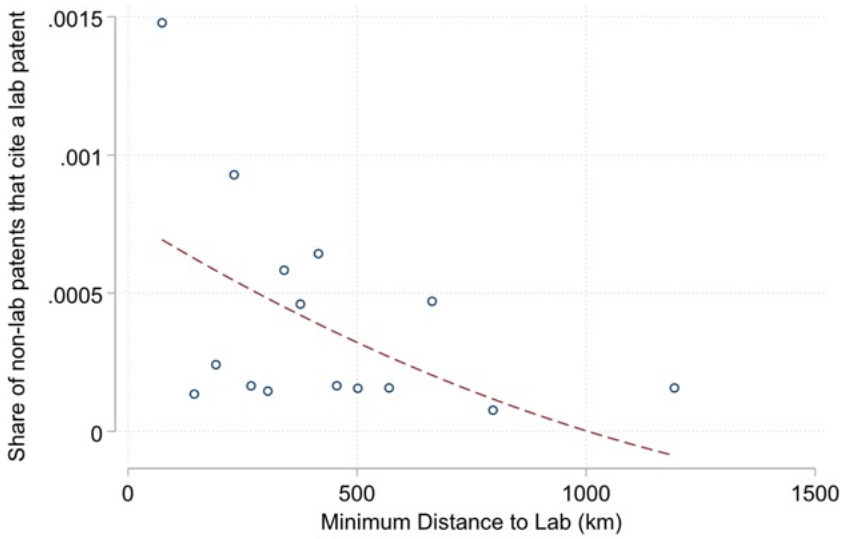
over time. This strong co-movement, while correlational in nature, suggests that lab scientists might be exerting an influence on the broader innovation ecosystem. In other words, the presence of a lab appears to stimulate not only in-house innovation but also indirectly boosts patenting activity among non-lab entities within the same county.

4.3.2 Spillovers over Space

Additional evidence of spillovers can be seen in figure 4 which plots the share of total non-lab patents that cite any of the lab patents as a function of proximity to one of the labs. A

county distance's is based on the distance of its centroid to its nearest lab county's centroid. The data is averaged across time between 1950 and 1975 at the county level.[23] The closer a county is to a lab county (figure 4), the higher its share of patents citing a lab patent (p-value = 0.033). We interpret this as evidence of spatial spillovers, indicating that the labs are positively influencing innovation in nearby regions.[24]

Figure 4: Counties closer to the lab cite the labs more



Binscatter plot: Lab patents being cited decline the further a county is from its nearest lab (p-value = 0.033).

4.3.3 Spillovers in Patent Fields

We also provide additional evidence of spillovers by showing how the International Patent Classification (IPC) codes of non-lab patents shift towards those more commonly associated with the labs after their establishment.[25] To analyse this we use linked data from Patstat (patent category) and HistPat (patent county locations) and identify non-lab patents in treated counties and compare them with patents in the runner-up counties to see if there

[23] By 1950, 10 of the labs were already established.

[24] When disaggregating further, we find that the average share of non-lab patents citing lab patents is 0.5% for treated counties compared to 0.05% for non-treated counties in the same treated states. When looking at counties in other non-treated states this value is 0.03%.

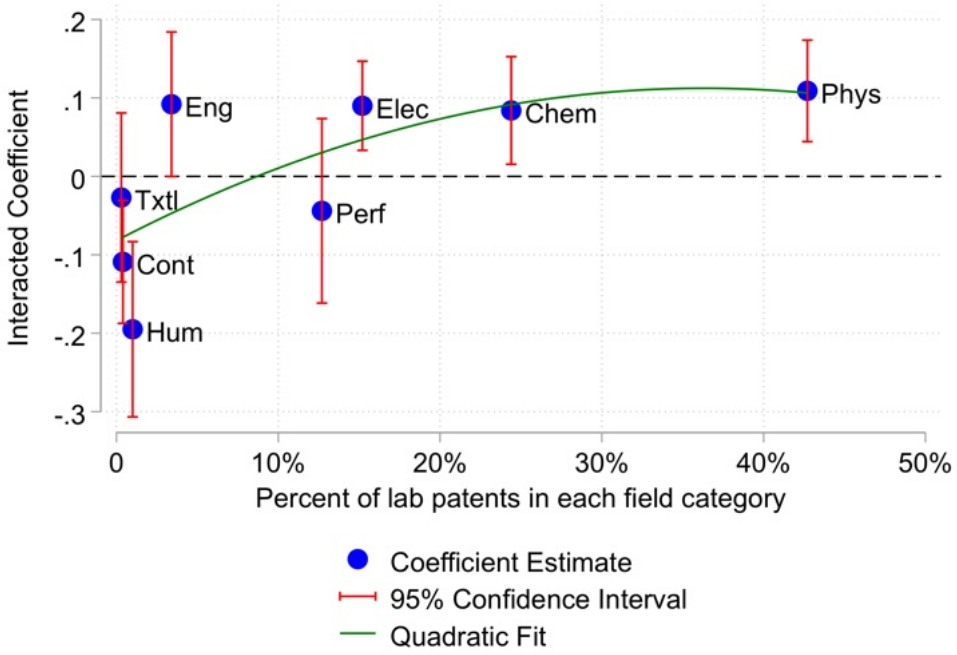
[25] The International Patent Classification categorises patents at various levels. The highest level comprises 8 categories: Chemistry, Metallurgy; Fixed Constructions; Electricity; Mechanical Engineering, Lighting, Heating, Weapons; Human Necessities; Performing Operations, Transporting; Physics; Textiles, Paper. While self-explanatory, for those interested, more details about these categories can be found on the website of the World Intellectual Property Organization.

was a shift towards categories that the labs patent in primarily. We run the following specification-

$$\text{ipc code}_{ict}^{J} = \alpha + \beta_1 \times \text{Treated}_c + \beta_2 \times \text{Post}_t + \beta_3(\text{Treated}_c \times \text{Post}_t) + \epsilon_{ict} \tag{1}$$

where $\text{ipc code}_{ict}^{J}$ is a dummy variable that equals 1 if patent i from county c at time t belongs to category J, and $J = \{$Chemistry/Metallurgy, Construction, Electricity, Engineering, Human Necessities, Performing Operations, Physics, Textiles$\}$; Treated_c indicates if county c has a lab and Post_t equals 1 if the time period t falls after the establishment of a laboratory. For this regression we use the following labs and their respective runner-up locations: Idaho Lab, Los Alamos Lab, Oak Ridge lab, Savannah River Lab, Pacific Northwest Lab.

Figure 5: Non-lab patents in treated counties shift toward lab topic areas



Notes: Diff-in-diff coefficient plot for the 8 ipc categories (equation 1). The higher a lab's share in a particular field, the more non-lab patents in treated counties shift their own fields. Legend: Phys (Physics), Chem (Chemistry and Metallurgy), Elec (Electricity), Perf (Performing Operations), Eng (Engineering), Txtl (Textiles and Paper), Const (Construction), Hum (Human Necessities). The sample includes non-lab patents in the following lab counties and their runner-up counties: Idaho Lab, Los Alamos Lab, Oak Ridge lab, Savannah River Lab, Pacific Northwest Lab.

Our coefficient of interest is β_3, which captures the relative shift in patent category for

non-lab patents post a lab's founding. Figure 5 plots the estimates for the 8 IPC categories. It is striking that for 3 of the 4 categories that the labs patent in significantly—physics, chemistry, electricity—we see that non-lab patents in treated counties see a relative increase in their share of these categories, while for the 3 categories that the lab minimally engages in—human necessities, construction, textiles—non-lab patents in treated counties fall or don't change.

Lastly, to get a sense of the geographic spread of lab patent citations, figure A.7 in the Appendix gives a visual representation of counties that cite all the lab patents up to 1965. A cursory look appears to show a clustering in the coastal regions, though we also observe higher citations from non-lab patents in counties closer to the labs. Again, we interpret this as knowledge spillovers (Kerr and Kominers, 2015).

4.4 Informal Interactions

The extent of innovation spillovers may seem surprising in the early years of the labs, when most were focused on nuclear weapons technology, which would seem to have few civilian spinoffs. In addition, security around the labs was tight, reducing (one might think) "ideas in the air." [26] In fact however, there was significant technology transfer. Some of this was due to labs' work in the fundamental science behind nuclear technologies, science that had other applications. Eventually, some applications of this fundamental science (eg., in optics) led Sandia National Lab to become a leader in solar technology (Schrank, 2015). Another source of spillovers was extensive interaction with equipment suppliers. At Lawrence Livermore, at least since the 1960s there was significant interaction with the computer and laser industries, who developed equipment, instruments or other components needed for the labs' experiments. Especially before the 1982 Stevenson-Wydler Act clarified the ownership of labs' discoveries, frequently labs and their suppliers did not seek to patent their discoveries (Jaffe and Lerner, 2001). Thus, our measurement of both lab and non-lab patenting likely understates the true amount of innovative activity sparked by the labs.

In some cases, the innovation was only indirectly related to the labs' missions. Some of the 1950s non-lab patents were for simple inventions such as arrow straightener, floor receptacle, or water cooler.[27] For example, according to his obituary, Harold Groves was a veteran who worked in Los Alamos but not as a scientist; "Harold, who was always inventing something, consulted with several scientists and built a 'bow' that would change his life." A few years later, Harold left for Albuquerque to set up a bow business where he manufactured

[26] See Gross (2023) as an example of this effect.

[27] A few patent numbers of non-lab patents in Los Alamos county in the mid 1950s: 2876822 (Arrow runout tester and straightener), 2723535 (Portable refreshment cooler).

Groves Spitfire Bows.[28]

Overall, these results offer compelling evidence that the establishment of labs has spurred an increase in innovative activities within their host counties, both directly through lab patents and indirectly through spillover effects on non-lab innovation.

4.5 Economic Effects

While it is reassuring to observe that the labs begin to affect local patenting activity—as expected given their role in knowledge creation—this naturally raises the question of whether these innovative activities translate into broader economic benefits. Notably, several labs were originally established with a mandate to promote the economic well-being of the local community. Today, this principle remains a shared goal across all labs.[29] To rigorously assess these impacts, we adopt a three-pronged approach: first, we analyze aggregate economic outcomes; second, we investigate how variations in lab budgets correlate with these aggregate outcomes; and lastly, we examine effects at the per-capita level.

4.5.1 Retail Sales

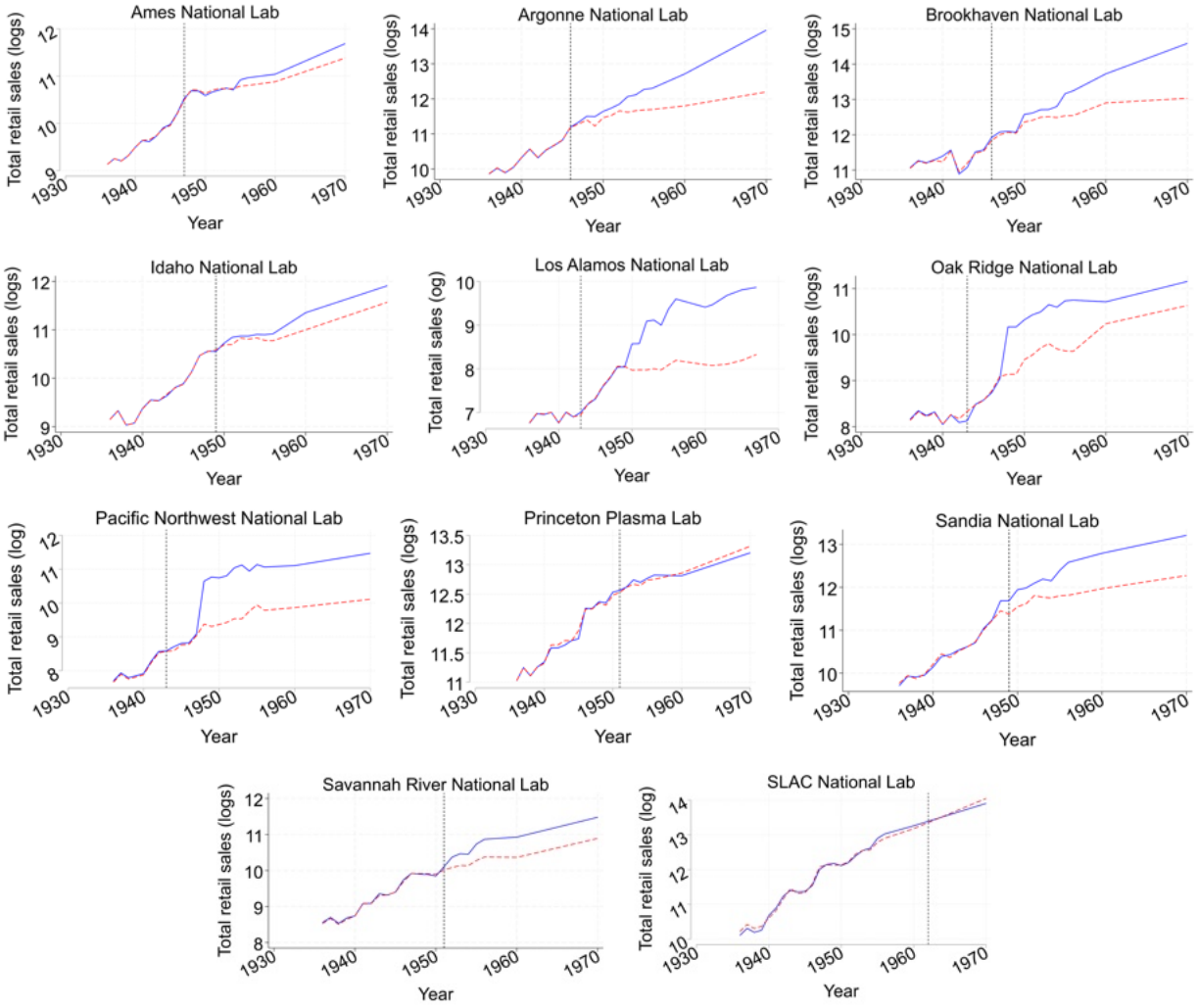
Figure 6 shows synthetic control plots for aggregate retail sales, derived from the Survey of Buying Power. This dataset spans from 1936 to 1971, providing a rich temporal perspective on economic performance at the county level. Given that Berkeley Lab was founded before the start of this dataset, we have excluded it from this segment of the analysis. The figure reveals several important patterns. For 9 out of the 11 cases examined, there is a noticeable increase in aggregate retail sales over the long run, typically manifesting around five years after the lab's founding. This delayed onset suggests that the economic benefits associated with the lab's establishment may require time to fully materialize, likely due to initial setup costs and the gradual diffusion of technological and managerial practices into the local economy.

It might seem concerning that some labs, like Argonne Lab in DuPage county IL, are located in close proximity to large cities like Chicago. As an additional robustness check we use the other 3 counties adjacent to Chicago metro area—Kane, Lake, Will—as a control group and find a similar divergence of DuPage county in its retail sales (Appendix figure A.4).

[28] Details of Grove's account can be found in this obituary written by his son in 1997.

[29] Contemporary evidence of this can be found in the economic mission outlined in most labs' individual websites. For example, Idaho National Lab states that it "is committed to fostering small business growth and economic development not only in southeastern Idaho but throughout the state. Remarkably, for every job at INL, two jobs are generated in other industries, with nearly 12,600 other Idaho jobs tied to our operations."

Figure 6: Retail sales changes (synthetic control method)



Notes: Blue (solid) lines plot the time series of the treated county and the red (dashed) ones plot the synthetic controls. The vertical dashed line in each subfigure indicates the founding year of the labs.

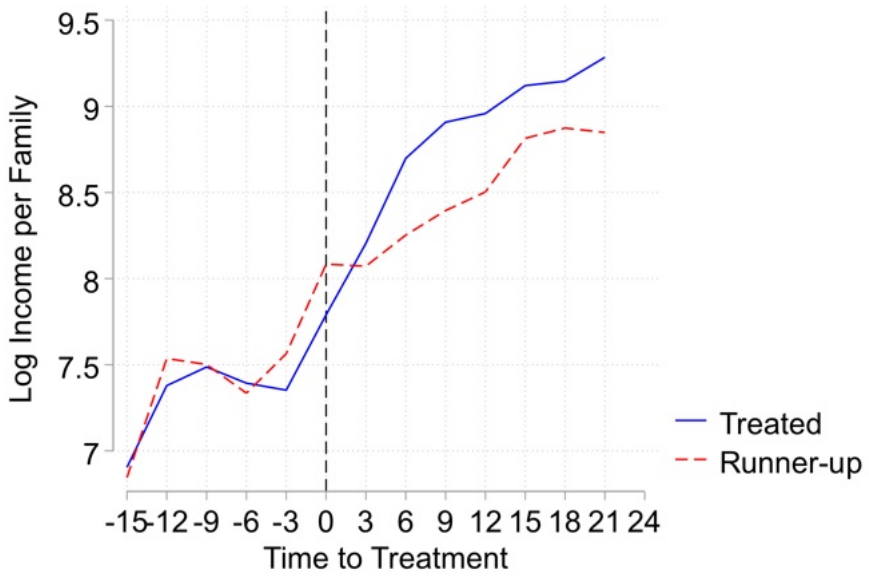
Conducting inference for the retail sales estimation (Appendix table A3), the results are slightly mixed. For example, Ames and Idaho labs get p-values that indicate non-significance even at the 10% level despite their graphs showing divergence in the longer run. This is likely due to the lagged response to the treatment as mentioned above and that the inference based-method assigns zero effects for the early years post treatment. But as Abadie (2021) states, despite this, one can be convinced of a treatment effect by having sufficient data for the post-treatment periods "if the effect of an intervention is expected to arise gradually over time."

Lastly, while an upward trend is common, there is also heterogeneity in the magnitude

and persistence of these effects. For instance, Argonne National Lab in Illinois shows a steadily widening gap between the treated county and its synthetic control, whereas Idaho National Lab experiences a divergence about six years after the lab's founding and then the gap stabilizes over time.

4.5.2 Household Income

Figure 7: Household income rises for treated counties relative to the runner-ups



Notes: The sample includes the following counties and their runner-ups: Idaho Lab, Los Alamos Lab, Oak Ridge lab, Savannah River Lab, Pacific Northwest Lab, and Brookhaven Lab. The data is averaged at three-year intervals.

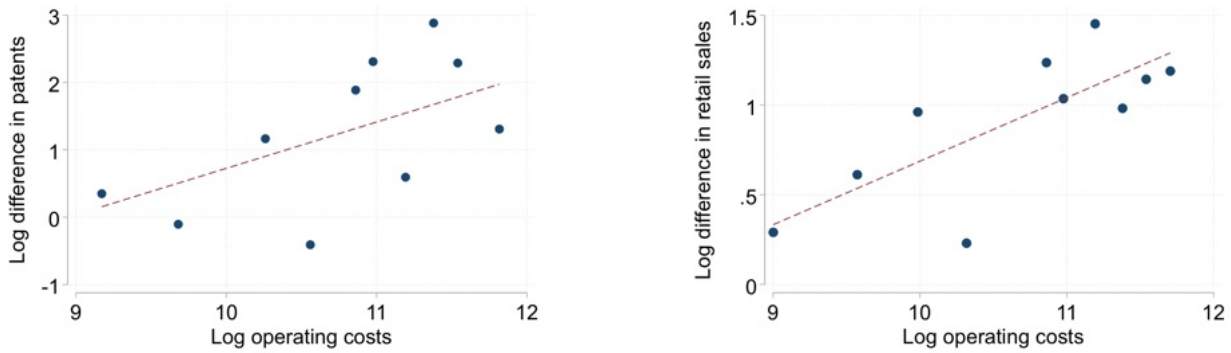
Using additional county-level data from the Sales Management's Survey of Buying Power (mentioned in section 3.1.1), we look at trends in income per family in the decades after a lab is established.[30] The sample includes the cases for which runner-ups are identified: Idaho Lab, Los Alamos Lab, Oak Ridge lab, Savannah River Lab, Pacific Northwest Lab, and Brookhaven Lab. The data is averaged at three-year intervals. Figure 7 plots the time series for the cases for which runner-up locations are identified. There is an initial divergence that then stabilizes about a decade after lab founding to about 25% higher per family income. This suggests the labs had economic effects not just at the aggregate level (Figure 6) but that they also generated increased incomes at the household level. While the aggregate changes

[30]The construction for the income per-family county estimates are outlined in section 3.1.1.

suggest increased economic activity, the increases in per-family income are indicative of the average residents becoming better off.

4.5.3 Dose-Response Relationship

Figure 8: Patents and retail sales rise with lab operating costs



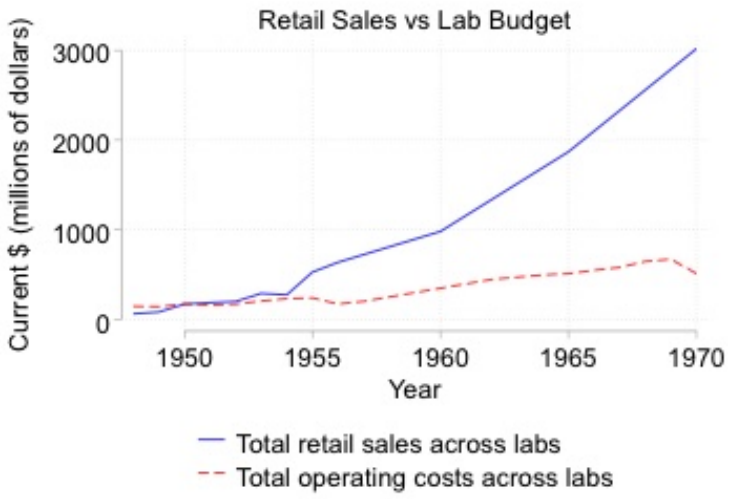
Notes: Binned scatter plots of patents (left) and retail sales (right) with respect to lab operating costs (both measured in logs).

To further understand the link between lab activity and economic performance, we utilize archival data on lab operating costs for selected years (1957, 1965, and 1970). This analysis, depicted in Figure 8, examines the correlation between a national lab's operating expenses and the logarithmic difference in aggregate retail sales between the treated county and its synthetic control. The results indicate a robust elasticity relationship: we estimate a regression coefficient of 0.4 (p-value = 0.001), suggesting that higher operating costs, which may proxy for a lab's scale or intensity of activity, are associated with larger economic expansion. For patents, the estimated coefficient is 0.6 (p = 0.55).[31]

The multiplier effect of the dose-intensity impact grows over time to become much larger than that of a typical Keynesian stimulus. Chodorow-Reich (2019) summarizes a large literature on local multipliers and finds a preferred point estimate of 1.8, consistent with several other estimates in the literature (Nakamura and Steinsson, 2014; Shoag, 2013; Fishback and Kachanovskaya, 2015; Clemens and Miran, 2012; Serrato and Wingender, 2016). However, over time, US national labs appear to develop significantly larger local multipliers. Taking

[31] We also obtained annual reports for Brookhaven lab between 1950 and 1967 from the Hathitrust repository. This was the only lab that provided annual information from its early years—the lab was founded in 1947—up to two decades later. Figure A.5 in the Appendix plots the time series of the lab's operating costs (black-dotted, left axis) and the following research activities: patent applications (blue-dashed), patents granted (black-dashed), and number of publications divided by 50 (red-solid). We do find that these three measures do move up with increases in the lab's spending, therefore likely reflecting a dose-response. For example, for every $1 million in operating costs, the lab produces about 30 publications.

Figure 9: Trends in lab budget and economic activity



the sum of retail sales in excess of synthetic controls as an outcome, the volume of economic activity is initially 0.48 times total lab operating costs in 1948, rising to 2.2 times in 1955, 3.6 times in 1965, and exceeding 5 by 1970 (see figure 9).

These large multipliers align more closely with the higher estimates found in Deleidi and Mazzucato (2021); Ciaffi et al. (2024), which suggest that R&D spending generates disproportionately larger long-run effects. While macroeconomic estimates are primarily derived from time-series evidence[32]—making them potentially sensitive to specification choices and time windows (Boysen-Hogrefe, 2025)—our findings lend empirical credibility to the broader conclusion that direct R&D expenditure has significantly greater stimulative impacts than general government spending.

4.6 Income Effect on Prior Residents: Linked Individual Census Data

Moving beyond aggregate county-level measures, we also explore the impact of lab establishment on individual wages. Leveraging the U.S. Census' ability to link individuals across the 1940 and 1950 census rounds, we investigate how residents of treated counties fared in the decade following a lab's founding. Between these census years, eight labs were established, including Ames National Lab, Argonne National Laboratory, Brookhaven National Laboratory, Idaho National Laboratory, Los Alamos National Laboratory, Oak Ridge National Laboratory, Pacific Northwest National Lab, and Sandia National Laboratories. By identi-

[32] Alternately, Dyevre (2023) uses an instrumental variable approach and Fieldhouse and Mertens (2023) use narratively identified R&D shocks to estimate productivity impacts.

Table 1: Individual level wage regressions for stayers in the same county between 1940 and 1950

VARIABLES	(1) Wage change (logs)	(2) Wage change (logs)	(3) Wage change (logs)	(4) Wage change (logs)
Treated county	0.040**	0.013	0.022	0.005
	(0.020)	(0.020)	(0.021)	(0.021)
College × Treat		0.274***		0.252***
		(0.074)		(0.076)
Professional × Treat			0.143**	0.086
			(0.061)	(0.063)
Observations	13,278	13,278	13,278	13,278
R-squared	0.026	0.028	0.027	0.028

Notes: The sample is restricted to male stayers in the same county between 1940 and 1950, and who were in the labor force in 1940 and under 65 in 1940. Robust standard errors in parentheses. All regressions control for 1940 baseline wages. College is a dummy variable that equals 1 if an individual has completed college. Professionals are industry types as defined by the IPUMS 1940 census. These include educational services, medical and other health services, legal, engineering, and miscellaneous professional services, charitable, religious, and membership organizations. *** p<0.01, ** p<0.05, * p<0.1

fying counterfactual locations for four of these labs (Brookhaven, Idaho, Los Alamos, and Pacific Northwest), we construct 4 treatment clusters and 6 control clusters. We focus on "stayers"—individuals who remained in their respective counties in both census periods—to mitigate concerns that higher-income individuals might self-select into counties with labs. Though this movement of people is itself alab effect, it, nonetheless, is interesting to explore the impact of the lab on those who might not have been endogenously pulled in. The sample is restricted to men who were in the labor force in both years and under 65 in 1940. The estimation follows the familiar form:

$$\Delta \text{wage}_{i,c,s} = \alpha + \beta \times \text{Treated}_c + \delta \mathbf{X}_i + \Gamma_s + \epsilon_{i,c,s} \tag{2}$$

where $\text{wage}_{i,c,s}$ denotes the log wage difference between 1950 and 1940 for individual i in county c in state s; Treated_c equals 1 if county c received a lab between the two census rounds and 0 otherwise; X_i represents a vector of individual characteristics, including baseline wage, age and employment status in 1940; and Γ_s captures state fixed effects. Table 1 presents the results of our estimation. Column (1) indicates that treatment status is associated with a significant wage premium. On average, individuals residing in treated counties experienced a 4–5% increase in wages compared to those in counterfactual locations. We further explore heterogeneity by certain baseline characteristics. Looking at columns 2 to 4 we find that the

labs had differential impacts on the wages of states with a large boost for college educated workers as well as those in professional industries.

Lab histories provide examples of mechanisms behind the higher wages for existing residents when labs come in. Some residents were hired directly, as "carpenters, furnace stokers, and janitors" or lab technicians and secretaries; others contracted with the labs or did business with such contractors (such as the residents of San Ildefonso Pueblo near Los Alamos national lab, who rented reservation land to a construction company.[33] Another mechanism by which labs affect local residents was through education spillovers. We owe the presence of one of our colleagues at Case Western, Prof. Jenny Hawkins, to this effect. She was raised in a family where girls did not think of college as appropriate for them. However, she grew up in the vicinity of Idaho National Lab, which organized Saturday field trips to the labs for girls. When she learned what scientists did, she was intrigued, and wanted to be like them. She found her way to a BA at Idaho State University, and eventually to a PhD in economics. More directly, schools near the labs were often excellent and generously subsidized by the Atomic Energy Commission (Hunner, 2007; Schrank, 2015).

5 Conclusion

This paper examines the long-run effects of the National Laboratory System on local innovation and economic development. Leveraging a synthetic control approach, archival research, and counterfactual site comparisons, we document substantial and persistent differences in patenting activity, retail sales, and individual income growth between counties that hosted national labs and otherwise similar counties that did not. Our findings indicate that these labs were not only critical centers of scientific advancement but also played a broader role in fostering self-sustaining innovation clusters and regional economic growth.

The results highlight several key mechanisms driving these impacts. First, the labs appear to have generated substantial knowledge spillovers, as evidenced by the increase in non-lab patenting activity in treated counties, the co-movement between the numbers of lab and non-lab patents, and the observed shifts in non-lab patents toward the technological domains of the labs. Second, the economic benefits of the labs extended beyond direct employment effects, with significant wage increases for individuals residing in treated counties, even for those not directly affiliated with the labs, including non-college graduates. Third, we find a positive correlation between lab operating costs and economic outcomes, suggesting that continued federal investment in these research hubs reinforced their long-term impact.

[33]Not all the spillover effects were positive. Access to open space was restricted, and experiments "detonated blasts that at times scattered radioactive particles over surrounding areas" (Hunner, 2007).

Our study contributes to ongoing policy discussions on place-based innovation strategies. As policymakers consider new federal investments in regional technology hubs—such as those proposed in the CHIPS and Science Act, the NSF Regional Engines program, and the Build Back Better Regional Challenge—our findings provide historical evidence that such investments can create durable economic and innovation clusters. The lessons from the National Laboratory System suggest that adequately-funded research institutions have the potential to catalyze long-term regional development, even in areas without preexisting innovation ecosystems.

Our study looks at the establishment of national labs – the extensive margin. In many cases, there was very little human activity in the region before the establishment of the lab.[34] As noted above, we find evidence that lab investments created a durable innovation eco-system. The absence of previous activity makes it easier to attribute the cause of the innovation to the establishment of the lab. However, this factor should not be interpreted as a finding that funds for innovation are most effectively spent in rural areas. It may be that the presence of complementary institutions (such as pre-existing universities or innovative firms) would increase multipliers even more and reduce the size of investment required to achieve takeoff. This possibility would be consistent with Gruber and Johnson (2019), who propose directing significant innovation spending to build technology hubs in cities that have already significant science prowess.

Future research could explore how modern national labs and similar research institutions interact with today's innovation landscape, particularly in an era of increasing private sector R&D investment and evolving regional economic conditions. Further analysis of the specific conditions that enhance knowledge spillovers, firm creation, and commercialization of lab-generated innovations could provide additional insights for designing effective technology-based regional development policies.

Taken together, our results suggest that large-scale, place-based scientific investments can have profound and lasting effects on local economies and benefit most of their residents. By fostering knowledge spillovers and stimulating economic growth, national labs have demonstrated that government-backed research institutions can serve as powerful engines of regional development.

[34]For example, the population of Los Alamos County grew from 200 in 1943 to more than 12,000 in 1957, of whom almost half worked at the lab (Hunner, 2007 p4, 30).

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of economic literature*, 59(2):391–425.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, 105(490):493–505.

Andrews, M. J. (2023). How do institutions of higher education affect local invention? evidence from the establishment of us colleges. *American Economic Journal: Economic Policy*, 15(2):1–41.

Azoulay, P., Graff Zivin, J. S., Li, D., and Sampat, B. N. (2019). Public r&d investments and private-sector patenting: evidence from nih funding rules. *The Review of economic studies*, 86(1):117–152.

Boysen-Hogrefe, J. (2025). Is the supermultiplier currently nil?-a replication study of deleidi and mazzucato (2021). *Research Policy*, 54(3):105176.

Buera, F. J. and Kaboski, J. P. (2012). The rise of the service economy. *American Economic Review*, 102(6):2540–2569.

Chatterji, A., Glaeser, E., and Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innovation policy and the economy*, 14(1):129–166.

Chodorow-Reich, G. (2019). Geographic cross-sectional fiscal spending multipliers: What have we learned? *American Economic Journal: Economic Policy*, 11(2):1–34.

Ciaffi, G., Deleidi, M., and Mazzucato, M. (2024). Measuring the macroeconomic responses to public investment in innovation: evidence from oecd countries. *Industrial and Corporate Change*, 33(2):363–382.

Clemens, J. and Miran, S. (2012). Fiscal policy multipliers on subnational government spending. *American Economic Journal: Economic Policy*, 4(2):46–68.

Crease, R. P. (1999). *Making Physics: A Biography Of Brookhaven National Laboratory, 1946–1972. University of Chicago Press.*

Dale, B. C. and Moy, T. D. (2000). The rise of federally funded research and development centers. Technical Report SAND2000-2212, Sandia National Labs., Albuquerque, NM, and Livermore, CA (US).

De Rassenfosse, G., Dernis, H., and Boedt, G. (2014). An introduction to the patstat database with example queries. *Australian economic review*, 47(3):395–408.

Deleidi, M. and Mazzucato, M. (2021). Directed innovation policies and the supermultiplier: An empirical assessment of mission-oriented policies in the us economy. *Research Policy*, 50(2):104151.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *American economic review*, 106(3):479–524.

District, M., OF, M., Energy, D. o., and Engineers, U. S. A. C. o. (2020). *Manhattan District History: Book I - General; Volume I - General*. Manhattan District History. Nimble Books LLC.

Dyevre, A. (2023). Public r&d spillovers and productivity growth. Working paper. Accessed: 2025-03-20.

Eckert, F. et al. (2022). The early county business pattern files: 1946-1974. Technical Report w30578, National Bureau of Economic Research.

Feldman, M. P. and Francis, J. (2002). The entrepreneurial spark: Individual agents and the formation of innovative clusters. In *Complexity and Industrial Clusters: Dynamics and Models in Theory and Practice*, pages 195–212. Springer.

Fieldhouse, A. J. and Mertens, K. (2023). *The returns to government R&D: Evidence from US appropriations shocks*. Federal Reserve Bank of Dallas, Research Department.

Fishback, P. and Kachanovskaya, V. (2015). The multiplier for federal spending in the states during the great depression. *The Journal of Economic History*, 75(1):125–162.

Furman, N. S. (1989). Sandia national laboratories: The postwar decade. Technical Report SAND-89-2261-Pt. 1, Sandia National Lab. (SNL-NM), Albuquerque, NM (United States).

Glaeser, E. L. (2010). *Agglomeration economics*. University of Chicago Press.

Glaeser, E. L. and Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature*, 47(4):983–1028.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of political economy*, 118(3):536–598.

Grier, K. and Maynard, N. (2016). The economic consequences of hugo chavez: A synthetic control analysis. *Journal of Economic Behavior & Organization*, 125:1–21.

Gross, D. P. (2023). The hidden costs of securing innovation: The manifold impacts of compulsory invention secrecy. *Management Science*, 69(4):2318–2338.

Gross, D. P. and Sampat, B. N. (2023). America, jump-started: World war ii r&d and the takeoff of the us innovation system. *American Economic Review*, 113(12):3323–3356.

Gruber, J. and Johnson, S. (2019). *Jump-starting America: How breakthrough science can revive economic growth and the American dream*. Hachette UK.

Holl, J. M. (1994). The national reactor testing station: The atomic energy commission in idaho, 1949-1962. *The Pacific Northwest Quarterly*, 85(1):15–24.

Howard, G., Weinstein, R., and Yang, Y. (2024). Do universities improve local economic resilience? *Review of Economics and Statistics*, 106(4):1129–1145.

Howell, S. T. (2024). Government intervention in innovation. *Annual Review of Financial Economics*, 16.

Hunner, J. (2007). *Inventing Los Alamos: The growth of an atomic community*. University of Oklahoma Press.

Jaffe, A. B. and Lerner, J. (2001). Reinventing public r&d: Patent policy and the commercialization of national laboratory technologies. *RAND Journal of Economics*, pages 167–198.

Johnson, L. and Schaffer, D. (1994). *Oak Ridge National Laboratory: The First Fifty Years*. *University of Tennessee Press*.

Jones, V. C. (1985). *Manhattan, the Army and the atomic bomb*, volume 8. US Government Printing Office.

Kantor, S. and Whalley, A. T. (2023). Moonshot: Public r&d and growth. Technical report, National Bureau of Economic Research.

Kerr, W. R. (2010). Breakthrough inventions and migrating clusters of innovation. *Journal of urban Economics*, 67(1):46–60.

Kerr, W. R. and Kominers, S. D. (2015). Agglomerative forces and cluster shapes. *Review of Economics and Statistics*, 97(4):877–899.

Kline, P. and Moretti, E. (2014). People, places, and public policy: Some simple welfare economics of local economic development programs. *Annu. Rev. Econ.*, 6(1):629–662.

Magness, P. W. and Makovi, M. (2023). The mainstreaming of marx: Measuring the effect of the russian revolution on karl marx's influence. *Journal of Political Economy*, 131(6):1507–1545.

Moretti, E. (2021). The effect of high-tech clusters on the productivity of top inventors. *American Economic Review*, 111(10):3328–3375.

Moretti, E., Steinwender, C., and Van Reenen, J. (2025). The intellectual spoils of war? defense r&d, productivity, and international spillovers. *Review of Economics and Statistics*, 107(1):14–27.

Myers, K. R. and Lanahan, L. (2022). Estimating spillovers from publicly funded r&d: Evidence from the us department of energy. *American Economic Review*, 112(7):2393–2423.

Nakamura, E. and Steinsson, J. (2014). Fiscal stimulus in a monetary union: Evidence from us regions. *American Economic Review*, 104(3):753–792.

Petralia, S., Balland, P.-A., and Rigby, D. L. (2016). Unveiling the geography of historical patents in the united states from 1836 to 1975. *Scientific data*, 3(1):1–14.

Porter, M. E. and Stern, S. (2001). Innovation: location matters. *MIT Sloan management review*.

Ramsey, N. F. (1966). Early history of associated universities and brookhaven national laboratory. brookhaven lecture series no. 55. Technical report, Brookhaven National Laboratory (BNL), Upton, NY (United States).

Reed, M., Swanson, M., Tyson, J., and Gillett, T. (2013). Bringing it to form—a thematic study of svannah river sites's separation processes, f and h areas. *New South Associate Technical Report*, 2202.

Ruggles, S., Nelson, M. A., Sobek, M., Fitch, C. A., Goeken, R., Hacker, J. D., Roberts, E., and Warren, J. R. (2024). Ipums ancestry full count data: Version 4.0 [dataset]. minneapolis, mn: Ipums.

Sales and Management Magazine (1936-1972). *Survey of Buying Power*. Bill Brothers Publications.

Schrank, A. (2015). Green capitalists in a purple state: Sandia national laboratories and the renewable energy industry in new mexico. In *State of Innovation*, pages 96–108. Routledge.

Serrato, J. C. S. and Wingender, P. (2016). Estimating local fiscal multipliers. Technical report, National Bureau of Economic Research.

Shoag, D. (2013). Using state pension shocks to estimate fiscal multipliers since the great recession. *American Economic Review*, 103(3):121–124.

Tartari, V. and Stern, S. (2021). More than an ivory tower: The impact of research institutions on the quantity and quality of entrepreneurship. Technical Report w28846, National Bureau of Economic Research.

Tarter, C. B. (2018). *The American Lab: An Insider's History Of The Lawrence Livermore National Laboratory. Johns Hopkins University Press.*

U.S. Department of Energy (2023). 75 breakthroughs by america's national laboratories. Accessed: 2025-03-20.

Westwick, P. J. (2003). *The National Labs: Science In An American System, 1947–1974. Harvard University Press.*

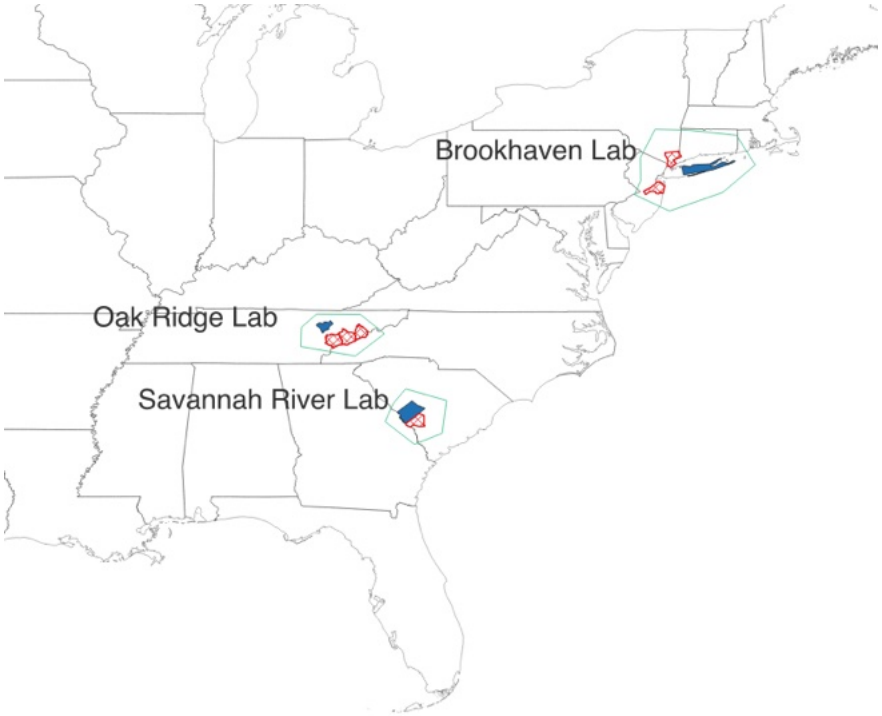
A Appendix

A.1 Sources for Identifying Runner-Up Locations

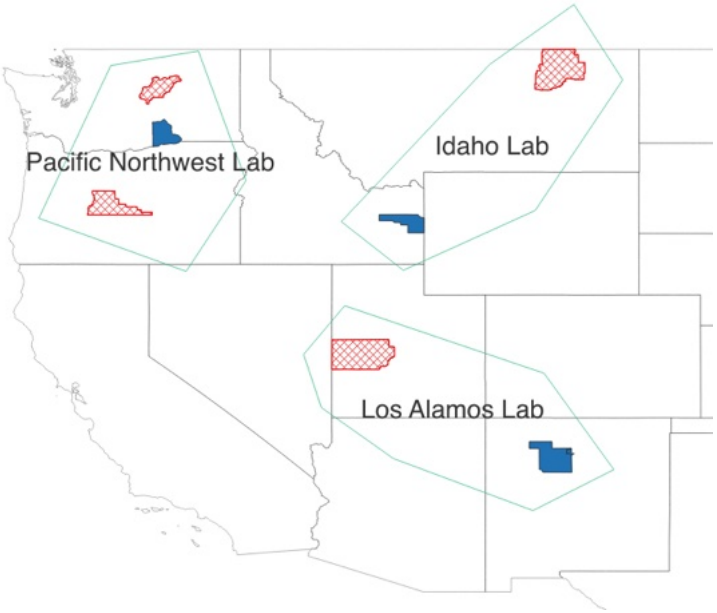
Lab Name	Sources
Brookhaven Lab	Ramsey (1966), pp. 4 - 6
Idaho National Lab	Holl (1994), pg. 16
Oak Ridge Lab	Jones (1985), pg. 70
Los Alamos Lab	Link 1, Link 2
Pacific Northwest Lab (earlier called Hanford Site)	District et al. (2020), pp. 2.4 - 2.7
Savannah River Lab	Reed et al. (2013), pg. 27

Table A1: Lab names and document titles/links of their respective runner-up sources

A.2 Runner-Up Locations



(a) Labs in the eastern states

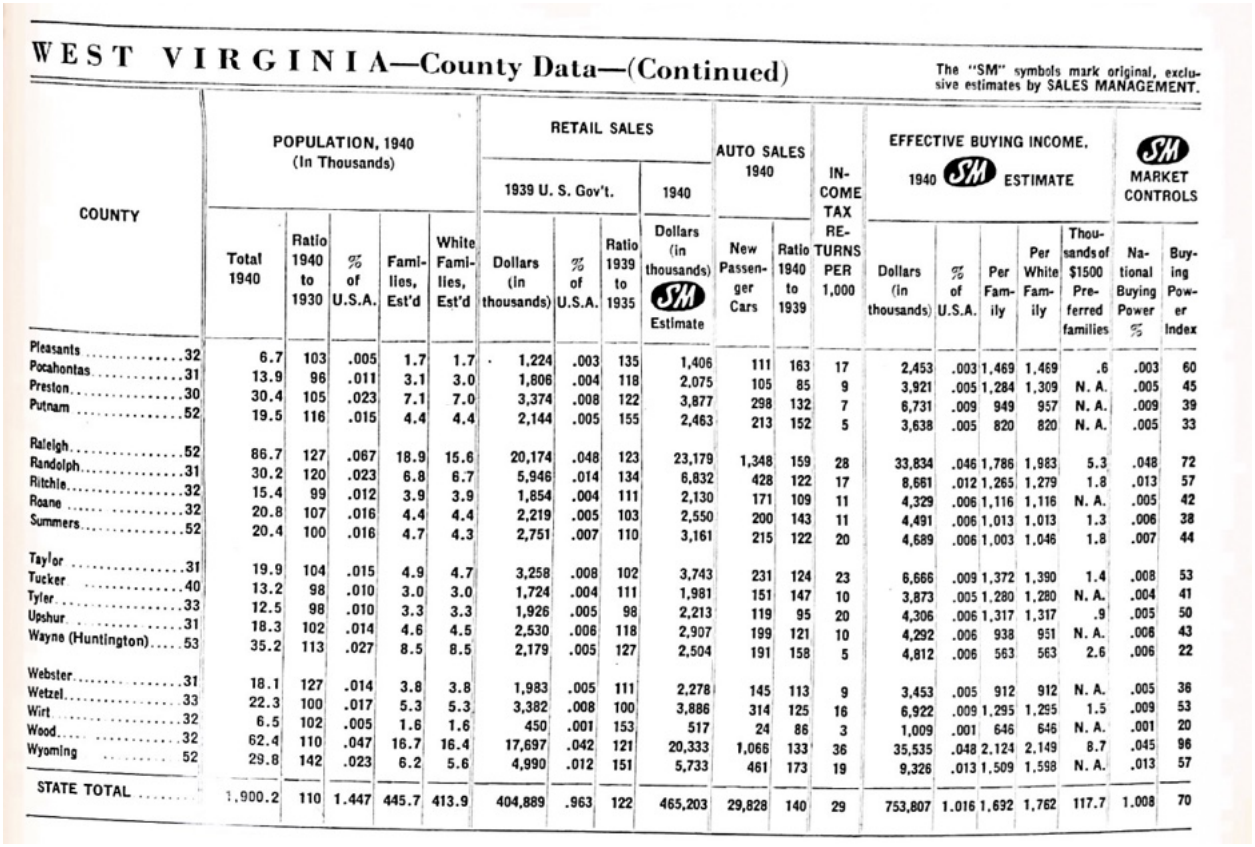


(b) Labs in the western states

Figure A.1: Map of lab counties (blue) and runner-up counties (red, hashed). Los Alamos county in New Mexico was carved out of Sandoval county in 1949, and so for all our analysis we merge them into one.

A.3 Novel Data Collection

Figure A.2: A scan from the 1941 issue of the Sales and Management's Survey of Buying Power



WEST VIRGINIA—County Data—(Continued)

The "SM" symbols mark original, exclusive estimates by SALES MANAGEMENT.

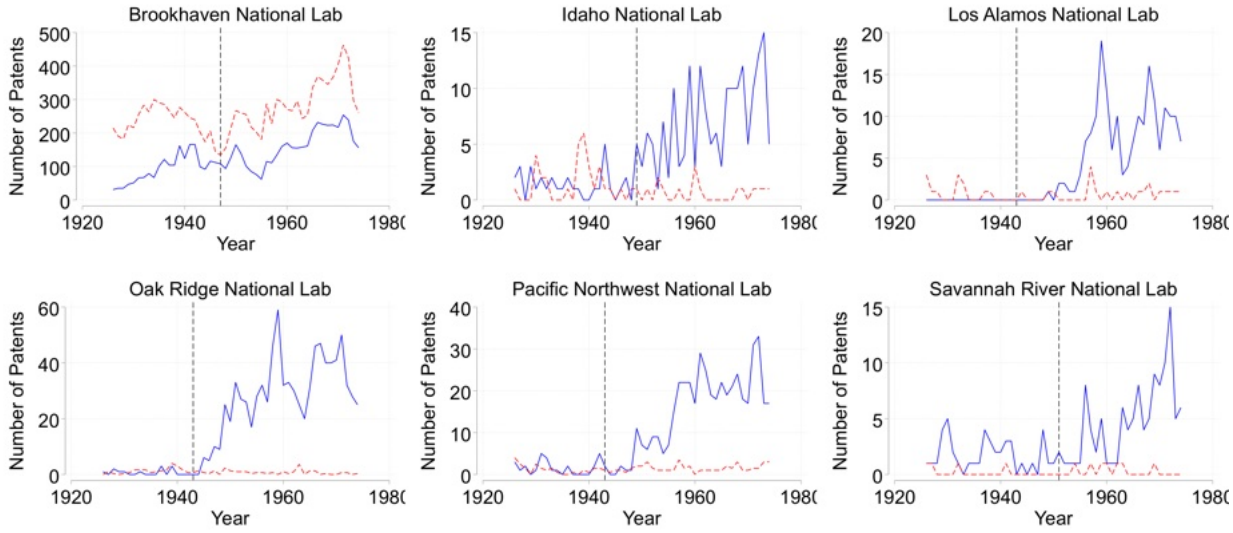
COUNTY	POPULATION, 1940 (In Thousands)					RETAIL SALES				AUTO SALES 1940		INCOME TAX RETURNS PER 1,000	EFFECTIVE BUYING INCOME, 1940 SM ESTIMATE					SM MARKET CONTROLS	
						1939 U. S. Gov't.			1940										
	Total 1940	Ratio 1940 to 1930	% of U.S.A.	Families, Est'd	White Families, Est'd	Dollars (in thousands)	% of U.S.A.	Ratio 1939 to 1935	Dollars (in thousands) SM Estimate	New Passenger Cars	Ratio 1940 to 1939		Dollars (in thousands)	% of U.S.A.	Per Family	Per White Family	Thousands of $1500 Preferred families	National Buying Power %	Buying Power Index
Pleasants 32	6.7	103	.005	1.7	1.7	1,224	.003	135	1,406	111	163	17	2,453	.003	1,469	1,469	.6	.003	60
Pocahontas 31	13.9	96	.011	3.1	3.0	1,806	.004	118	2,075	105	85	9	3,921	.005	1,284	1,309	N. A.	.005	45
Preston 30	30.4	105	.023	7.1	7.0	3,374	.008	122	3,877	298	132	7	6,731	.009	949	957	N. A.	.009	39
Putnam 52	19.5	116	.015	4.4	4.4	2,144	.005	155	2,463	213	152	5	3,638	.005	820	820	N. A.	.005	33
Raleigh 52	86.7	127	.067	18.9	15.6	20,174	.048	123	23,179	1,348	159	28	33,834	.046	1,786	1,983	5.3	.048	72
Randolph 31	30.2	120	.023	6.8	6.7	5,946	.014	134	6,832	428	122	17	8,661	.012	1,265	1,279	1.8	.013	57
Ritchie 32	15.4	99	.012	3.9	3.9	1,854	.004	111	2,130	171	109	11	4,329	.006	1,116	1,116	N. A.	.005	42
Roane 32	20.8	107	.016	4.4	4.4	2,219	.005	103	2,550	200	143	11	4,491	.006	1,013	1,013	1.3	.006	38
Summers 52	20.4	100	.016	4.7	4.3	2,751	.007	110	3,161	215	122	20	4,689	.006	1,003	1,046	1.8	.007	44
Taylor 31	19.9	104	.015	4.9	4.7	3,258	.008	102	3,743	231	124	23	6,666	.009	1,372	1,390	1.4	.008	53
Tucker 40	13.2	98	.010	3.0	3.0	1,724	.004	111	1,981	151	147	10	3,873	.005	1,280	1,280	N. A.	.004	41
Tyler 33	12.5	98	.010	3.3	3.3	1,926	.005	98	2,213	119	95	20	4,306	.006	1,317	1,317	.9	.005	50
Upshur 31	18.3	102	.014	4.6	4.5	2,530	.006	118	2,907	199	121	10	4,292	.006	938	951	N. A.	.006	43
Wayne (Huntington) 53	35.2	113	.027	8.5	8.5	2,179	.005	127	2,504	191	158	5	4,812	.006	563	563	2.6	.006	22
Webster 31	18.1	127	.014	3.8	3.8	1,983	.005	111	2,278	145	113	9	3,453	.005	912	912	N. A.	.005	36
Wetzel 33	22.3	100	.017	5.3	5.3	3,382	.008	100	3,886	314	125	16	6,922	.009	1,295	1,295	1.5	.009	53
Wirt 32	6.5	102	.005	1.6	1.6	450	.001	153	517	24	86	3	1,009	.001	646	646	N. A.	.001	20
Wood 32	62.4	110	.047	16.7	16.4	17,697	.042	121	20,333	1,066	133	36	35,535	.048	2,124	2,149	8.7	.045	96
Wyoming 52	29.8	142	.023	6.2	5.6	4,990	.012	151	5,733	461	173	19	9,326	.013	1,509	1,598	N. A.	.013	57
STATE TOTAL	1,900.2	110	1.447	445.7	413.9	404,889	.963	122	465,203	29,828	140	29	753,807	1.016	1,692	1,762	117.7	1.008	70

A.4 Balance Checks

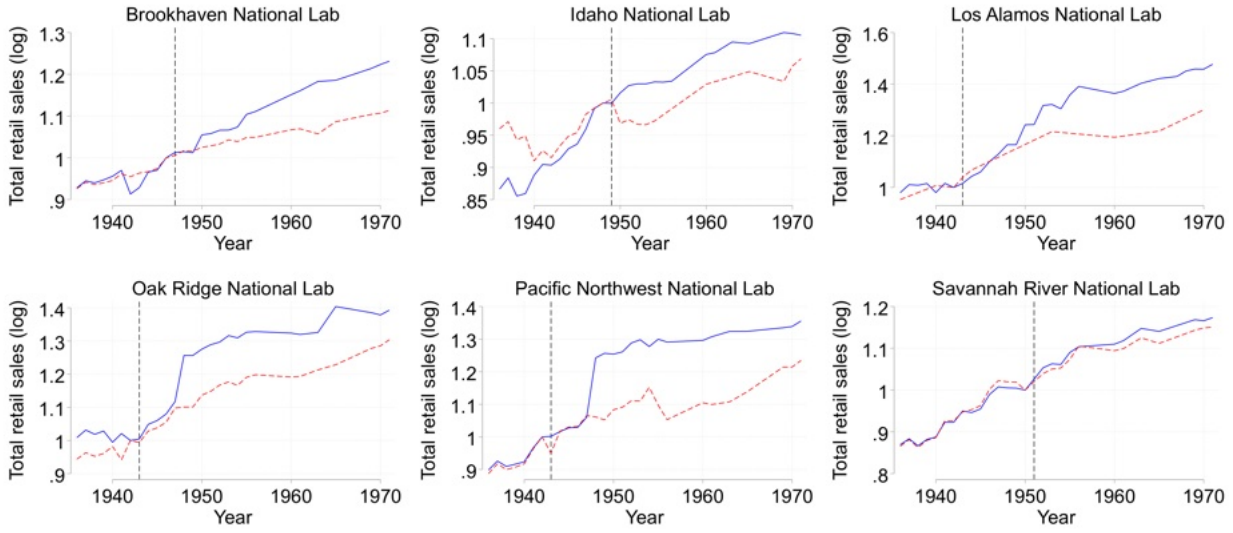
Table A2: Balance check. Data from the decennial census (IPUMS) aggregated to the county level. The sample consists of the following labs: Ames National Lab, Argonne National Lab, Berkeley Lab, Brookhaven, Idaho, Los Alamos, Oak Ridge, PNNL, Princeton, Plasma Savannah, SLAC, Sandia. Educational attainment is measured as in the IPUMS Database (ranges from 0 to 20). Share of farmers is total farmers divided by total employed.

Variables	**Treated**	**Synthetic Control**	**p-val**
Age (1940)	29.7	29.8	0.940
Log population (1940)	11.03	11.15	0.821
Labor Force Participation Rate % (1940)	52.8	52.2	0.739
Educational attainment (1940)	9.8	9.6	0.810
Wages (current $) (1940)	318	274	0.461
Share of farmers % (1940)	7.5	8.7	0.571
Share of manufacturing workers % (1940)	6.3	8.0	0.468
Share of whites % (1940)	91.7	92.9	0.967
Share of foreign-born % (1940)	14.6	15.0	0.952

A.5 Time Series Plots of Patents and Retail Sales using Runner-Up Locations as Controls



(a)



(b)

Figure A.3: Time series of patents (a) and log aggregate retail sales (b). Blue solid lines indicate treated counties and the red dashed are the runner-up ones. For retail sales the values are normalized to equal 1 at treatment start year.

A.6 Inference: P-Values from the Synthetic Control Plots of Tables 2 and 6

Table A3: P-values of synthetic control regressions. N/A for Berkeley lab since the retail sales series begins three years after its founding.

Lab	Patents (p-val)	Retail Sales (p-val)
Ames	0.030**	0.316
Argonne	0.009***	0.031**
Berkeley	0.087*	N/A
Brookhaven	0.021**	0.033**
Idaho	0.023**	0.756
Los Alamos	0.033**	0.000***
Oak Ridge	0.010***	0.021**
PNNL	0.052*	0.026**
Princeton Plasma	0.300	0.850
Savannah	0.333	0.111
SLAC	0.072*	0.407
Sandia	0.000***	0.035**

A.7 Time Series of Retail Sales of Suburban Counties of the Chicago Metro Area

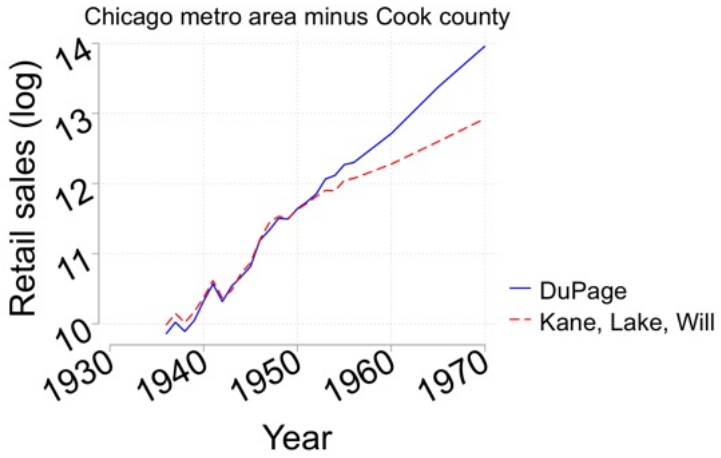


Figure A.4: Retail sales for counties adjacent to the Chicago metro area. Argonne Lab located in DuPage, IL. The red line is normalized so that it equals the blue line in 1940.

A.8 Budget Correlates of Brookhaven Lab in Suffolk County, NY

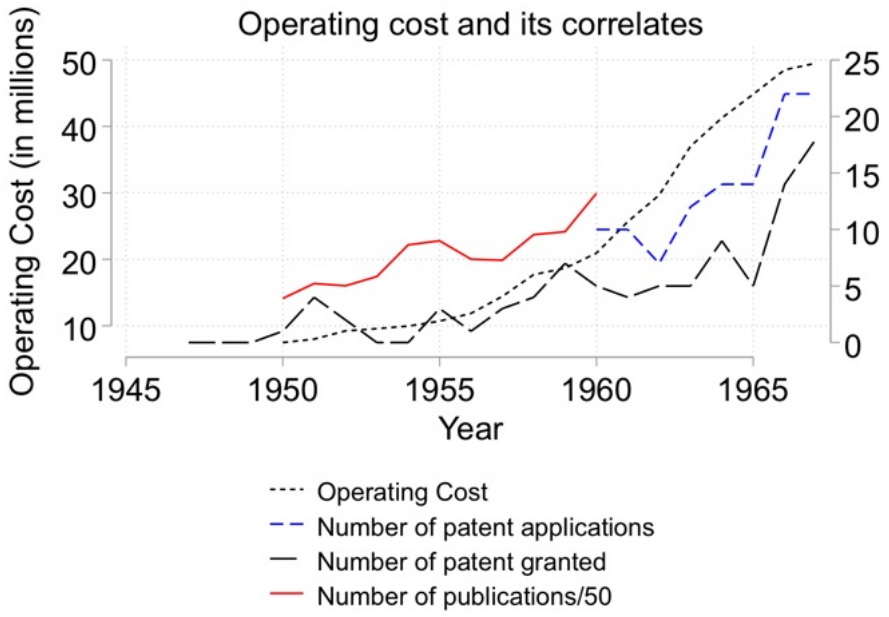
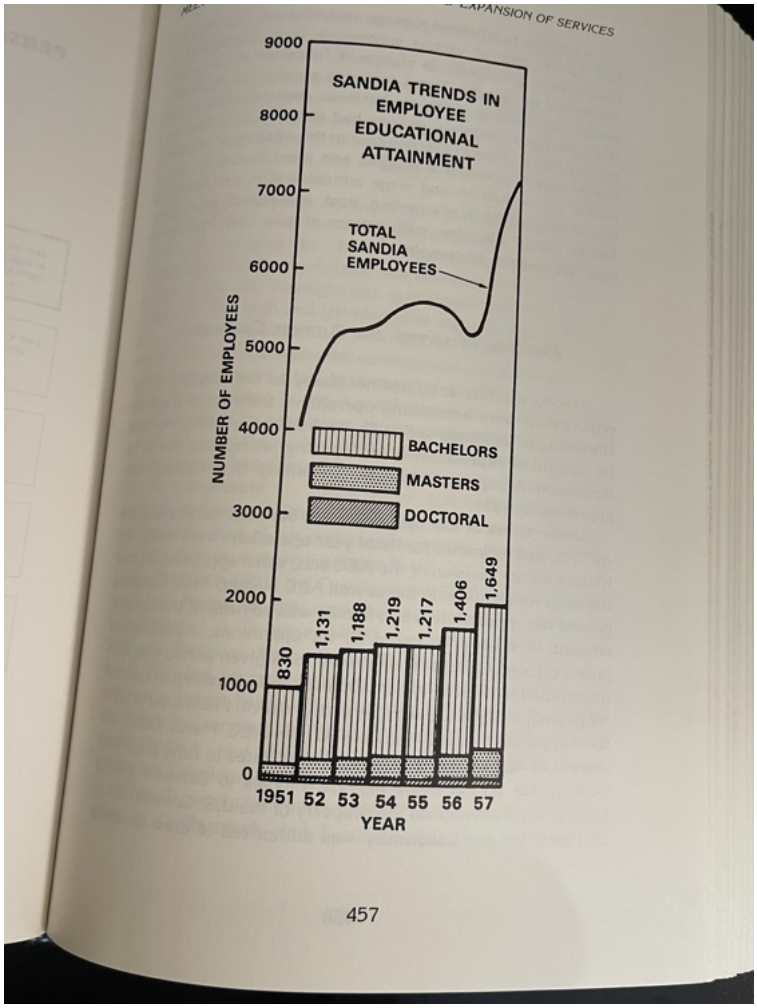


Figure A.5: Lab budget correlates over time for Brookhaven Lab in Suffolk county, NY

A.9 Trends in Sandia Lab's Employees by Education

Figure A.6: Taken from Furman (1989)



A.10 Regression Table for the Plots from Figure 5

$$\text{ipc code}^{J}_{ict} = \alpha + \beta_1 \times \text{Treated}_c + \beta_2 \times \text{Post}_t + \beta_3(\text{Treated}_c \times \text{Post}_t) + \epsilon_{ict}$$

Table A4: Diff-in-diff results of ipc-code changes in treated counties relative to runner-up counties

VARIABLES	(1) Physics (42.7%)	(2) Chemistry & Metallurgy (24.4%)	(3) Electricity (15.2%)	(4) Performing Operations (12.7%)	(5) Engineering (3.4%)	(6) Human Necessities (1%)	(7) Construction (0.4%)	(8) Textiles & Paper (0.3%)
$\text{Treated}_c \times \text{Post}_t$	0.109***	0.084**	0.090***	-0.044	0.092*	-0.195***	-0.109***	-0.027
	(0.033)	(0.035)	(0.029)	(0.060)	(0.049)	(0.057)	(0.040)	(0.055)
Observations	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401
R-squared	0.051	0.043	0.012	0.009	0.017	0.017	0.028	0.039

Notes: The unit of observation is a patent. Each outcome variable is a dummy variable that indicates if the patent belong to the outcome IPC category. Post equals 1 if time period of patent grant is after lab's founding in the respective county and 0 otherwise. Percentages in parenthesis below the outcome variable names indicate the share of lab patents in each outcome category. Standard errors are two-way clustered at the county year level. *** p<0.01, ** p<0.05, * p<0.1

A.11 County Map of Non-Lab Patents Citing Lab Patents

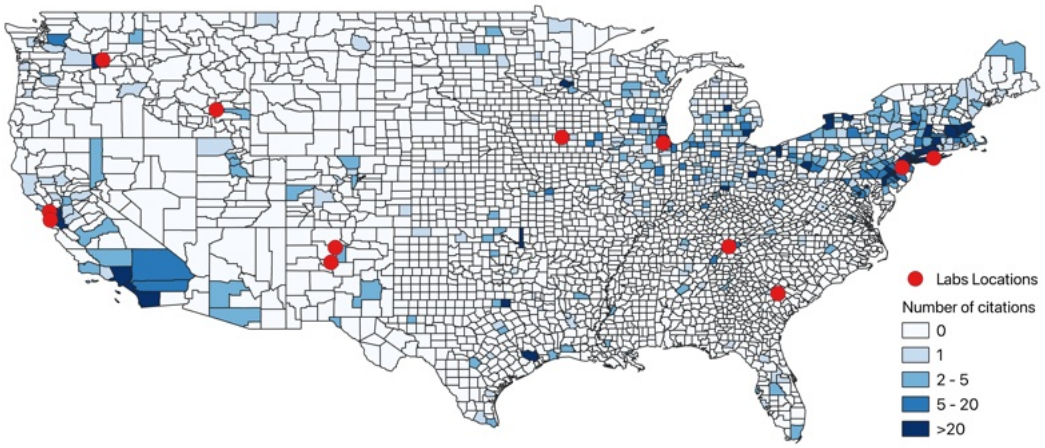
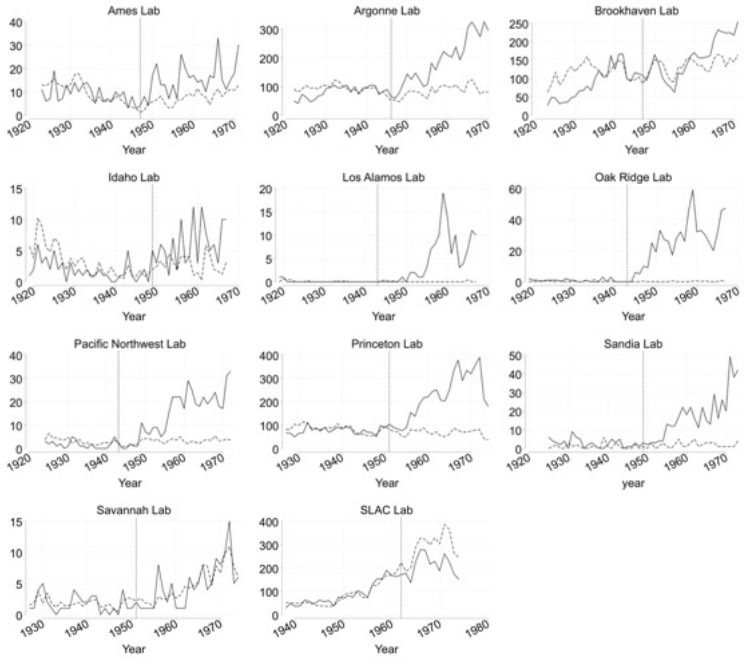
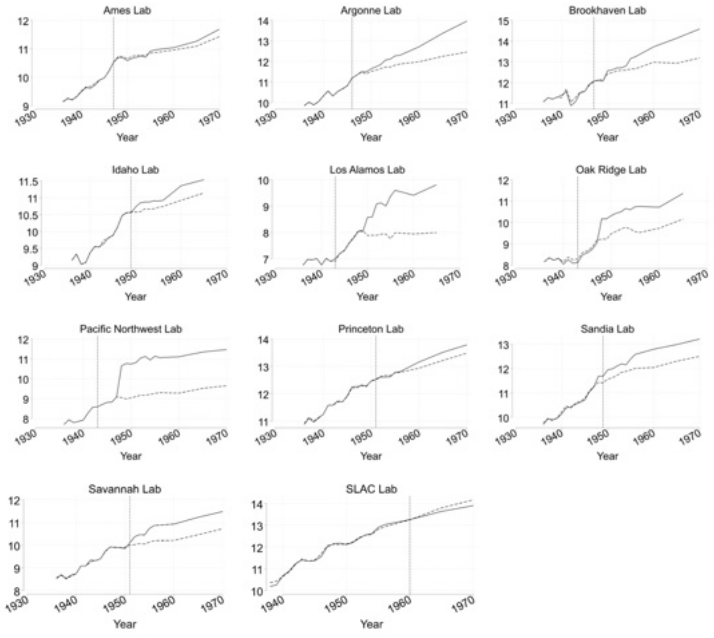


Figure A.7: This sample excludes patents assigned to the labs. Red dots indicate lab locations.

A.12 Robustness to Using a Common Set of Predictor Variables



(a) On the y-axis is the number of patents.



(b) On the y-axis is the log of total retail sales

Figure A.8: Synthetic control plots using a common set of predictors for both patents and retails sales.